# SOME REFINEMENTS OF SHANNON'S INEQUALITIES

M. MATIĆ<sup>1</sup>, C. E. M. PEARCE<sup>2</sup> and J. PEČARIĆ<sup>2</sup>

(Received 12 May, 1999)

#### Abstract

We refine Shannon's inequality, in its discrete and integral forms, by presenting upper estimates of the difference between its two sides.

## 1. Introduction

A fundamental result related to the notion of the Shannon entropy is the inequality

$$\sum_{i=1}^{n} p_i \log \frac{1}{p_i} \le \sum_{i=1}^{n} p_i \log \frac{1}{q_i},$$
(1.1)

which is valid for all positive real numbers  $p_i$  and  $q_i$  with  $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$ . Here as subsequently 'log' refers to logarithms to base b for some fixed b > 1. Equality holds in (1.1) if and only if  $q_i = p_i$  for all i. For details see [8, pp. 635– 650]. This result, sometimes called the fundamental lemma of information theory, has extensive applications (see, for example, [7]).

The following two theorems (see [7, pp. 278–279]) extend (1.1) and we can call them Shannon's discrete and integral inequalities respectively.

THEOREM 1. Let I be a finite or countable set of integers and  $\{p_i, i \in I\}$  a set of positive real numbers with  $\sum_{i \in I} p_i = 1$ . If  $\{q_i, i \in I\}$  is a set of nonnegative real numbers with  $\sum_{i \in I} q_i = \alpha > 0$ , then

$$\sum_{i \in I} p_i \log \frac{1}{p_i} \le \sum_{i \in I} p_i \log \frac{1}{q_i} + \log \alpha, \qquad (1.2)$$

with equality if and only if  $q_i = \alpha p_i$  for all  $i \in I$ .

<sup>&</sup>lt;sup>1</sup>FESB, Mathematics Department, University of Split, R. Boškovića bb, 21000 Split, Croatia.

<sup>&</sup>lt;sup>2</sup>Applied Mathematics Department, University of Adelaide, Adelaide SA 5005, Australia.

<sup>©</sup> Australian Mathematical Society 2002, Serial-fee code 1446-1811/02

THEOREM 2. Let I be a measurable subset of the real line and let p(x) be a positive integrable function defined on I with  $\int_{I} p(x) dx = 1$ . If q(x) is a nonnegative integrable function defined on I with  $\int_{I} q(x) dx = \alpha > 0$ , then

$$\int_{I} p(x) \log \frac{1}{p(x)} dx \le \int_{I} p(x) \log \frac{1}{q(x)} dx + \log \alpha, \tag{1.3}$$

with equality if and only if  $q(x) = \alpha p(x)$  a.e. on I.

We can allow  $p_i$  in (1.2) to be nonnegative with the usual convention  $0 \log 0 = 0$ . Also if  $p_i > 0$  and  $q_i = 0$  for some  $i \in I$ , then (1.2) is uninteresting, since the right-hand side is infinite in that case. Similarly, if the Lebesgue measure of the set  $\{x \in I : q(x) = 0\}$  is positive, then the right-hand side of (1.3) is infinite. We shall exclude such cases from consideration. Also, since  $\sum_{i \in I} p_i = 1$  implies that  $0 \le p_i \le 1$  for all  $i \in I$ , we have  $S_p := \sum_{i \in I} p_i \log(1/p_i) \ge 0$ . So if  $S_q := \sum_{i \in I} p_i \log(1/q_i)$  is finite, then (1.2) implies that  $S_p$  is finite too. On the other hand, the finiteness of  $\int_I p(x) \log(1/q(x)) dx$  in (1.3) doesn't imply that of  $\int_I p(x) \log(1/p(x)) dx$ . For example, take b = e, I = (0, 1/e),  $p(x) = 1/(x \ln^2 x)$  and q(x) = e for  $x \in I$ . This gives  $\alpha = 1$ ,  $\int_I p(x) dx = 1$  and  $\int_I p(x) \ln(1/q(x)) dx = -1$ , but  $\int_I p(x) \ln(1/p(x)) dx = -\infty$ .

Our underlying motivation is to estimate the difference between the two sides of the relevant Shannon inequality. This entails a discussion of whether the appropriate generalization of the entropy (the left-hand side of (1.2) or (1.3)) converges. In the discrete case a key role is played by the existence and value of  $v := \sum_{i} ip_{i}$ .

Suppose a discrete-valued random variable assumes value  $x_i$  with probability  $p_i$   $(i \in I)$ . Its entropy is independent of the values  $x_i$   $(i \in I)$  provided that they are all distinct. Suppose we form equivalence classes of such random variables according to the sets  $\{p_i; i \in I\}$ . We may select a canonical representative from each class for which  $x_i = i$   $(i \in I)$ . The parameter  $\nu$  can then be interpreted as the mean of that representative.

In the integral case the identities of the values assumed by a random variable are not lost, and v is the same as the mean  $\mu$  of the probability distribution with density function  $p(\cdot)$ .

In Theorem 1 we must distinguish between the case when I is finite (when without loss in generality we can take  $I = \{1, 2, ..., n\}$  for some integer  $n \ge 2$ ) and the case when I is countable (when we may take I = N).

To obtain counterparts of some inequalities in information theory in the case when I is finite, the following result was used in [3].

Let  $\xi_k \in (0, \infty)$  and  $p_k \ge 0$  with  $\sum_{k=1}^n p_k = 1$ . We then have

$$0 \le \log\left(\sum_{k=1}^{n} p_k \xi_k\right) - \sum_{k=1}^{n} p_k \log \xi_k \le \frac{1}{\ln b} \left[\sum_{k=1}^{n} \frac{p_k}{\xi_k} \sum_{j=1}^{n} p_j \xi_j - 1\right].$$
(1.4)

Shannon's inequalities

To prove counterparts of (1.2) (in the case when *I* is countable) and of (1.3), we need a better and more general result than (1.4). This we prove in Section 2. We proceed to counterparts of (1.2) and (1.3) in Sections 3 and 4.

## 2. Preliminary results

We begin by establishing an upper bound estimate for the difference between the two sides of the well-known Jensen's inequality. In integral version, Jensen's inequality states the following (see for example [11, Theorem 3.3]).

If  $(\Omega, \mathscr{A}, \mu)$  is a probability space and f a real function in  $L^{1}(\mu)$  with a < f(x) < b for all  $x \in \Omega$   $(-\infty \le a < b \le \infty)$ , then for any function  $\varphi$  defined and convex on (a, b) we have

$$\varphi\left(\int_{\Omega} f \, d\mu\right) \leq \int_{\Omega} \varphi(f) d\mu. \tag{2.1}$$

In the case when  $\varphi$  is strictly convex on (a, b), equality holds if and only if f is constant a.e. on  $\Omega$ .

