TAMKANG JOURNAL OF MATHEMATICS
Volume 22, Number 4, Winter 1991

LOWER BOUNDS ON $L^t_{1:1}(D)$ IN TERMS OF RENYI ENTROPY

RAJANI SHARMA, PRITI GUPTA AND R. K. TUTEJA

Abstract. In this paper we obtain the lower bounds for the exponentiated mean codeword length (as defined by Campbell [4]) for one-one codes of size D by using the functions which represent possible transformations from one-one codes of size D to uniquely decodable codes.

1. Introduction

The average length of a noiseless uniquely decodable code for a discrete random variable X satisfies

$$H(X) + 1 > L_{UD} \geq H(X), \tag{1.1}$$

where $H(X)$ is the Shannon's entropy of the random variable X. Shannon's restriction of encoding or description of X to prefix codes is highly motivated by the implicit assumption that the descriptions will be concatenated and must be uniquely decodable. Since there is the same set of allowed codeword lengths for uniquely decodable and instantaneous codes c.f. [1], [2], the mean codeword length L is the same for both set of codes. In some communication situations, a single random variable X is being smitted tran instead of a sequence of random variables. For this, Leung-Yan-Cheong and Cover [3] taken one to one codes, i.e. the codes which assign different binary codeword to each outcome of the random variable, without regard to the constraint that the concatenation of the descriptions must be uniquely decodable.

Campbell [4] introduced the *onentiated exp mean codeword length,* given by

$$L^t_{UD} = \frac{1}{t}\log\left(\sum_{i=1}^{n} p_i D^{t\ell_i}\right) \tag{1.2}$$

where D represent the size of the code alphabet and ℓ_i, $i = 1, 2, \ldots, n$ are the lengths of the codewords associated with the values of X and proved the following "noiseless coding theorem."

$$H_\alpha(X) + 1 > L^t_{UD} \geq H_\alpha(X), \tag{1.3}$$

Receive March 25, 1991.

Where $H_\alpha(X)$ is the Renyi entropy [5] of order-α, with $\alpha = \frac{1}{1+t}$, given by

$$H_\alpha(X) = \frac{1}{1-\alpha}\log\left(\sum_{i=1}^{n} p_i^\alpha\right); \qquad \alpha \neq 1,\ \alpha > 0. \tag{1.4}$$

In limiting case, as $t \to 0$, L^t_{UD} reduces to ordinary mean codeword length L.

Kieffer [6], by defining a class of decision rules, showed that $H_\alpha(X)$ is the best decision rule for deciding which of two sources can be coded with smaller expected cost for sequences of length n, as $n \to \infty$, where the cost of encoding a sequence is assumed to be a function only of the codeword length. Jelinek [7] showed that coding with respect to L^t is useful in minimizing the problem of buffer overflow which occurs when the source symbols are being produced at a fixed rate, and the codewords must be stored temporarily in a finite buffer. A simple generalisation of Huffman algorithm solves the problem of minimizing L^t, was studied by Parker [8]. The determination of L^t along with its significance to the minimal expense coding problem is given by Aczel [9].

In this paper, we will define the codes which assign D alphabet one to one codeword to each outcome of the random variable. By defining the transformations from 1 : 1 to uniquely decodable codes for D alphabet, we are obtaining lower bounds for $L^t_{1:1}(D)$ in terms of (1.4).

In limiting case as $t \to 0$, and the codes are binary then results obtained reduces to that of Leung-Yan-Cheong and Cover [3].

2. Some Possible Transformations from 1 : 1 to UD Codes of Size D-Alphabet

Let $X = \{x_1, x_2, \ldots, x_n\}$ be an experiment which take finite number of values with probability distribution $P = \{p_1, p_2, \ldots, p_n\} \in \Delta_n$. where $\Delta_n = \{P = (p_1, \ldots, p_n);\ p_i > 0,\ \sum_{i=1}^{n} p_i = 1\}$ be the set of all complete finite discrete probability distribution and $p_1 \geq p_2 \geq \ldots \geq p_n$. Let ℓ_i, $i = 1, 2, \ldots, n$ be the length of condewords in the best 1 : 1 code alphabet of order D, i.e. $\{0, 1, \ldots\ldots (D-1);\ 00, \ldots (D-1)(D-1);\ (0\,0\,0), \ldots\ldots\}$ for encoding the random variable X.

