International Journal of Pure and Applied Mathematics

Volume 1 No. 4 2002, 379-387

ON SOME BOUNDS IN CODING THEORY

Y.J. Cho[1 §], M. Matić[2], J. Pečarić[3]

[1]Dept. of Mathematics
College of Education
Gyeongsang National University
Chinju 660-701, KOREA

[2]Dept. of Mathematics
University of Split
R. Boškovića bb, 21000 Split, CROATIA

[3]Faculty of Textile Technology
University of Zagreb
Pierottijeva 6, 10000 Zagreb, CROATIA

Abstract: In this paper, we prove some extensions of the results given in N. M. Dragomir et al. [2] and S. Roman [1] and some related results in coding theory.

AMS Subject Classification: 26A33
Key Words: Jensen's inequality, noiseless coding theorem, average codeword length

1. Introduction

Let us consider an encoding scheme $(c_1, c_2, \cdots, c_n)$ for a probability distribution $(p_1, p_2, \cdots, p_n)$. Recall that the average codeword length of an encoding scheme $(c_1, c_2, \cdots, c_n)$ for $(p_1, p_2, \cdots, p_n)$ is

Received: February 28, 2002

§Correspondence author

$$\mathrm{AveLen}(c_1, c_2, \cdots, c_n) = \sum_{i=1}^{n} p_i \mathrm{len}(c_i),$$

where $\mathrm{len}(c_i)$ $(i = 1, 2, \cdots, n)$ denotes the length of c_i, and the r-ary entropy of a probability distribution (or of a source) is given by

$$H_r(p_1, p_2, \cdots, p_n) = \sum_{i=1}^{n} p_i \log_r \frac{1}{p_i}.$$

The following result is well-known (see [1, Theorem 2.3.1]):

Theorem 1.1. *Let $C = (c_1, c_2, \cdots, c_n)$ be an instantaneous (decipherable) encoding scheme for $P = (p_1, p_2, \cdots, p_n)$. Then we have*

$$H_r(p_1, p_2, \cdots, p_n) \leq \mathrm{AveLen}(c_1, c_2, \cdots, c_n) \tag{1.1}$$

with the equality, if and only if $\ell_i = \log_r(1/p_i)$ for all $i = 1, 2, \cdots, n$, where ℓ_i $(i = 1, 2, \cdots, n)$ denotes the length $\mathrm{len}(c_i)$.

We shall use the notation $\mathrm{MinAveLen}_r(p_1, p_2, \cdots, p_n)$ to denote the minimum average codeword length among all r-ary instantaneous encoding schemes for the probability distribution $P = (p_1, p_2, \cdots, p_n)$.

The following noiseless coding theorem is well-known (see [1, Theorem 2.3.2]):

Theorem 1.2. *For any probability distribution $P = (p_1, p_2, \cdots, p_n)$, we have*

$$\begin{aligned} H_r(p_1, p_2, \cdots, p_n) &\leq \mathrm{MinAveLen}_r(p_1, p_2, \cdots, p_n) \\ &< H_r(p_1, p_2, \cdots, p_n) + 1. \end{aligned} \tag{1.2}$$

In this paper, we give some extensions of the above results and some related results. Our results are, also, improvements of the results given in [2].

2. Preliminary Results

In this section, we give some preliminary results before proving our main theorems.

Lemma 2.1. *Let p_i and q_i be strictly positive real numbers for $i = 1, 2, \cdots, n$. If $P_n = \sum_{i=1}^{n} p_i$ and $Q_n = \sum_{i=1}^{n} q_i$,, then we have*

$$\begin{aligned} \frac{1}{\ln r}(P_n - Q_n) &\le P_n \log_r \left(\frac{P_n}{Q_n}\right) \\ &\le \sum_{i=1}^{n} p_i \log_r \frac{p_i}{q_i} \\ &\le P_n \log_r \left(\frac{1}{P_n} \sum_{i=1}^{n} \frac{p_i^2}{q_i}\right) \\ &\le \frac{1}{\ln r} \sum_{i=1}^{n} \left(\frac{p_i}{q_i} - 1\right) p_i, \end{aligned} \tag{2.1}$$

where $r \in \mathbb{R}$ with $r > 1$. The equality holds in all inequalities simultaneously, if and only if $p_i = q_i$ for $i = 1, 2, \cdots, n$.