For our purposes we need the following result.

THEOREM 3. Let  $(\Omega, \mathscr{A}, \mu)$  be a probability space. If  $f : \Omega \to (0, \infty)$  is such that f and 1/f are in  $L^1(\mu)$ , then we have

$$0 \leq \log\left(\int_{\Omega} f \ d\mu\right) - \int_{\Omega} \log f \ d\mu$$
$$\leq \log\left(\int_{\Omega} f \ d\mu \int_{\Omega} \frac{1}{f} \ d\mu\right) \leq \frac{1}{\ln b} \left[\int_{\Omega} f \ d\mu \int_{\Omega} \frac{1}{f} \ d\mu - 1\right].$$
(2.2)

Moreover, equality holds throughout if and only if f is constant a.e. on  $\Omega$ .

PROOF. If f is a positive function defined on  $\Omega$  such that f and 1/f are both in  $L^{1}(\mu)$ , then this is true for  $\varphi(f) = -\log f$  as well, since

$$0 < -\log f < \frac{1}{\ln b} \frac{1}{f}$$
 on  $\{x \in \Omega : 0 < f(x) < 1\}$ 

and

$$0 \le \log f < \frac{1}{\ln b} f \quad \text{on } \{x \in \Omega : 1 \le f(x)\}.$$

The first inequality in (2.2) is, in fact, Jensen's inequality for the convex function  $\varphi(t) = -\log t$ . For f replaced by 1/f, this inequality becomes

$$\log\left(\int_{\Omega}\frac{1}{f}\,d\mu\right)\geq\int_{\Omega}\log\frac{1}{f}\,d\mu=-\int_{\Omega}\log f\,d\mu.$$

Thus we have

$$\log\left(\int_{\Omega} f \ d\mu\right) - \int_{\Omega} \log f \ d\mu \leq \log\left(\int_{\Omega} f \ d\mu\right) + \log\left(\int_{\Omega} \frac{1}{f} \ d\mu\right)$$

which is the second inequality in (2.2). The third inequality is a simple consequence of the elementary result

$$\log x \le \frac{1}{\ln b}(x-1) \quad (\forall x > 0).$$

Dragomir and Goh [3] have used their version of Jensen's inequality in the proof of the second inequality in (1.4). In our proof we have used only Jensen's inequality for the second inequality in (2.2), and the above elementary inequality in the proof of the last part of (2.2).

We have applications for the following two special cases of Theorem 3.

COROLLARY 1. Let I be a finite or countable set of integers and  $\{p_i, i \in I\}$  a set of positive real numbers with  $\sum_{i \in I} p_i = 1$ . For  $\xi_i \in (0, \infty)$   $(i \in I)$  such that  $\sum_{i \in I} p_i \xi_i < \infty$  and  $\sum_{i \in I} p_i / \xi_i < \infty$ , we have

$$0 \le \log\left(\sum_{i \in I} p_i \xi_i\right) - \sum_{i \in I} p_i \log \xi_i$$
  
$$\le \log\left[\sum_{i \in I} p_i \xi_i \sum_{i \in I} \frac{p_i}{\xi_i}\right] \le \frac{1}{\ln b}\left[\sum_{i \in I} p_i \xi_i \sum_{i \in I} \frac{p_i}{\xi_i} - 1\right],$$

with equality throughout if and only if  $\xi_i = c$  for all  $i \in I$ .

**PROOF.** Set  $\Omega = I$ ,  $\mu(\{i\}) = p_i$  and  $f(i) = \xi_i$  for  $i \in I$  and then apply Theorem 3.

COROLLARY 2. Let I be a measurable subset of the real line and p(x) a positive integrable function defined on I with  $\int_I p(x) dx = 1$ . If  $\xi(x)$  is a positive integrable function defined on I and such that  $\int_I p(x)\xi(x) dx < \infty$  and  $\int_I (p(x)/\xi(x)) dx < \infty$ , then we have

$$0 \le \log\left(\int_{I} p(x)\xi(x) \, dx\right) - \int_{I} p(x) \log \xi(x) \, dx$$
  
$$\le \log\left[\int_{I} p(x)\xi(x) \, dx \int_{I} \frac{p(x)}{\xi(x)} \, dx\right] \le \frac{1}{\ln b} \left[\int_{I} p(x)\xi(x) \, dx \int_{I} \frac{p(x)}{\xi(x)} \, dx - 1\right],$$

with equality throughout if and only if  $\xi(x) = c$  a.e. on I.

PROOF. Set  $\Omega = I$ ,  $\mathscr{A} =$  the  $\sigma$ -algebra of measurable subsets of I and  $\mu(A) = \int_A p(x) dx$  for  $A \in \mathscr{A}$ . Now apply Theorem 3 with  $f = \xi$ .

496

Shannon's inequalities

The following two lemmas give discrete and integral versions of the well-known Grüss inequality. For an account of the Grüss inequality see, for example, [8, pp. 295-310].

LEMMA 1. Let I be a finite or countable set of integers and  $\{p_i, i \in I\}$  a set of positive real numbers such that  $P := \sum_{i \in I} p_i < \infty$ . If  $\{a_i, i \in I\}$  and  $\{b_i, i \in I\}$  are bounded sets of real numbers satisfying

$$\alpha \leq a_i \leq A, \quad \beta \leq b_i \leq B \quad \text{for all } i \in I,$$

then

$$\left|P\sum_{i\in I}p_ia_ib_i-\sum_{i\in I}p_ia_i\sum_{i\in I}p_ib_i\right|\leq \frac{1}{4}(A-\alpha)(B-\beta)P^2.$$

LEMMA 2. Let I be a measurable subset of the real line and p(x) a positive integrable function on I such that  $P =: \int_{I} p(x) dx < \infty$ . If f(x) and g(x) are bounded measurable functions on I with

$$\alpha \leq f(x) \leq A$$
,  $\beta \leq g(x) \leq B$  a.e. on I,

then

$$\left|P\int_{I}p(x)f(x)g(x)\,dx-\int_{I}p(x)f(x)\,dx\int_{I}p(x)g(x)\,dx\right|\leq\frac{1}{4}(A-\alpha)(B-\beta)P^{2}.$$

Lemmas 1 and 2 can easily be proved directly, following the lines of the standard proof for the classical Grüss result. They are also direct consequences [2, Proposition 3.2]. We omit the details.

We need also the following counterpart of the weighted Cauchy inequality (see [8, p. 125]).

