

ON SOME INEQUALITIES FOR ENTROPIES OF DISCRETE PROBABILITY DISTRIBUTIONS

CVETAN JARDAS¹, JOSIP PEČARIĆ², RAJKO ROKI³ and NIKOLA SARAPA⁴

(Received 13 February 1997; revised 28 August 1997)

Abstract

In this paper we prove some inequalities for finite sums and infinite series with positive terms. As an application of these results we obtain some inequalities for entropies of discrete probability distributions.

1. Introduction

Let $(p_k > 0, k = 0, 1, 2, \dots)$ be a discrete probability distribution (finite or countably infinite). The Shannon entropy of this distribution is defined by

$$H = - \sum_k p_k \log p_k. \quad (1)$$

The quantity H is a measure of information of the distribution $(p_k > 0, k = 0, 1, 2, \dots)$ and it plays a key role in information theory. For this reason in many applications of discrete probability distributions, it is important to find lower and upper bound for entropy H .

In [1] the following result has been proved.

Let $(p_n > 0, n \geq 0)$ be a countably infinite probability distribution such that

$$\sup_{n \geq 0} \left(p_n^{-1} \sum_{k=n+1}^{\infty} p_k \right) \leq \lambda < \infty. \quad (2)$$

¹Department of Mathematics, University of Rijeka, 51000 Rijeka, Omladinska 14, Croatia.

²Faculty of Textile Technology, University of Zagreb, 10000 Zagreb, Pierottijeva 6, Croatia.

³Faculty of Food Technology and Biotechnology, University of Zagreb, 10000 Zagreb, Pierottijeva 6, Croatia.

⁴Department of Mathematics, University of Zagreb, 10000 Zagreb, Bijenička 30, Croatia.

© Australian Mathematical Society 1999, Serial-fee code 0334-2700/99

Then

$$H \leq F(\lambda), \quad (3)$$

where

$$F(x) = (x + 1) \log(x + 1) - x \log x, \quad x > 0. \quad (4)$$

Equality holds in (3) if $p_n = \frac{\lambda^n}{(\lambda + 1)^n}$ ($n \geq 0$).

This result has been generalized in [3] where the lower and upper bounds for H have been obtained by use of the function F .

In [4] the authors found another upper bound for H , where the probability distribution satisfies (2).

By the use of the function F defined by (4), we will find in this paper a lower bound for the entropy of a finite probability distribution which is non-increasing in mean, and an upper bound for the entropy of a finite probability distribution which is non-decreasing in mean.

First we have the following definition.

DEFINITION. A sequence $(a_k, k \in \mathbf{N}) \subset \mathbf{R}$ is *non-increasing in mean* if

$$\frac{1}{n} \sum_{k=1}^n a_k \geq \frac{1}{n+1} \sum_{k=1}^{n+1} a_k, \quad n \in \mathbf{N}. \quad (5)$$

The sequence $(a_k, k \in \mathbf{N}) \subset \mathbf{R}$ is *non-decreasing in mean* if we reverse the inequality in (5).

In a similar way we define when a finite sequence $(a_k, k = 1, 2, \dots, n) \subset \mathbf{R}$ is non-increasing in mean or non-decreasing in mean.

It is well known and easy to see that if $(a_k, k \in \mathbf{N})$ is non-increasing or non-decreasing, then it is also non-increasing in mean or non-decreasing in mean, respectively, but the reverse implications do not hold in general.

It is also easy to check that $(a_k, k \in \mathbf{N})$ is non-increasing in mean if and only if

$$S_k = \sum_{i=1}^k a_i \geq k a_k, \quad k \in \mathbf{N}. \quad (6)$$

The sequence $(a_k, k \in \mathbf{N})$ is non-decreasing in mean if and only if

$$S_k \leq k a_k, \quad k \in \mathbf{N}. \quad (7)$$

We need the following proposition (for the proof see [4, Corollary 1]).

PROPOSITION. Let $f(x) = \frac{1 - (1 - x)^p}{x}$, $x \in (0, 1)$.

(i) If $0 < p < 1$ or $p > 2$, then f is strictly convex on $(0, 1)$. If $1 < p < 2$, then f is strictly concave on $(0, 1)$.