Proof. The well known Jensen's inequality for a concave function $x \mapsto \log_r x$ states

$$\log_r \left(\frac{1}{P_n} \sum_{i=1}^{n} p_i x_i\right) \ge \frac{1}{P_n} \sum_{i=1}^{n} p_i \log_r x_i \tag{2.2}$$

with the equality, if and only if $x_1 = x_2 = \cdots = x_n$. Setting $x_i = q_i/p_i$ $(i = 1, 2, \cdots, n)$ in (2.2), we get

$$\log_r \left(\frac{Q_n}{P_n}\right) \ge \frac{1}{P_n} \sum_{i=1}^{n} p_i \log_r \frac{q_i}{p_i}, \tag{2.3}$$

which is equivalent to the second inequality in (2.1). The first inequality in (2.1) is a simple consequence of the elementary inequality

$$\log_r x \le \frac{1}{\ln r}(x - 1), \tag{2.4}$$

which holds for all $x > 0$ with the equality, if and only if $x = 1$. Namely, we have

$$\begin{aligned} P_n \log_r \left(\frac{P_n}{Q_n}\right) &= -P_n \log_r \left(\frac{Q_n}{P_n}\right) \\ &\ge -P_n \frac{1}{\ln r} \left(\frac{Q_n}{P_n} - 1\right) \\ &= \frac{1}{\ln r}(P_n - Q_n). \end{aligned} \tag{2.5}$$

On the other hand, (2.2) for $x_i = p_i/q_i$ $(i = 1, 2, \cdots, n)$ becomes

$$\log_r\left(\frac{1}{P_n}\sum_{i=1}^{n}\frac{p_i^2}{q_i}\right) \geq \frac{1}{P_n}\sum_{i=1}^{n} p_i \log_r \frac{p_i}{q_i}, \tag{2.6}$$

which is equivalent to the third inequality in (2.1). The last inequality in (2.1) is obtained using (2.4), similarly as (2.5) is obtained. If $p_i = q_i$ for all $i = 1, \cdots, n$, then, obviously, we have equalities throughout in (2.1).

Conversely, if we have an equality either in the second inequality in (2.1) or in the third one, then we have an equality either in (2.3) or in (2.6), respectively. This, in both cases, implies that $p_i/q_i = \alpha$ $(i = 1, 2, \cdots, n)$ for some constant $\alpha \in \mathbb{R}$, and therefore $P_n = \alpha Q_n$. If, additionally, the equality holds either in the first inequality in (2.1) or in the last one, then we get $P_n = \alpha P_n$, which implies $\alpha = 1$, that is, $p_i = q_i$ for all $i = 1, \cdots, n$. Thus the equalities hold in all inequalities in (2.1) simultaneously, if and only if $p_i = q_i$ for all $i = 1, 2, \cdots, n$. This completes the proof. □

Corollary 2.2. *Let $P = (p_1, p_2, \cdots, p_n)$ be a probability distribution, i.e. $p_i \in (0, 1]$ $(i = 1, \cdots, n)$ and $\sum_{i=1}^{n} p_i = 1$. Let $Q = (q_1, q_2, \cdots, q_n)$ have the property that $q_i \in (0, 1]$ and $\sum_{i=1}^{n} q_i \leq 1$. Then we have*

$$\begin{aligned} 0 &\leq \frac{1}{\ln r}\left(1 - \sum_{i=1}^{n} q_i\right) \\ &\leq -\log_r\left(\sum_{i=1}^{n} q_i\right) \\ &\leq \sum_{i=1}^{n} p_i \log_r \frac{1}{q_i} - \sum_{i=1}^{n} p_i \log_r \frac{1}{p_i} \\ &\leq \log_r\left(\sum_{i=1}^{n}\frac{p_i^2}{q_i}\right) \\ &\leq \frac{1}{\ln r}\left(\sum_{i=1}^{n}\frac{p_i^2}{q_i} - 1\right), \end{aligned} \tag{2.7}$$

where $r \in \mathbb{R}$ with $r > 1$. The equalities hold in all the inequalities in (2.7) simultaneously, if and only if $p_i = q_i$ for all $i = 1, 2, \cdots, n$.

Proof. The proof is obvious by Lemma 2.1, taking into account, that $\sum_{i=1}^{n} p_i = 1$ and $\sum_{i=1}^{n} q_i \leq 1$. □

Remark 1. The above results are improvements of the related results from [2], while the above Corollary 2.2 is a further improvement of Lemma 1.2.2 from the book [1], which plays a very important role in obtaining the basic inequalities for entropy, conditional entropy and mutual information, etc. In fact, it is an refinement of the well-known Shannon's inequality (see [3, Chapter XXIII]).

3. Main Results

Now, we give our main result in this paper and some related results.