LEMMA 3. Let  $(a_i)$ ,  $(b_i)$  and  $(w_i)$  be positive n-tuples such that

$$0 < m \leq a_i/b_i \leq M, \quad (i = 1, 2, ..., n).$$

Then

$$\sum_{i=1}^{n} w_i a_i^2 \sum_{j=1}^{n} w_j b_j^2 - \left(\sum_{i=1}^{n} w_i a_i b_i\right)^2 \le \left(\sqrt{M} - \sqrt{m}\right)^2 \sum_{i=1}^{n} w_i a_i b_i \sum_{j=1}^{n} w_j b_j^2.$$
(2.3)

When  $n \to \infty$ , (2.3) still holds if all the sums involved converge. Hence (2.3) holds with  $\sum_{i \in I}$  in place of  $\sum_{i=1}^{n}$  for any finite or countable discrete set *I*.

## 3. A counterpart of the discrete Shannon inequality

First we prove our main result in this section.

THEOREM 4. Let I be a finite or countable set of integers and  $\{p_i, i \in I\}$  and  $\{q_i, i \in I\}$  sets of positive numbers such that  $\sum_{i \in I} p_i = 1$  and  $\alpha := \sum_{i \in I} q_i < \infty$ . Suppose  $S_q$  is finite. Then  $S_p$  is finite and we have  $0 < S_p \leq S_q + \log \alpha$ . If additionally  $\sum_{i \in I} p_i^2/q_i < \infty$ , then

$$0 \le S_q - S_p + \log \alpha \le \log \left[ \alpha \sum_{i \in I} \frac{p_i^2}{q_i} \right] \le \frac{1}{\ln b} \left[ \alpha \sum_{i \in I} \frac{p_i^2}{q_i} - 1 \right], \quad (3.1)$$

with equality throughout if and only if  $q_i = \alpha p_i$  for all  $i \in I$ .

PROOF. Since  $\sum_{i \in I} p_i = 1$  we have that  $p_i < 1$  and so  $\log(1/p_i) > 0$  for all  $i \in I$ . This implies  $0 < S_p \le \infty$ . Similarly, since  $\sum_{i \in I} q_i < \infty$ , we have that  $q_i < 1$  and  $\log(1/q_i) > 0$  for all but finitely many  $i \in I$ , which implies that  $-\infty < S_q \le \infty$ . As remarked in the introduction,  $S_q$  finite implies  $S_p$  finite, so our first assertion is a consequence of Theorem 1.

To prove the second assertion, set  $\xi_i = q_i/p_i$   $(i \in I)$ . We have  $\sum_{i \in I} p_i \xi_i = \alpha < \infty$ and  $\sum_{i \in I} p_i/\xi_i = \sum_{i \in I} p_i^2/q_i < \infty$  and we can apply Corollary 1 to obtain

$$0 \le \log \alpha - \sum_{i \in I} p_i \log \frac{q_i}{p_i} \le \log \left[ \alpha \sum_{i \in I} \frac{p_i^2}{q_i} \right] \le \frac{1}{\ln b} \left[ \alpha \sum_{i \in I} \frac{p_i^2}{q_i} - 1 \right]$$

Since  $S := \sum_{i \in I} p_i \log(q_i/p_i)$  is finite and  $S_q$  is finite, we have that  $S_q + S = S_p$ is finite, so we can write  $S = S_p - S_q$  and (3.1) is proved. Moreover, by Corollary 1, we know that equality holds throughout if and only if  $\xi_i = c$ , that is, if  $q_i = cp_i$  for all  $i \in I$ . Since  $\sum_{i \in I} q_i = \alpha$  and  $\sum_{i \in I} p_i = 1$  we have  $c = \alpha$ .

The assumptions about the finiteness of the sums involved can be omitted when I is finite.

**THEOREM 5.** Suppose the assumptions of Theorem 4 are satisfied and

$$0 < m \leq p_i/q_i \leq M$$
 for all  $i \in I$ .

Then

$$0 \le \sum_{i \in I} p_i \log \frac{1}{q_i} - \sum_{i \in I} p_i \log \frac{1}{p_i} + \log \alpha \le \log \frac{(M+m)^2}{4Mm} \le \frac{1}{4 \ln b} \frac{(M-m)^2}{Mm}.$$

**PROOF.** Because  $1/M \le q_i/p_i \le 1/m$  for all  $i \in I$  we may apply Lemma 1 with  $a_i = q_i/p_i$ ,  $b_i = p_i/q_i$  for  $i \in I$  and P = 1 to obtain

$$\alpha \sum_{i \in I} \frac{p_i^2}{q_i} - 1 = \sum_{i \in I} p_i \frac{q_i}{p_i} \sum_{j \in I} p_j \frac{p_j}{q_j} - \sum_{i \in I} p_i \frac{q_i}{p_i} \frac{p_i}{q_i} \le \frac{1}{4} \left( \frac{1}{m} - \frac{1}{M} \right) (M - m).$$

Therefore we have

$$\alpha \sum_{i \in I} \frac{p_i^2}{q_i} \le \frac{(M-m)^2}{4Mm} + 1 = \frac{(M+m)^2}{4Mm}.$$

The desired result follows by Theorem 4.

THEOREM 6. Under the assumptions of Theorem 5 we have

$$0 \leq \sum_{i \in I} p_i \log \frac{1}{q_i} - \sum_{i \in I} p_i \log \frac{1}{p_i} + \log \alpha$$
$$\leq \log \left[ \alpha \left( \sqrt{M} - \sqrt{m} \right)^2 + 1 \right] \leq \frac{\alpha}{\ln b} \left( \sqrt{M} - \sqrt{m} \right)^2.$$

**PROOF.** Set  $w_i = 1/q_i$ ,  $a_i = p_i$  and  $b_i = q_i$  for  $i \in I$ . Now apply Lemma 3 (with  $\sum_{i \in I}$  in place of  $\sum_{i=1}^{n}$ ) to obtain

$$\alpha \sum_{i \in I} \frac{p_i^2}{q_i} - 1 = \sum_{i \in I} \frac{1}{q_i} q_i^2 \sum_{j \in I} \frac{1}{q_j} p_j^2 - \left( \sum_{i \in I} \frac{1}{q_i} p_i q_i \right)^2$$
  
$$\leq \left( \sqrt{M} - \sqrt{m} \right)^2 \sum_{i \in I} \frac{1}{q_i} p_i q_i \sum_{j \in I} \frac{1}{q_j} q_j^2 = \alpha \left( \sqrt{M} - \sqrt{m} \right)^2.$$

Therefore we have

$$\alpha \sum_{i=1}^{n} \frac{p_i^2}{q_i} \leq \alpha \left(\sqrt{M} - \sqrt{m}\right)^2 + 1.$$

The desired result follows by (3.1).