(ii) If $0 < p < 1$, then $f'(x) > 0$, $x \in (0, 1)$. If $p > 1$, then $f'(x) < 0$, $x \in (0, 1)$.

(iii) For all $p > 0$, we have $f(x) > 0$, $x \in (0, 1)$.

In the following sections we suppose that all series are convergent.

2. Results

We first prove the following theorem.

THEOREM 1. Let $(a_k > 0, k = 1, 2, \dots, n)$ be non-increasing in mean. Then

(i) For all $p, 0 < p < 1$, we have

$$\left(\sum_{k=1}^n a_k\right)^p \leq \sum_{k=1}^n [k^p - (k - 1)^p] a_k^p + \sum_{k=2}^n \frac{1}{k} a_k^{-1} D_k (ka_k - S_k), \tag{8}$$

where

$$D_k = p a_k S_{k-1}^{p-1} + S_{k-1}^p - S_k^p = a_k^2 S_k^{p-2} f' \left(\frac{a_k}{S_k}\right), \quad k = 2, 3, \dots, n, \tag{9}$$

where f is the function from the Proposition.

For all $p, 1 < p < 2$, the opposite inequality holds in (8).

(ii) For all $p, p > 2$, we have

$$\left(\sum_{k=1}^n a_k\right)^p \geq \sum_{k=1}^n [k^p - (k - 1)^p] a_k^p + \sum_{k=2}^n a_k S_k^{p-2} C_k (ka_k - S_k), \tag{10}$$

where

$$C_k = k^{1-p} [p(k - 1)^{p-1} + (k - 1)^p - k^p] = \frac{1}{k} f' \left(\frac{1}{k}\right), \quad k = 2, 3, \dots, n. \tag{11}$$

Equalities hold in (8) and (10) if and only if $a_1 = a_2 = \dots = a_n$.

PROOF. (i) For $0 < p < 1$, the function f from the Proposition is strictly convex on $(0, 1)$, so for all $x, y \in (0, 1)$ we have

$$f(x) \leq f(y) + (x - y)f'(x). \tag{12}$$

Substituting $x = a_k/S_k$ and $y = 1/k$ into (12) and using the fact that $S_k - a_k = S_{k-1}$ (for $k = 2, 3, \dots, n$), we get

$$S_k^p - S_{k-1}^p \leq f\left(\frac{1}{k}\right) a_k S_k^{p-1} + \left(\frac{a_k}{S_k} - \frac{1}{k}\right) a_k S_k^{p-1} f'\left(\frac{a_k}{S_k}\right), \quad k = 2, 3, \dots, n, \quad (13)$$

where $f\left(\frac{1}{k}\right) = [k^p - (k-1)^p]k^{1-p} > 0$ and $f'\left(\frac{a_k}{S_k}\right) = (p a_k S_{k-1}^{p-1} + S_{k-1}^p - S_k^p) S_k^{2-p} a_k^{-2} > 0$, by the Proposition from the Introduction. Since $0 < p < 1$, it follows from (6) that $a_k S_k^{p-1} \leq k^{p-1} a_k^p$ ($k = 2, 3, \dots, n$) so we conclude that

$$S_k^p - S_{k-1}^p \leq [k^p - (k-1)^p] a_k^p + \frac{1}{k} f'\left(\frac{a_k}{S_k}\right) a_k S_k^{p-2} (k a_k - S_k), \quad k = 2, 3, \dots, n. \quad (14)$$

Summing the inequalities in (14) for all $k = 2, 3, \dots, n$ leads to (8) since

$$\sum_{k=2}^n (S_k^p - S_{k-1}^p) + a_1^p = S_n^p = \left(\sum_{k=1}^n a_k \right)^p.$$

We can also conclude that in this case we have $D_k > 0$ ($k = 2, 3, \dots, n$).

For $1 < p < 2$, the function f is strictly concave on $(0, 1)$ so we have the opposite inequalities in (12) and (13). Since $1 < p < 2$, it follows from (6) that $a_k S_k^{p-1} \geq k^{p-1} a_k^p$ ($k = 2, 3, \dots, n$) so we have the opposite inequality in (14). Summing these inequalities for all $k = 2, 3, \dots, n$, we get the opposite inequality in (8). In this case we have $D_k < 0$ ($k = 2, 3, \dots, n$).