Theorem 3.1. *Let $C = (c_1, c_2, \cdots, c_n)$ and $P = (p_1, p_2, \cdots, p_n)$ be as in Theorem 1.1. Then we have the following inequalities:*

$$\begin{aligned} 0 &\leq \frac{1}{\ln r}\left(1 - \sum_{i=1}^{n} \frac{1}{r^{\ell_i}}\right) \\ &\leq -\log_r\left(\sum_{i=1}^{n} \frac{1}{r^{\ell_i}}\right) \\ &\leq \mathrm{AveLen}(c_1, c_2, \cdots, c_n) - H_r(p_1, p_2, \cdots, p_n) \\ &\leq \log_r\left(\sum_{i=1}^{n} p_i^2 r^{\ell_i}\right) \\ &\leq \frac{1}{\ln r}\left(\sum_{i=1}^{n} p_i^2 r^{\ell_i} - 1\right). \end{aligned} \tag{3.1}$$

The equalities hold in all the inequalities in (3.1) *simultaneously, if and only if $\ell_i = \log_r(1/p_i)$ for $i = 1, 2, \cdots, n$.*

Proof. The conclusions follow from Corollary 2.2 for $q_i = 1/r^{\ell_i}$ $(i = 1, 2, \cdots, n)$, since we have $q_i \in (0, 1]$ and $\sum_{i=1}^{n} q_i = \sum_{i=1}^{n} 1/r^{\ell_i} \leq 1$ by Kraft's theorem (see [1, Theorem 2.1.2]). $\square$

Theorem 3.2. *Let the assumptions of Theorem 3.1 be satisfied and let*

$$0 < m \leq p_i r^{\ell_i} \leq M \quad (i = 1, 2, \cdots, n). \tag{3.2}$$

Then we have

$$\max\{0, \log_r m\} \leq \text{AveLen}(c_1, c_2, \cdots, c_n) - H_r(p_1, p_2, \cdots, p_n) \leq \log_r M. \quad (3.3)$$

Proof. From (3.2), it follows that

$$mp_i \leq p_i^2 r^{\ell_i} \leq Mp_i \qquad (i = 1, 2, \cdots, n)$$

and so, since $\sum_{i=1}^n p_i = 1$, we have

$$m \leq \sum_{i=1}^n p_i^2 r^{\ell_i} \leq M. \tag{3.4}$$

Similarly, from (3.2), we have

$$\frac{p_i}{M} \leq r^{-\ell_i} \leq \frac{p_i}{m} \quad (i = 1, 2, \cdots, n),$$

which implies that

$$\frac{1}{M} \leq \sum_{i=1}^n r^{-\ell_i} \leq \frac{1}{m}. \tag{3.5}$$

Further, by Kraft's theorem, we have

$$\sum_{i=1}^n r^{-\ell_i} \leq 1. \tag{3.6}$$

Now, the inequality (3.6) and the second inequality in (3.5) yield the estimate

$$-\log_r \left(\sum_{i=1}^n r^{-\ell_i} \right) \geq \max\{0, \log_r m\},$$

which together with the third inequality in (3.1) give the first inequality in (3.3).

The second inequality in (3.3) is a simple consequence of the second inequality in (3.4) and the fourth inequality in (3.1). This completes the proof.□

Theorem 3.3. *Let $P = (p_1, p_2, \cdots, p_n)$ be a given probability distribution, and let $r \in \mathbb{N}$ with $r \geq 2$. If $\varepsilon > 0$ is given, and there exist natural numbers $\ell_1, \ell_2, \cdots, \ell_n$ such that*

$$\log_r\left(\frac{1}{p_i}\right) \leq \ell_i \leq \log_r\left(\frac{1}{p_i}\right) + \varepsilon \tag{3.7}$$

for all $i \in \{1, 2, \cdots, n\}$, then there exists an instantaneous r-ary code $C = (c_1, c_2, \cdots, c_n)$ with codeword lengths $\mathrm{len}(c_i) = \ell_i$ $(i = 1, 2, \cdots, n)$, such that

$$\begin{aligned} H_r(p_1, p_2, \cdots, p_n) \leq \mathrm{AveLen}(c_1, c_2, \cdots, c_n) \\ \leq H_r(p_1, p_2, \cdots, p_n) + \varepsilon. \end{aligned} \tag{3.8}$$

Proof. Note that (3.7) is equivalent to the following:

$$1 \leq p_i r^{\ell_i} \leq r^{\varepsilon},$$

which is the condition (3.2) with $m = 1$ and $M = r^{\varepsilon}$. From the second inequality in (3.5), we have

$$\sum_{i=1}^{n} r^{-\ell_i} \leq 1$$

and, by Kraft's theorem, there exists an instantaneous r-ary code $C = (c_1, c_2, \cdots, c_n)$ such that $\mathrm{len}(c_i) = \ell_i$ $(i = 1, 2, \cdots, n)$. Obviously, it follows from (3.3) that the inequalities (3.8) are valid. This completes the proof. $\square$