If X is a discrete random variable with finite range  $\{x_1, x_2, ..., x_n\}$  and probability distribution  $p_i := P\{X = x_i\} \ge 0$   $(\sum_{i=1}^n p_i = 1)$ , the entropy of X is defined by

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

with the usual convention  $0 \log 0 = 0$ . We can apply Theorem 4 with  $q_i = 1/n$  (i = 1, 2, ..., n) to obtain

$$0 \le \log n - H_b(X) \le \log \left[ n \sum_{i=1}^n p_i^2 \right] \le \frac{1}{\ln b} \left[ n \sum_{i=1}^n p_i^2 - 1 \right].$$
(3.2)

Equality holds throughout if and only if  $p_i = 1/n$  for all *i*. The above result is an improvement of some results from [3, 4, 6].

[7]

The first inequality in (3.2), that is,

$$H_b(X) \le \log n, \tag{3.3}$$

shows that the entropy function  $H_b(X)$  achieves its maximum value on the discrete uniform probability distribution. The following generalization of this inequality is a simple consequence of a well-known majorization theorem ([10, pp. 319–320]).

Suppose  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{y} = (y_1, \dots, y_n)$  and let  $x_{[1]} \ge x_{[2]} \ge \dots \ge x_{[n]}$  and  $y_{[1]} \ge y_{[2]} \ge \dots \ge y_{[n]}$  be their ordered components. The *n*-tuple  $\mathbf{y}$  is said to majorize  $\mathbf{x}$  (or  $\mathbf{x}$  is majorized by  $\mathbf{y}$ ) and we write  $\mathbf{y} \succ \mathbf{x}$  (or  $\mathbf{x} \prec \mathbf{y}$ ) if

$$\sum_{i=1}^{m} x_{[i]} \le \sum_{i=1}^{m} y_{[i]} \text{ holds for } m = 1, \dots, n-1, \text{ and } \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i.$$

Let I be an interval in **R** and **x**, **y** two *n*-tuples such that  $x_i, y_i \in I$  (i = 1, ..., n). The majorization theorem states that

$$\sum_{i=1}^n \varphi(x_i) \leq \sum_{i=1}^n \varphi(y_i)$$

holds for every continuous convex function  $\varphi : I \to \mathbf{R}$  if and only if  $\mathbf{y} \succ \mathbf{x}$ .

Suppose X and Y are discrete random variables with finite ranges and respective probability distributions  $\mathbf{p} = \{p_i\}_{i=1}^n$  and  $\mathbf{q} = \{q_i\}_{i=1}^n \left(\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1\right)$  such that  $\mathbf{p} \succ \mathbf{q}$ . Since the function  $\varphi(x) = x \log x$  is convex,  $-x \log x$  is concave and we have the reversed inequality

$$H_b(X) \le H_b(Y). \tag{3.4}$$

It is well-known (and easily established) that  $\mathbf{p} \succ (1/n, ..., 1/n)$ , so the last inequality gives (3.3).

An extension of (3.4) was recently considered in [1]. The incomplete Rényi entropy was defined by

$$H_b^k(\mathbf{p}) = \frac{\sum_{i=1}^k p_i \log 1/p_i}{\sum_{i=1}^k p_i}, \quad k = 1, ..., n,$$

and the inequality

$$H_b^k(\mathbf{p}) \leq H_b^k(\mathbf{q})$$

proposed in the case when  $p_1 \ge \cdots \ge p_n$  and  $q_1 \ge \cdots \ge q_n$  and  $\mathbf{p} > \mathbf{q}$ . However the proof is not correct: it depends on a step which states that because

$$\log\left(\frac{1}{\sum_{i=1}^{k} q_i}\right) \ge H_b^k(\mathbf{q}) - \log k, \quad \log\left(\frac{1}{\sum_{i=1}^{k} p_i}\right) \ge H_b^k(\mathbf{p}) - \log k$$

and

[9]

$$\log\left(\frac{1}{\sum_{i=1}^{k} p_i}\right) \le \log\left(\frac{1}{\sum_{i=1}^{k} q_i}\right)$$

for  $1 \le k \le n$ , we must have  $H_b^k(\mathbf{p}) - \log k \ge H_b^k(\mathbf{q}) - \log k$ . This need not hold.

In the case when X is a discrete random variable with countable range  $\{x_i\}_{i=1}^{\infty}$  and probability distribution  $p_i := P\{X = x_i\} > 0$   $(\sum_{i=1}^{\infty} p_i = 1)$ , we can define the entropy of X by

$$H_b(X) := \sum_{i=1}^{\infty} p_i \log 1/p_i,$$
 (3.5)

but now only for those random variables X for which the sum is finite. In that case it is evident that  $H_b(X) > 0$  since  $0 < p_i < 1$  for all *i*. Since there is no uniform probability distribution on an infinite sample space, we can't extend our finite-range results to the general case of a countable range. In that case the main question is finding sufficient conditions on the probability distribution of X under which the entropy  $H_b(X)$  exists. Here we offer one possible answer to this question. A central role is played by the quantity  $v := \sum_{i=1}^{\infty} ip_i$  when it exists. We begin with an example.

Let X be a discrete random variable with a geometric probability distribution

$$p_i = P\{X = i\} = (1 - a)a^{i-1}, \quad i = 1, 2, ..., 0 < a < 1.$$

This random variable has a finite mean

$$\mu = \nu = \sum_{i=1}^{\infty} i(1-a)a^{i-1} = \frac{1}{1-a} > 1$$

and entropy

$$H_b(X) = \sum_{i=1}^{\infty} (1-a)a^{i-1} \left[ -\log(1-a) - (i-1)\log a \right] = \log \frac{a}{1-a} - \frac{1}{1-a}\log a.$$

Since  $a = 1 - 1/\mu$ , we get

$$H_b(X) = \log \frac{\mu^{\mu}}{(\mu - 1)^{\mu - 1}}.$$

THEOREM 7. Let X be a discrete random variable with countable range  $\{x_1, x_2, ...\}$ and probability distribution  $p_i = P\{X = x_i\} > 0$   $\left(\sum_{i=1}^{\infty} p_i = 1\right)$  such that  $v < \infty$ . Then the entropy  $H_b(X)$  defined by (3.5) is finite and

$$0 < H_b(X) \le \log \frac{\nu^{\nu}}{(\nu-1)^{\nu-1}}.$$

If in addition  $\sum_{i=1}^{\infty} \left[ v^i / (v-1)^{i-1} \right] p_i^2 < \infty$ , then

$$0 \le \log \frac{\nu^{\nu}}{(\nu-1)^{\nu-1}} - H_b(X)$$
  
$$\le \log \left[ \sum_{i=1}^{\infty} \frac{\nu^i}{(\nu-1)^{i-1}} p_i^2 \right] \le \frac{1}{\ln b} \left[ \sum_{i=1}^{\infty} \frac{\nu^i}{(\nu-1)^{i-1}} p_i^2 - 1 \right], \quad (3.6)$$

with equality throughout if and only if  $p_i = (v - 1)^{i-1}/v^i$  for all  $i \in \mathbb{N}$ .