When the code alphabet is of order D, then the set of available codewords is

$$\{0, 1, \ldots\ (D-1);\ 00, 01, \ldots, (D-1)(D-1);\ \ldots\},$$

and by inspection, we precisely have

$$\ell_1 = 1,\ \ell_2 = 1, \ldots\ldots, \ell_D = 1,\ \ell_{D+1} = 2, \ldots\ldots, \ell_{D(D+1)} = 2, \ldots\ldots \text{ etc.}$$

and thus

$$\ell_i = \left\lceil \log\left(\frac{D-1}{D} i + 1\right)\right\rceil, \tag{2.1}$$

where $\lceil X \rceil$ denotes the smallest integer greater than or equal to X.

If f be any function such that $\sum_i D^{-f(\ell_i)} \le 1$. Then from Kraft's inequality, the set of lengths $\{\lceil f(\ell_i)\rceil\}$ yields acceptable word lengths for a prefix (or uniquely decodable) code. If f is integer valued and $\sum_i D^{-f(\ell_i)} > 1$, then $\{f(\ell_i)\}$ can not yield a prefix code.

Theorem 1. *The following functions represent possible transformations from* $1:1$ *to* UD *codes.*

I) $f(\ell_i) = \ell_i + a\lceil\log \ell_i\rceil + \log(\frac{D^a - D + 1}{D^a - D})$, *where* $a > 1$, *and* $D \ge 2$.

II) $f(\ell_i) = \ell_i + D\lfloor\log(\ell_i + 1)\rfloor$.

Proof. (I) Defining

$$S = \sum_{i=1}^{\infty} D^{-f(\ell_i)},$$

$$= \sum_{i=1}^{\infty} D^{-(\ell_i + a\lceil\log \ell_i\rceil + r)},$$

$$\text{where } a > 1, \quad \text{and} \quad r = \log(\frac{D^a - D + 1}{D^a - D}),$$

$$= \sum_{i=1}^{\infty} D^{-\ell_i} D^{-a\lceil\log \ell_i\rceil} D^{-r},$$

$$= D^{-r} \sum_{i=1}^{\infty} \frac{D^{-\ell_i}}{D^{a\lceil\log \ell_i\rceil}}.$$

But there are D^k, $1:1$ codewords of length k, therefore

$$S = D^{-r} \sum_{\ell=1}^{\infty} \frac{1}{D^{a\lceil\log \ell\rceil}},$$

$$= D^{-r}(\frac{1}{D^0} + \frac{D^0}{D^a} + \frac{D^1}{D^{2a}} + \frac{D^2}{D^{3a}} + \cdots + \frac{D^{k-1}}{D^{ka}} + \cdots),$$

$$= D^{-r}(1 + \frac{1}{D^a} + \frac{1}{D^{2a-1}} + \frac{1}{D^{3a-2}} + \cdots + \frac{1}{D^{ka-k+1}} + \cdots),$$

$$= D^{-r}(1 + \frac{\frac{1}{D^a}}{1 - \frac{1}{D^a - 1}}),$$

$$= D^{-r}(\frac{D^a - D + 1}{D^a - D}),$$

S diverges if $a \le 1$, to make $S \le 1$, it is sufficient (and necessary) to have $a > 1$, and $r \ge \log(\frac{D^a - D + 1}{D^a - D})$.

This completes the proof of (i).

II) Defining, in this case also.

$$\begin{aligned} S &= \sum_{i=1}^{\infty} D^{-f(\ell_i)} \\ &= \sum_{i=1}^{\infty} D^{-\ell_i} D^{-D\lfloor \log(\ell_i+1) \rfloor}, \end{aligned}$$