(ii) If $p > 2$, the function f is strictly convex on $(0, 1)$. Therefore for all $x, y \in (0, 1)$ we have

$$f(x) \geq f(y) + (x - y)f'(y). \quad (15)$$

Substituting $x = \frac{a_k}{S_k}$ and $y = \frac{1}{k}$ into (15) we get

$$S_k^p - S_{k-1}^p \geq f\left(\frac{1}{k}\right) a_k S_k^{p-1} + \left(\frac{a_k}{S_k} - \frac{1}{k}\right) a_k S_k^{p-1} f'\left(\frac{1}{k}\right), \quad k = 2, 3, \dots, n, \quad (16)$$

where $f\left(\frac{1}{k}\right) = [p(k-1)^{p-1} + (k-1)^p - k^p]k^{2-p} < 0$. It follows from (6) that $a_k S_k^{p-1} \geq k^{p-1} a_k^p$ ($k = 2, 3, \dots, n$), so we get from (16) that

$$S_k^p - S_{k-1}^p \geq [k^p - (k-1)^p] a_k^p + a_k S_k^{p-2} + \frac{1}{k} f'\left(\frac{1}{k}\right) (k a_k - S_k), \quad k = 2, 3, \dots, n. \quad (17)$$

Summing the inequalities in (17) for all $k = 2, 3, \dots, n$ leads to (10). We also have $C_k < 0$ ($k = 2, 3, \dots, n$).

The last statement follows from the fact that equalities hold in (12) and (15) if and only if $x = y$ and therefore equalities hold in (8) and (10) if and only if $S_k = ka_k$ ($k = 2, 3, \dots, n$) and this is equivalent to $a_1 = a_2 = \dots = a_n$.

In the same way one can prove the following theorem.

THEOREM 2. *Let $a_k > 0$, ($k = 1, 2, \dots, n$) be non-decreasing in mean. Then*

(i) *For all p , $0 < p < 1$, inequality (10) holds (here $C_k > 0$).*

For all p , $1 < p < 2$, the opposite inequality holds in (10) (here $C_k < 0$).

(ii) *For all p , $p > 2$, inequality (8) holds (with $D_k < 0$).*

Equalities hold in (i) and (ii) if and only if $a_1 = a_2 = \dots = a_n$.

REMARK 1. It is easy to check that for $p = 1$ and $p = 2$ we have equalities in (8) and (10).

Following the arguments used in the proofs of Theorem 1 and Theorem 2 one can prove the following theorem.

THEOREM 3. (a) *Let $(a_n > 0, n \in \mathbb{N})$ be non-increasing in mean. Then:*

(i) *For all p , $0 < p < 1$, we have*

$$\left(\sum_{n=1}^{\infty} a_n \right)^p \leq \sum_{n=1}^{\infty} [n^p - (n-1)^p] a_n^p + \sum_{n=2}^{\infty} \frac{1}{n} a_n^{-1} D_n (n a_n - S_n), \quad (18)$$

where D_n ($n \geq 2$) is given by (9) (here $D_n > 0$).

For all p , $1 < p < 2$, the opposite inequality holds in (18) (here $D_n < 0$).

(ii) *For all p , $p > 2$, we have*

$$\left(\sum_{n=1}^{\infty} a_n \right)^p \geq \sum_{n=1}^{\infty} [n^p - (n-1)^p] a_n^p + \sum_{n=2}^{\infty} a_n S_n^{p-2} C_n (n a_n - S_n), \quad (19)$$

where C_n ($n \geq 2$) is given by (11) (here $C_n < 0$).

(b) *Let $(a_n > 0, n \in \mathbb{N})$ be non-decreasing in mean.*

(iii) *For all p , $0 < p < 1$, inequality (19) holds (here $C_n > 0$).*

For all p , $1 < p < 2$, the opposite inequality holds in (19) (here $C_n < 0$).

(iv) *For all p , $p > 2$, inequality (18) holds (with $D_n < 0$).*

REMARK 2. For $p = 1$ and $p = 2$ we have equalities in (18) and (19).

The next theorem is a consequence of Theorem 1 and Theorem 2.