PROOF. Set  $q_i = (1 - a)a^{i-1}$  ( $i \in \mathbb{N}$ ), where a is any real number with 0 < a < 1. We have  $\alpha = \sum_{i=1}^{\infty} q_i = 1$  and

$$S_q = \sum_{i=1}^{\infty} p_i \left[ -\log(1-a) - (i-1)\log a \right] = \log \frac{a}{1-a} - \nu \log a < \infty.$$

By Theorem 4 we know that  $S_p = \sum_{i=1}^{\infty} p_i \log(1/p_i) = H_b(X) < \infty$  and

$$0 < S_p \leq S_q + \log \alpha = \log \frac{a}{1-a} - \nu \log a.$$

It is easily shown that for 0 < a < 1 the function

$$f(a) = \log(a/(1-a)) - \nu \log a$$

achieves its minimal value at  $a = (\nu - 1)/\nu$  and that the minimal value is

$$\log\left[\nu^{\nu}/(\nu-1)^{\nu-1}\right].$$

(Note that  $\nu = \sum_{i=1}^{\infty} ip_i > \sum_{i=1}^{\infty} p_i = 1$ .) For  $a = (\nu - 1)/\nu$  we have  $(1 - a)a^{i-1} = (\nu - 1)^{i-1}/\nu^i$  and the desired results follow by Theorem 4 with  $q_i = (\nu - 1)^{i-1}/\nu^i$ .

The first inequality in (3.6) has been proved in [5].

The first upper bound for the difference  $\log \left[ \nu^{\nu} / (\nu - 1)^{\nu - 1} \right] - H_b(X)$  given by (3.6) is nontrivial only in the case when

$$\sum_{i=1}^{\infty} \frac{\nu^{i}}{(\nu-1)^{i-1}} p_{i}^{2} < \frac{\nu^{\nu}}{(\nu-1)^{\nu-1}},$$

while the last upper bound is nontrivial in the case when

$$\frac{1}{\ln b} \left[ \sum_{i=1}^{\infty} \frac{\nu^{i}}{(\nu-1)^{i-1}} p_{i}^{2} - 1 \right] < \log \frac{\nu^{\nu}}{(\nu-1)^{\nu-1}},$$

502

since  $H_b(X) > 0$ . The last condition is equivalent to

$$\sum_{i=1}^{\infty} \frac{\nu^{i}}{(\nu-1)^{i-1}} p_{i}^{2} < \ln \frac{\nu^{\nu} e}{(\nu-1)^{\nu-1}}.$$

Theorem 7 shows that the entropy  $H_b(X)$  is approximately  $\log K$ , where  $K := \nu^{\nu}/(\nu - 1)^{\nu-1}$ , whenever X has a probability distribution 'close enough' to the geometric probability distribution  $q_i = (1 - a)a^{i-1}$  ( $i \in \mathbb{N}$ ) with  $a = (\nu - 1)/\nu$ , where  $\nu$  is the mean of X.

THEOREM 8. Under the assumptions of Theorem 7, if

$$0 < L \le \frac{\nu^i}{(\nu-1)^{i-1}} p_i \le U \quad \text{for all } i \in \mathbb{N},$$

then

$$0 \le \log K - H_b(X) \le \log \frac{(U+L)^2}{4LU} \le \frac{1}{4\ln b} \frac{(U-L)^2}{LU}.$$
 (3.7)

Also we have

$$0 \le \log K - H_b(X) \le \log \left[ \left( \sqrt{U} - \sqrt{L} \right)^2 + 1 \right] \le \frac{1}{\ln b} \left( \sqrt{U} - \sqrt{L} \right)^2.$$
(3.8)

PROOF. To obtain (3.7) apply Theorem 5, and to obtain (3.8) apply Theorem 6, with L, U in place of m, M and with  $q_i = (\nu - 1)^{i-1}/\nu^i$   $(i \in \mathbb{N})$ .

Since  $H_b(X) > 0$ , the first upper bound for the difference  $D := \log K - H_b(X)$  given in (3.7) is nontrivial only in the case when

$$\frac{(U+L)^2}{4LU} < K.$$
 (3.9)

If we set  $\rho = U/L$ , then (3.9) is equivalent to  $(\rho + 1)^2/\rho < 4K$ , that is,

$$\rho^2 - 2(2K - 1)\rho + 1 < 0$$

or

$$\left[2K - 1 + 2\sqrt{K(K-1)}\right]^{-1} < \rho < 2K - 1 + 2\sqrt{K(K-1)}.$$

Since  $\rho \ge 1$ , it follows that the first upper bound in (3.7) gives a nontrivial upper bound for *D* only in the case when  $\rho = U/L < 2K - 1 + 2\sqrt{K(K-1)}$ . Similarly, the second upper bound in (3.7) gives a nontrivial upper bound for *D* only in the case when

$$\frac{1}{4\ln b} \frac{(U-L)^2}{LU} < \log K$$
(3.10)

504

or  $\rho < 1 + 2 \ln K + 2\sqrt{(1 + \ln K) \ln K}$ . Likewise (3.8) gives a nontrivial first upper bound for D only when  $\sqrt{U} - \sqrt{L} < \sqrt{K-1}$  and a nontrivial second upper bound only when  $\sqrt{U} - \sqrt{L} < \sqrt{\ln K}$ .

Shannon's inequality can be given in the form

$$L(\mathbf{p},\mathbf{q})\geq 0,$$

where  $L(\mathbf{p}, \mathbf{q}) := \sum_{i \in I} p_i \log(p_i/q_i)$  represents the Kulback-Leibler distance or relative entropy between the distributions  $\mathbf{p} = \{p_i\}_{i \in I}$  and  $\mathbf{q} = \{q_i\}_{i \in I}$ . Alencar and Assis [1] established the inequality

$$L(\mathbf{p}, \mathbf{q}) + H(\mathbf{p}) \ge \frac{d(\mathbf{p}, \mathbf{q})}{2\ln b},$$
(3.11)

where  $d(\mathbf{p},\mathbf{q}) := \sum_{i \in I} |p_i - q_i|$  is the variational distance between the probability distributions **p** and **q**. The quantity  $L(\mathbf{p}, \mathbf{q}) + H(\mathbf{p})$  is referred to as 'inaccuracy'. The following stronger result also holds.