Using the fact that there are D^k, $1:1$ codewords of length k

$$\begin{aligned} S &= \sum_{\ell=1}^{\infty} \frac{1}{D^{D\lfloor \log(\ell+1) \rfloor}} \\ &= (\frac{1}{D^D})^D + (\frac{1}{D^{2D}})^{D^2} + (\frac{1}{D^{3D}})^{D^3} + \cdots + (\frac{1}{D^{kD}})^{D^k} + \ldots, \\ &= (\frac{1}{D^{D-1}}) + (\frac{1}{D^{2D-2}}) + (\frac{1}{D^{3D-3}}) + \cdots, +(\frac{1}{D^{kD-k}}) + \cdots, \\ &= (\frac{1}{D^{(D-1)}}) + (\frac{1}{D^{2(D-1)}}) + (\frac{1}{D^{3(D-1)}}) + \cdots + (\frac{1}{D^{k(D-1)}}) + \cdots, \\ S &= (\frac{\frac{1}{D^{D-1}}}{1 - \frac{1}{D^{D-1}}}), \\ &= (\frac{1}{D^{D-1}})(\frac{1}{\frac{D^{D-1}-1}{D^{D-1}}}), \\ &= (\frac{1}{D^{D-1}-1}) \text{ where } D \geq 2. \end{aligned}$$

If $D = 2$, $S = 1$, $D = 3$, $S < 1$, and so on.

Hence $S \leq 1$, for $D \geq 2$.

This completes the proof of (ii).

III) *Lower Bounds for* $L_{1:1}^t(D)$

3. Lower Bounds For $L_{1:1}^t(D)$

The exponentiated mean codeword length (1.2) of size D for $1:1$ codes is defined by

$$L_{1:1}^t(D) = \frac{1}{t}\log\Big[\sum_{i=1}^{n} p_i D^{t\lceil \log(\frac{D-1}{D}i+1)\rceil}\Big] \tag{3.1}$$

We will now prove the following theorem, which gives lower bounds on $L_{1:1}^t(D)$ in terms of Renyi entropy when the code alphabet are of size D. We now make use of theorem 1 to prove some lower bounds on $L_{1:1}^t(D)$ in terms of the Renyi entropy.

Theorem 2. *The exponentiated mean codeword length $L^t_{1:1}(D)$ of the best 1 : 1 code of size D, satisfies the following lower bounds*

I) $L^t_{1:1}(D) \geq H_\alpha(X) - \frac{\alpha}{1-\alpha}\log\Big[\sum_{i=1}^{n} p_i\{\lceil\log(\frac{D-1}{D}i+1)\rceil\}^{a\cdot\frac{1-\alpha}{\alpha}}\Big] - a - r$; *where* $a > 1$, *and* $r = \log(\frac{D^a - D + 1}{D^a - D})$.

II) $L^t_{1:1}(D) \geq H_\alpha(X) - \frac{\alpha}{1-\alpha}\log\Big[\sum_{i=1}^{n} p_i\{\lfloor\log(\frac{D-1}{D}i+1)\rfloor+1\}^{D\cdot\frac{1-\alpha}{\alpha}}\Big]$; *where* $D > 2$.

Proof. (1) From the fact that the exponentiated mean codeword length $L^t_{UD}(D)$ of uniquely decodable code is $\geq H_\alpha(X)$. So we can write,

$$\frac{\alpha}{1-\alpha}\log\Big\{\sum_{i=1}^{n} p_i D^{\ell_i(\frac{1-\alpha}{\alpha})}\Big\} \geq H_\alpha(X),$$

$$\log\Big\{\sum_{i=1}^{n} p_i D^{\ell_i(\frac{1-\alpha}{\alpha})}\Big\} \geq \frac{1-\alpha}{\alpha}H_\alpha(X),$$

$$\log\Big\{E(D^{\ell(\frac{1-\alpha}{\alpha})}\Big\} \geq \frac{1-\alpha}{\alpha}\cdot H_\alpha(X),$$

$$E(D^{\ell\cdot(\frac{1-\alpha}{\alpha})}) \geq D^{\frac{1-\alpha}{\alpha}\cdot H_\alpha(X)},$$

Now from theorem I(i), we have

$$E\Big\{D^{\frac{1-\alpha}{\alpha}(\ell + a\lceil\log\ell\rceil + r)}\Big\} \geq D^{\frac{1-\alpha}{\alpha}\cdot H_\alpha(X)},$$

$$E(D^{\frac{1-\alpha}{\alpha}\ell})\, E(D^{a\cdot\frac{1-\alpha}{\alpha}(1+\log\ell)})(D^{\frac{1-\alpha}{\alpha}\cdot r}) \geq D^{\frac{1-\alpha}{\alpha}\cdot H_\alpha(X)},$$

$$E(D^{\frac{1-\alpha}{\alpha}\ell})\, E(D^{a\cdot\frac{1-\alpha}{\alpha}(1+\log\ell)}) \geq D^{\frac{1-\alpha}{\alpha}\cdot(H_\alpha(X)-r)},$$