THEOREM 4. *Let $(p_k > 0, k = 1, 2, \dots, n)$ be a probability distribution with entropy $H = -\sum_{k=1}^n p_k \log p_k$ and let $P_k = \sum_{i=1}^k p_i$ ($k = 1, 2, \dots, n$).*

(i) *If the sequence $(p_k, k = 1, 2, \dots, n)$ is non-increasing in mean, then we have*

$$\sum_{k=2}^n F(k-1)p_k \leq H + \sum_{k=2}^n \frac{1}{k} p_k^{-1} (kp_k - P_k)(P_k \log P_k - P_k \log P_{k-1} - p_k). \tag{20}$$

(ii) *If the sequence $(p_k, k = 1, 2, \dots, n)$ is non-decreasing in mean, then we have*

$$H + \sum_{k=2}^n p_k P_k^{-1} (kp_k - P_k)[k \log k - k \log(k-1) - 1] \leq \sum_{k=2}^n F(k-1)p_k. \tag{21}$$

Equalities hold in (20) and (21) if $p_k = 1/n$ ($k = 1, 2, \dots, n$).

PROOF. (i) We can use Theorem 1 for $1 < p < 2$, $a_k = p_k$ and $S_k = P_k$ ($k = 1, 2, \dots, n$). Set

$$a = a(p) = \sum_{k=1}^n [k^p - (k-1)^p] p_k^p,$$

$$b = b(p) = \sum_{k=2}^n \frac{1}{k} p_k^{-1} (kp - P_k)(p p_k P_{k-1}^{p-1} + P_{k-1}^p - P_k^p).$$

It is obvious that $\lim_{p \rightarrow 1} a(p) = 1$, $\lim_{p \rightarrow 1} b(p) = 0$. It follows from (8) that

$$a + b \leq 1, \tag{22}$$

and therefore

$$\frac{\log a}{p-1} \leq \frac{\log(1-b)}{p-1}. \tag{23}$$

In the limiting case where $p \rightarrow 1+$, the left-hand side of (23) gives

$$\sum_{k=2}^n F(k-1)p_k + \sum_{k=1}^n p_k \log p_k,$$

while the right-hand side of (23) becomes

$$\sum_{k=2}^n \frac{1}{k} p_k^{-1} (kp_k - P_k)(P_k \log P_k - P_k \log P_{k-1} - p_k),$$

together forming inequality (20).

(ii) We use Theorem 2 for $1 < p < 2$, $a_k = p_k$ and $S_k = P_k$ ($k = 1, 2, \dots, n$). Set

$$c = c(p) = \sum_{k=2}^n p_k P_k^{p-2} (kp_k - P_k) k^{1-p} [p(k-1)^{p-1} + (k-1)^p - k^p].$$

It is obvious that $\lim_{p \rightarrow 1} c(p) = 0$. It follows from (10) that

$$a + c \geq 1, \quad (24)$$

and therefore

$$\frac{\log a}{p-1} \geq \frac{\log(1-c)}{p-1}. \quad (25)$$

As we have already seen, in the limiting case where $p \rightarrow 1+$, the left-hand side of (25) gives $\sum_{k=2}^n F(k-1)p_k - H$, while the right-hand side of (25) becomes

$$\sum_{k=2}^n p_k P_k^{-1} (kp_k - P_k) [k \log k - k \log(k-1) - 1],$$

together forming inequality (21).

The last statement follows from the fact that $p_k = 1/n$ ($k = 1, 2, \dots, n$) is equivalent to $P_k = kp_k$ ($k = 1, 2, \dots, n$) and in that case we easily find $\sum_{k=2}^n F(k-1)p_k = H = \log n$.

REMARK 3. We can get inequalities (20) and (21) by the use of Theorem 1 and Theorem 2 for the case $0 < p < 1$ and by taking limits as $p \rightarrow 1-$.

References

- [1] J.-P. Allouche, M. Mendes France and G. Tenenbaum, "Entropy: An inequality", *Tokyo J. Math.* **11** (1988) 323–328.
- [2] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1952).
- [3] C. Jardaš, J. Pečarić, R. Roki and N. Sarapa, "On an inequality for the entropy of a probability distribution", submitted for publication.
- [4] C. Jardaš, J. Pečarić, R. Roki and N. Sarapa, "On some inequalities for series and probability entropies", submitted for publication.
- [5] J. Pečarić and L.-E. Persson, "On an inequality of Hardy–Littlewood–Pólya", *Math. Gazette* **79** (1995) 383–385.