THEOREM 9. With the above notation,

$$L(\mathbf{p},\mathbf{q}) + H(\mathbf{p}) \ge -\log\left[1-\frac{1}{2}d(\mathbf{p},\mathbf{q})\right] \ge \frac{d(\mathbf{p},\mathbf{q})}{2\ln b}.$$

**PROOF.** Jensen's inequality gives  $\sum_{i \in I} p_i \log q_i \le \log \sum_{i \in I} p_i q_i$ , so

$$\sum_{i \in I} p_i \log p_i - \sum_{i \in I} p_i \log \frac{p_i}{q_i} \le \log \sum_{i \in I} p_i q_i$$
$$L(\mathbf{p}, \mathbf{q}) + H(\mathbf{p}) \ge -\log \sum_{i \in I} p_i q_i.$$
(3.12)

or

$$L(\mathbf{p}, \mathbf{q}) + H(\mathbf{p}) \ge -\log \sum_{i \in I} p_i q_i.$$
(3.12)

We now address the right-hand side of (3.12). We have

$$-\sum_{i\in I} p_i q_i = \frac{1}{2} \sum_{i\in I} \left[ p_i (1-q_i) + q_i (1-p_i) \right] - 1.$$
(3.13)

Moreover,

$$p_i(1-q_i) + q_i(1-p_i) \ge |p_i - q_i|$$
. (3.14)

By symmetry it suffices to show this for  $p_i > q_i$ , when (3.14) becomes

$$p_i - p_i q_i + q_i - p_i q_i \ge p_i - q_i$$

or the trivial inequality  $2q_i(1-p_i) \ge 0$ .

Hence (3.14) holds and we have from (3.13) that

$$\sum_{i \in I} p_i q_i \le 1 - \frac{1}{2} \sum_{i \in I} |p_i - q_i| = 1 - \frac{1}{2} d(\mathbf{p}, \mathbf{q})$$

and so from (3.12) that

$$L(\mathbf{p},\mathbf{q}) + H(\mathbf{p}) \geq -\log\left[1-\frac{1}{2}d(\mathbf{p},\mathbf{q})\right].$$

The elementary inequality  $\log x \le (x-1)/\ln b$  or  $-\log x \ge (-x+1)/\ln b$  yields

$$-\log\left[1-\frac{1}{2}d(\mathbf{p},\mathbf{q})\right] \geq \frac{1}{\ln b}\left[-1+\frac{1}{2}d(\mathbf{p},\mathbf{q})+1\right] = \frac{1}{2\ln b}d(\mathbf{p},\mathbf{q}).$$

#### 4. A counterpart to Shannon's integral inequality

We begin this section with a counterpart of the integral Shannon inequality.

THEOREM 10. Let I be a measurable subset of the real line and p(x) and q(x) positive integrable functions on I such that  $\int_{I} p(x) dx = 1$  and  $\alpha := \int_{I} q(x) dx < \infty$ . Suppose that for b > 1 at least one of the integrals

$$J_p := \int_I p(x) \log \frac{1}{p(x)} dx \quad and \quad J_q := \int_I p(x) \log \frac{1}{q(x)} dx$$

is finite. If  $\int_{I} (p^2(x)/q(x)) dx < \infty$ , then both  $J_p$  and  $J_q$  are finite and

$$0 \le J_q - J_p + \log_b \alpha \le \log \left[ \alpha \int_I \frac{p^2(x)}{q(x)} dx \right] \le \frac{1}{\ln b} \left[ \alpha \int_I \frac{p^2(x)}{q(x)} dx - 1 \right], \quad (4.1)$$

with equality throughout if and only if  $q(x) = \alpha p(x)$  a.e. on I.

PROOF. Set  $\xi(x) = q(x)/p(x)$  ( $x \in I$ ). We have  $\int_I p(x)\xi(x) dx = \alpha < \infty$  and  $\int_I (p(x)/\xi(x)) dx = \int_I (p^2(x)/q(x)) dx < \infty$ . So we can apply Corollary 2 to obtain

$$0 \le \log \alpha - \int_{I} p(x) \log \frac{q(x)}{p(x)} dx \le \log \left[ \alpha \int_{I} \frac{p^{2}(x)}{q(x)} dx \right] \le \frac{1}{\ln b} \left[ \alpha \int_{I} \frac{p^{2}(x)}{q(x)} dx - 1 \right],$$

and we know that  $J := \int_{I} p(x) \log(q(x)/p(x)) dx$  is finite. Now if  $J_{p}$  is finite then  $J_{p} - J = J_{q}$  is finite too; if  $J_{q}$  is finite then  $J_{q} + J = J_{p}$  is finite too. So we can write  $J = J_{p} - J_{q}$  and (4.1) is proved. Moreover, by Corollary 2, equality holds throughout (4.1) if and only if  $\xi(x) = c$  a.e. on *I*, that is, q(x) = cp(x) a.e. on *I*. Since  $\int_{I} p(x) dx = 1$  and  $\int_{I} q(x) dx = \alpha$ , we have  $c = \alpha$ .

[13]

Now we proceed to a result analogous to Theorem 5.

THEOREM 11. Suppose the assumptions of Theorem 10 be satisfied and

$$0 < m \leq \frac{p(x)}{q(x)} \leq M \quad a.e. \text{ on } I.$$

Then

$$0 \le \int_{I} p(x) \log \frac{1}{q(x)} dx - \int_{I} p(x) \log \frac{1}{p(x)} dx + \log \alpha$$
  
$$\le \log \frac{(M+m)^{2}}{4Mm} \le \frac{1}{4 \ln b} \frac{(M-m)^{2}}{Mm}.$$

PROOF. Set f(x) = q(x)/p(x), g(x) = p(x)/q(x) for  $x \in I$  and P = 1. Then by Lemma 2

$$\alpha \int_{I} \frac{p^{2}(x)}{q(x)} dx - 1 = \int_{I} p(x) \frac{q(x)}{p(x)} dx \int_{I} p(x) \frac{p(x)}{q(x)} dx - \int_{I} p(x) \frac{q(x)}{p(x)} \frac{p(x)}{q(x)} dx$$
$$\leq \frac{1}{4} \left(\frac{1}{m} - \frac{1}{M}\right) (M - m),$$

that is,

$$\alpha \int_{I} \frac{p^2(x)}{q(x)} dx \leq \frac{(M+m)^2}{4Mm}.$$

Theorem 10 provides the desired result.

We can prove an integral version of (3.2) and use it to obtain a result analogous to Theorem 6. So under the assumptions of Theorem 11, we have

$$0 \leq \int_{I} p(x) \log \frac{1}{q(x)} dx - \int_{I} p(x) \log \frac{1}{p(x)} dx + \log \alpha$$
$$\leq \log \left[ \alpha \left( \sqrt{M} - \sqrt{m} \right)^{2} + 1 \right] \leq \frac{\alpha}{\ln b} \left( \sqrt{M} - \sqrt{m} \right)^{2}. \tag{4.2}$$

The notion of entropy  $H_b(X)$  can be extended to the case of a general random variable X, by approximating X by discrete random variables. In the case when X is nondiscrete,  $H_b(X)$  is usually infinite. For example, this always happens when X is continuous (see [7, p. 38]).