Form Jensen's inequality and convexity of $-\log\ell$, we have

$$E(D^{\frac{1-\alpha}{\alpha}\ell})\cdot E(D^{a\cdot\frac{1-\alpha}{\alpha}})\, E(D^{a\cdot\frac{1-\alpha}{\alpha}\log\ell}) \geq D^{\frac{1-\alpha}{\alpha}\cdot(H_\alpha(X)-r)},$$

$$E(D^{\frac{1-\alpha}{\alpha}\ell})\cdot E(D^{\log\ell^{\cdot a(\frac{1-\alpha}{\alpha})}})(D^{a\cdot\frac{1-\alpha}{\alpha}}) \geq D^{\frac{1-\alpha}{\alpha}\cdot(H_\alpha(X)-r)},$$

$$E(D^{\frac{1-\alpha}{\alpha}\ell})\cdot E(\ell^{\cdot a(\frac{1-\alpha}{\alpha})}) \geq D^{\frac{1-\alpha}{\alpha}\cdot(H_\alpha(X)-r-a)},$$

But $E(D^{\frac{1-\alpha}{\alpha}\ell}) = (D^{L^t_{1:1}(D)})^{\frac{1-\alpha}{\alpha}}$.

So we have

$$(D^{L^t_{1:1}(D)})^{\frac{1-\alpha}{\alpha}} \geq \frac{D^{\frac{1-\alpha}{\alpha}(H_\alpha(X)-a-r)}}{E(\ell^{a\cdot\frac{1-\alpha}{\alpha}})},$$

$$(D^{L^t_{1:1}(D)}) \geq \frac{D^{(H_\alpha(X)-a-r)}}{\{E(\ell^{a\cdot\frac{1-\alpha}{\alpha}})\}^{\frac{\alpha}{1-\alpha}}},$$

Taking logarithm of both sides, we have

$$L^t_{1:1}(D) \geq H_\alpha(X) - a - r - \frac{\alpha}{1-\alpha}\log\Big\{E(\ell^{a\cdot\frac{1-\alpha}{\alpha}})\Big\}.$$

$$L^t_{1:1}(D) \geq H_\alpha(X) - \frac{\alpha}{1-\alpha}\log\Big\{\sum_{i=1}^n p_i \ell_i^{a\cdot\frac{1-\alpha}{\alpha}}\Big\} - a - r.$$

Or

$$L^t_{1:1}(D) \geq H_\alpha(X) - \frac{\alpha}{1-\alpha}\Big\{\sum_{i=1}^n p_i\big(\lceil\log(\frac{D-1}{D}i+1)\rceil\big)^{a\cdot\frac{1-\alpha}{\alpha}}\Big\} - a - r, \text{ where } t = \frac{1-\alpha}{\alpha}.$$

Or

$$L^t_{1:1}(D) \geq H_\alpha(X) - \text{ “extra positive term.”}.$$

II) From the fact that exponentiated mean codeword length $L^t_{UD}(D)$ for a uniquely decodable codes $\geq H_\alpha(X)$. So we can write

$$\frac{\alpha}{1-\alpha}\log\Big\{\sum_{i=1}^n p_i D^{\frac{1-\alpha}{\alpha}\ell_i}\Big\} \geq H_\alpha(X),$$

$$\log\Big\{\sum_{i=1}^n p_i D^{\frac{1-\alpha}{\alpha}\ell_i}\Big\} \geq \frac{1-\alpha}{\alpha}\cdot H_\alpha(X),$$

$$\log\Big\{E(D^{\frac{1-\alpha}{\alpha}\ell})\Big\} \geq \frac{1-\alpha}{\alpha}\cdot H_\alpha(X),$$

$$E(D^{\frac{1-\alpha}{\alpha}\ell}) \geq D^{\frac{1-\alpha}{\alpha}\cdot H_\alpha(X)}.$$