Remark 2. In [2], a result similar to that in Theorem 3.3 was proved with the condition:

$$\log_r\left(\frac{1}{p_i}\right) \leq \ell_i \leq \log\left(\frac{1 + \varepsilon \ln r}{p_i}\right)$$

for $i = 1, 2, \cdots, n$ instead of the condition (3.7). Using (2.4), it is easy to show that

$$\log_r\left(\frac{1 + \varepsilon \ln r}{p_i}\right) < \log_r\left(\frac{1}{p_i}\right) + \varepsilon.$$

Now, we give an extension of Theorem 1.2 by using Theorem 3.3 as follows:

Theorem 3.4. *Let r be a given natural number and $\varepsilon \in (0,1)$. If a probability distribution $P = (p_1, p_2, \cdots, p_n)$ satisfies the condition that every closed interval*

$$I_i = \left[\log_r\left(\frac{1}{p_i}\right), \log_r\left(\frac{1}{p_i}\right) + \varepsilon\right] \tag{3.9}$$

for $i \in \{1, 2, \cdots, n\}$ contains at least one natural number ℓ_i. Then for the probability distribution P, we have

$$\begin{aligned} H_r(p_1, p_2, \cdots, p_n) &\le \mathrm{MinAveLen}_r(c_1, c_2, \cdots, c_n) \\ &\le H_r(p_1, p_2, \cdots, p_n) + \varepsilon. \end{aligned} \tag{3.10}$$

Proof. Under the given assumptions we have

$$\sum_{i=1}^{n} \frac{1}{r^{\ell_i}} \le \sum_{i=1}^{n} p_i = 1$$

and, by Kraft's theorem, there exists an instantaneous code $C = (c_1, c_2, \cdots, c_n)$ such that $\mathrm{len}(c_i) = \ell_i$ for $i \in \{1, 2, \cdots, n\}$. For this code the condition (3.7) is valid and so, by Theorem 3.3, we have the inequalities (3.8). Taking the infimum in the inequalities (3.8) over all r-ary instantaneous codes, we get the inequalities (3.10). This completes the proof. □

Remark 3. In [2] the above result was proved with the shorter intervals

$$I_i = \left[\log_r\left(\frac{1}{p_i}\right), \log_r\left(\frac{1+\varepsilon}{p_i}\right)\right]$$

for $i \in \{1, 2, \cdots, n\}$.

Remark 4. We can take $\varepsilon = 1$ in the above theorem. It is clear, that we can find always a natural number ℓ_i such that $\ell_i \in I_i$ $(i = 1, 2, \cdots, n)$ in that case.

Theorem 3.5. *Let a_i $(i = 1, 2, \cdots, n)$ be a given natural numbers. If p_i $(i = 1, 2, \cdots, n)$ are such that*

$$\frac{1}{r^{a_i}} \le p_i \le \frac{r^{\varepsilon}}{r^{a_i}} \tag{3.11}$$

for $i = 1, 2, \cdots, n$ and $\sum_{i=1}^{n} p_i = 1$, then there exists an instantaneous code $C = (c_1, c_2, \cdots, c_n)$ with $\mathrm{len}(c_i) = a_i$, such that the inequalities (3.8) hold for the *probability distribution* $P = (p_1, p_2, \cdots, p_n)$. *Furthermore, for the distribution, we have the inequalities* (3.10).

Proof. The condition (3.11) is equivalent to the following:

$$\log_r\left(\frac{1}{p_i}\right) \le a_i \le \log_r\left(\frac{1}{p_i}\right) + \varepsilon$$

for $i = 1, 2, \cdots, n$, and so $a_i \in I_i$. Thus, applying the above results, we have the desired conclusion. This completes the proof. $\square$

Acknowledgements

The first author was supported by grant of the Korea Research Foundation (KRF-2001-005-D00002).

References

[1] S. Roman, *Coding and Information Theory*, Springer-Verlag, New York-Berlin-Heidelberg (1992).

[2] N.M. Dragomir, S.S. Dragomir, An inequality for logarithms and its application in coding theory, To appear in: *Indian J. of Math.*

[3] D.S. Mitrinović, J. E. Pečarić, A. M. Fink, *Classical and New Inequalities in Analysis*, Kluwer Academic Publishers, Dordrecht (1993).