In the case when X is a continuous random variable with density p(x) (a nonnegative measurable function on **R** such that  $\int_R p(x) dx = 1$ ), we may define the so-called differential entropy of X by

$$h_b(X) := \int_R p(x) \log \frac{1}{p(x)} \, dx \quad (b > 1)$$
(4.3)

whenever the integral exists. The following result (with b = e) is given in [9, p. 33].

THEOREM 12. Let X be a continuous random variable with density p(x) ( $x \in \mathbf{R}$ ).

(a) If X has a finite variance  $s^2$ , then  $h_b(X)$  exists and

$$h_b(X) \le \log\left(s\sqrt{2\pi e}\right),$$
 (4.4)

with equality if and only if X is a Gaussian random variable with variance  $s^2$ . (b) If p(x) = 0 for all x < 0 and X has finite mean  $\mu$ , then  $h_b(X)$  exists and

$$h_b(X) \le \log(\mu e),\tag{4.5}$$

with equality if and only if  $p(x) = (1/\mu)e^{-x/\mu}$  for  $x \ge 0$ .

(c) If p(x) = 0 for all  $x \in \mathbb{R} \setminus [c, d]$ , where  $-\infty < c < d < \infty$ , then  $h_b(X)$  exists and

$$h_b(X) \le \log(d-c),\tag{4.6}$$

with equality if and only if p(x) = 1/(d - c) for  $x \in [c, d]$ .

From (4.6) we conclude that  $h_b(X) < 0$  whenever d - c < 1. In fact the possibility  $h_b(X) = -\infty$  is not excluded in (4.4)-(4.6), as is seen from the following example. Define

$$p(x) = \begin{cases} 1/(x \ln^2 x), & x \in (0, 1/e), \\ 0, & x \in \mathbf{R} \setminus (0, 1/e). \end{cases}$$

We have

$$\int_{\mathbf{R}} p(x) \, dx = \int_{0}^{1/e} \frac{1}{x \ln^{2} x} \, dx = \left[ -\frac{1}{\ln x} \right]_{0}^{1/e} = 1$$

and

$$\int_{\mathbf{R}} xp(x) \, dx = \int_0^{1/e} \frac{1}{\ln^2 x} \, dx < \infty, \quad \int_{\mathbf{R}} x^2 p(x) \, dx = \int_0^{1/e} \frac{x}{\ln^2 x} \, dx < \infty,$$

since the functions  $f(x) = 1/\ln^2 x$  and  $g(x) = x/\ln^2 x$  are continuous, positive and bounded on (0, 1/e). Thus a random variable with density p(x) ( $x \in \mathbf{R}$ ) has finite mean and variance but its differential *e*-entropy is

$$h_e(X) = \int_0^{1/e} \frac{1}{x \ln x} \, dx + \int_0^{1/e} \frac{\ln(\ln^2 x)}{x \ln^2 x} \, dx = -\infty,$$

since the first integral is  $\ln(-\ln x)|_0^{1/e} = -\infty$  and the second (by the substitution  $t = \ln(-\ln x)$ ) is found to be  $-2(t+1)e^{-t}|_0^{\infty} = 2$ .

THEOREM 13. Let X be a continuous random variable with density p(x) ( $x \in \mathbf{R}$ ).

(a) If X has a finite mean  $\mu$  and variance  $s^2$  and

$$\int_{\mathbf{R}} p^2(x) \exp\left[\frac{1}{2s^2}(x-\mu)^2\right] dx < \infty,$$

then  $h_b(X)$  is finite and

$$0 \le \log\left(s\sqrt{2\pi e}\right) - h_b(X) \le \log\left\{s\sqrt{2\pi}\int_{\mathbf{R}}p^2(x)\exp\left[\frac{1}{2s^2}(x-\mu)^2\right]dx\right\}$$
$$\le \frac{1}{\ln b}\left\{s\sqrt{2\pi}\int_{\mathbf{R}}p^2(x)\exp\left[\frac{1}{2s^2}(x-\mu)^2\right]dx - 1\right\},$$

with equality throughout if and only if  $p(x) = (1/s\sqrt{2\pi}) \exp\left[-(x-\mu)^2/(2s^2)\right] a.e.$ on **R**.

(b) Suppose that p(x) = 0 for all x < 0 and that X has finite mean  $\mu$ . If

$$\int_0^\infty p^2(x) \exp(x/\mu) \, dx < \infty,$$

then  $h_b(X)$  is finite and

$$0 \le \log(\mu e) - h_b(X) \le \log\left[\mu \int_0^\infty p^2(x) \exp(x/\mu) \, dx\right]$$
$$\le \frac{1}{\ln b} \left[\mu \int_0^\infty p^2(x) \exp(x/\mu) \, dx - 1\right]$$

with equality throughout if and only if  $p(x) = (1/\mu) \exp(-x/\mu)$  a.e. on  $[0, \infty)$ . (c) Suppose that p(x) = 0 for  $x \in \mathbb{R} \setminus [c, d]$ , where  $-\infty < c < d < \infty$ . If

$$\int_c^{d'} p^2(x)\,dx\,<\infty,$$

then  $h_b(X)$  is finite and

$$0 \le \log(d-c) - h_b(X) \le \log\left[(d-c)\int_c^d p^2(x)\,dx\right]$$
$$\le \frac{1}{\ln b}\left[(d-c)\int_c^d p^2(x)\,dx - 1\right]$$

with equality throughout if and only if p(x) = 1/(d - c) a.e. on [c, d].

PROOF. (a) The finiteness of the variance  $s^2$  of X implies that  $\mu = \int_{\mathbf{R}} xp(x) dx$  and  $s^2 = \int_{\mathbf{R}} (x - \mu)^2 p(x) dx > 0$  are well-defined real numbers. So we can define

$$q(x) = (1/s\sqrt{2\pi}) \exp[-(x-\mu)^2/(2s^2)]$$

for  $x \in \mathbf{R}$ . We have  $\alpha = \int_{\mathbf{R}} q(x) dx = 1$  and

$$\int_{\mathbf{R}} p(x) \log \frac{1}{q(x)} dx = \frac{1}{\ln b} \int_{\mathbf{R}} p(x) \ln \frac{1}{q(x)} dx$$
$$= \frac{1}{\ln b} \int_{\mathbf{R}} p(x) \left[ \ln \left( s\sqrt{2\pi} \right) + \frac{(x-\mu)^2}{2s^2} \right] dx$$
$$= \frac{1}{\ln b} \left[ \ln \left( s\sqrt{2\pi} \right) + \frac{1}{2} \right] = \log \left( s\sqrt{2\pi e} \right).$$

Application of Theorem 10 with  $I = \mathbf{R}$  provides the desired results.