From the theorem I(ii), we have

$$E\Big\{D^{\frac{1-\alpha}{\alpha}(\ell + D\lfloor\log(\ell+1)\rfloor)}\Big\} \geq D^{\frac{1-\alpha}{\alpha}H_\alpha(X)},$$

$$E\Big\{D^{\frac{1-\alpha}{\alpha}(\ell + D(\log(\ell+1))}\Big\} \geq D^{\frac{1-\alpha}{\alpha}H_\alpha(X)},$$

$$E\Big\{D^{\frac{1-\alpha}{\alpha}\ell}\Big\}E\Big\{D^{\frac{1-\alpha}{\alpha}\cdot D\cdot\log(\ell+1)}\Big\} \geq D^{\frac{1-\alpha}{\alpha}H_\alpha(X)},$$

But

$$E(D^{\frac{1-\alpha}{\alpha}\ell}) = (D^{L^t_{1:1}(D)})^{\frac{1-\alpha}{\alpha}}$$

So we have

$$(D^{L^t_{1:1}(D)})^{\frac{1-\alpha}{\alpha}}\cdot E\Big\{(\ell+1)^{D\cdot\frac{1-\alpha}{\alpha}}\Big\} \geq D^{\frac{1-\alpha}{\alpha}\cdot H_\alpha(X)},$$

$$(D^{L^t_{1:1}(D)})^{\frac{1-\alpha}{\alpha}} \geq \frac{D^{\frac{1-\alpha}{\alpha}H_\alpha(X)}}{E\{(\ell+1)^{D\cdot\frac{1-\alpha}{\alpha}}\}},$$

$$(D^{L^t_{1:1}(D)}) \geq \frac{D^{H_\alpha(X)}}{\left[E\{(\ell+1)^{D\cdot(\frac{1-\alpha}{\alpha})}\}\right]^{\frac{\alpha}{1-\alpha}}},$$

Taking logarithm of both sides, we have

$$L^t_{1:1}(D) \geq H_\alpha(X) - \frac{\alpha}{1-\alpha}\log\Big\{E(\ell+1)^{D\cdot\frac{1-\alpha}{\alpha}}\Big\},$$

$$L^t_{1:1}(D) \geq H_\alpha(X) - \frac{\alpha}{1-\alpha}\log\Big\{\sum_{i=1}^{n}p_i(\lfloor\log(\frac{D-1}{D}i+1)\rfloor+1)^{D\cdot\frac{1-\alpha}{\alpha}}\Big\}.$$

Or

$$L^t_{1:1}(D) \geq H_\alpha(X) - \text{"extra positive term."}$$

It is obvious from above that minimum bounds for one to one codes of size D alphabet are less than $H_\alpha(X)$. So the extra positive term is significant and cannot be improved. Due to this, obtained lower bounds in this paper for one to one codes of size D alphabet are better than that of uniquely decipherable codes.

Acknowkedgements:

The authoes are thankful to referee for his valuable comments.

References

1. Abramson, N: *Information Theory and Coding,* McGraw-Hill. New York (1963).
2. Ash, R: "Information Theory," *Interscience Pub.* New York (1965).
3. Leung-Yan-Cheong, S. K. and Cover, T.: "Some Equivalence Between Shannon Entropy and Kolmogrov Complexity." Vol. IT-24, No. 3, p.331-338, May, 1978.

4. Campbell, L. L.: "A Coding THeorem and Renyi's Entropy." *Information and Control,* Vol. 8, p.423-429, (1965).
5. Renyi, A.: "On Measures of Entropy and Information" Proc. 4th Berk. Symp. Math. Stat. Prob. Vol. I. P. 546-561, (1961).
6. Kieffer, J. C.: "Variable length Source Coding with a Cost Depending only on the Codeword Length." *Information and Control.* Vol. 41, p. 136-146, (1979).
7. Jelinek, F.: "Buffer Overflow in Variable Length Coding of Fixed Rate Sources." *IEEE Trans. on Information Theory,* Vol. IT-14. No. 3, p. 490-501, May, 1980.
8. Parker Jr., D. S.: "Conditions for Optimally of the Huffman Algorithm." *Siam J. Comput.* Vol. 9, No. 3, p. 470-489, (1980).
9. Aczel, J.: "Determination of all Additive Quasiarithmeti Mean Codeword Lengths." *Z. Wahrs. Verw. Geb.* Vol. 29, p. 351-360, (1974).

Department of Mathematics, Maharshi Dayanand University, Rohtak-124001 (INDIA).