(b) Under the given assumptions, we have that  $\mu = \int_{\mathbb{R}} xp(x) dx = \int_{0}^{\infty} xp(x) dx$  is a well-defined positive number and we can define  $q(x) = (1/\mu) \exp(-x/\mu)$  ( $x \in I = [0, \infty)$ ). We have  $\alpha = \int_{I} q(x) dx = 1$  and

$$\int_{I} p(x) \log \frac{1}{q(x)} dx = \frac{1}{\ln b} \int_{I} p(x) \ln \frac{1}{q(x)} dx$$
$$= \frac{1}{\ln b} \int_{0}^{\infty} p(x) \left( \ln \mu + \frac{x}{\mu} \right) dx = \frac{1}{\ln b} (\ln \mu + 1) = \log(\mu e).$$

Again application of Theorem 10 yields the stated results. (c) Set I = [c, d] and q(x) = 1/(d - c),  $x \in I$ . It is evident that  $\alpha = \int_{I} q(x) dx = 1$  and

$$\int_{I} p(x) \log \frac{1}{q(x)} dx = \log(d-c) \int_{c}^{d} p(x) dx = \log(d-c).$$

We can apply Theorem 10 again to complete our proof.

Theorem 13 shows that  $h_b(X) \approx \log(s\sqrt{2\pi e})$  when the distribution of X is 'close' to the Gaussian distribution with variance  $s^2$ . Also,  $h_b(X) \approx \log(\mu e)$  if the distribution of X is 'close' to the exponential distribution with mean  $\mu$ . Finally  $h_b(X) \approx \log \ell$  whenever the distribution of X is 'close' to the uniform distribution over an interval of length  $\ell$ .

THEOREM 14. (a) Under the assumptions of Theorem 13 (a), if

$$0 < L \le p(x) \exp\left[\frac{1}{2s^2}(x-\mu)^2\right] \le U \quad a.e. \text{ on } \mathbf{R},$$

then

$$0 \le \log \left( s \sqrt{2\pi e} \right) - h_b(X) \le \log \frac{(U+L)^2}{4LU} \le \frac{1}{4 \ln b} \frac{(U-L)^2}{LU}.$$

(b) Under the assumptions of Theorem 13 (b), if

$$0 < L \le p(x) \exp(x/\mu) \le U \quad a.e. \text{ on } [0, \infty),$$

then

$$0 \le \log_b(\mu e) - h_b(X) \le \log \frac{(U+L)^2}{4LU} \le \frac{1}{4\ln b} \frac{(U-L)^2}{LU}$$

(c) Under the assumptions of Theorem 13 (c), if

$$0 < L \le p(x) \le U \quad a.e. \text{ on } [c, d],$$

then

$$0 \le \log_b(d-c) - h_b(X) \le \log \frac{(U+L)^2}{4LU} \le \frac{1}{4\ln b} \frac{(U-L)^2}{LU}$$

PROOF. (a) Apply Theorem 11 with  $s\sqrt{2\pi}L$  and  $s\sqrt{2\pi}U$  in place of *m* and *M* respectively and with q(x) as in the proof of Theorem 13 (a).

(b) Apply Theorem 11 with  $\mu L$  and  $\mu U$  in place of *m* and *M* respectively and with q(x) as in the proof of Theorem 13 (b).

(c) Apply Theorem 11 with (d - c)L and (d - c)U in place of m and M respectively and with q(x) as in the proof of Theorem 13 (c).

We can obtain upper bounds of the types

$$\log\left[s\sqrt{2\pi}\left(\sqrt{U}-\sqrt{L}\right)^{2}+1\right], \quad \log\left[\mu\left(\sqrt{U}-\sqrt{L}\right)^{2}+1\right], \\ \log\left[\left(d-c\right)\left(\sqrt{U}-\sqrt{L}\right)^{2}+1\right]$$

in an obvious way using the inequalities (4.2).

THEOREM 15. Suppose that the assumptions of Theorem 14 (c) are satisfied. Then

$$0 \le \log_b(d-c) - h_b(X) \le \log\left[\frac{(U-L)^2}{4}\min\left\{(d-c)^2, \frac{1}{LU}\right\} + 1\right]$$
$$\le \frac{(U-L)^2}{4\ln b}\min\left\{(d-c)^2, \frac{1}{LU}\right\}.$$

PROOF. Set I = [c, d] and p(x) = 1 and f(x) = g(x) = p(x) in Lemma 2 to obtain

$$(d-c)\int_{c}^{d}p^{2}(x)\,dx-1\leq\frac{1}{4}(U-L)^{2}(d-c)^{2},$$

that is,

$$(d-c)\int_{c}^{d}p^{2}(x)\,dx\leq \frac{1}{4}(U-L)^{2}(d-c)^{2}+1.$$

Now apply Theorem 13 (c) to obtain

$$0 \le \log(d-c) - h_b(X) \le \log\left[\frac{(U-L)^2}{4}(d-c)^2 + 1\right] \le \frac{(U-L)^2}{4\ln b}(d-c)^2.$$

Combine this with Theorem 14 (c) and the desired result follows.

# 5. Acknowledgments

The authors wish to acknowledge partial funding by a DSTO TSS grant.

#### References

- M. S. Alencar and F. M. Assis, "Inequalities involving the incomplete Renyi entropy, divergence and variational distance", in *Proceedings Globecom '98, Sydney* (1998).
- [2] S. S. Dragomir, "Grüss inequality in inner product spaces", RGMIA Research Report Collection 1 (1998) 67–71.
- [3] 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 (1996) 1-11.
- [4] M. Matić, C. E. M. Pearce and J. Pečarić, "Improvements of some bounds on entropy measures in information theory", *Math. Ineq. Appl.* 1 (1998) 295–304.
- [5] M. Matić, C. E. M. Pearce and J. Pečarić, "On an inequality for the entropy of a probability distribution", Acta Math. Hungarica 85 (1999) 345–349.
- [6] M. Matić, C. E. M. Pearce and J. Pečarić, "Refinements of some bounds in information theory", ANZIAM J. 42 (2001) 387–399.
- [7] R. J. McEliece, The theory of information and coding (Addison-Wesley, Reading, Mass., 1977).
- [8] D. S. Mitrinović, J. E. Pečarić and A. M. Fink, Classical and new inequalities in analysis (Kluwer, Dordrecht, 1993).
- [9] Z. Pauše, Uvod u teoriju informacije (Školska Knjiga, Zagreb, 1989).
- [10] J. E. Pečarić, F. Proschan and Y. L. Tong, Convex functions, partial orderings, and statistical applications (Academic Press, New York, 1992).
- [11] W. Rudin, Real and complex analysis (McGraw-Hill, London, 1970).