Fundamentals of Information Theory

Homework 3

Problem 1 *(10 points)*

(a) *Find the entropy rate of the two-state Markov chain with transition matrix*

$$P = \begin{bmatrix} 1-p_{01} & p_{01} \\ p_{10} & 1-p_{10} \end{bmatrix}.$$

(b) *What values of p_{01}, p_{01} maximize the rate of part (a)?*

(c) *Find the entropy rate of the two-state Markov chain with transition matrix*

$$P = \begin{bmatrix} 1-p & p \\ 1 & 0 \end{bmatrix}.$$

(d) *Find the maximum value of the entropy rate of the Markov chain of part (c). We expect that the maximizing value of p should be less than $\frac{1}{2}$, since the 0 state permits more information to be generated than the 1 state.*

Problem 2 *(10 points)*

(a) *Are Shannon codes compact codes? Why?*

(b) *Explain why you can obtain a prefix code according to the shannon coding algorithm?*

Problem 3 *(10 points) Consider the following method for generating a code for a random variable X which takes on m values $\{1, 2, \ldots, m\}$ with probabilities $p_1, p_2, \ldots p_m$. Assume that the probabilities are ordered so that $p_1 \geqslant p_2 \geqslant \ldots \geqslant p_m$. Define*

$$F_i = \sum_{k=1}^{i-1} p_k,$$

the sum of the probabilities of all symbols less than i. Then the codeword for i is the number $F_i \in [0,1]$ rounded off to l_i bits, where $l_i = \left\lceil \log \frac{1}{p_i} \right\rceil$.

(a) *Show that the code constructed by this process is prefix-free and the average length satisfies*

$$H(X) \leq L < H(X) + 1$$

(b) *Construct the code for the probability distribution $(0.5, 0.25, 0.125, 0.125)$.*

Problem 4 *(10 points) Consider the random variable*

$$X = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\ 0.49 & 0.26 & 0.12 & 0.04 & 0.04 & 0.03 & 0.02 \end{pmatrix}.$$

(a) *Find a binary Huffman code for X.*

(b) *Find the expected codelength for this encoding.*

(c) *Find a ternary Huffman code for X.*

Problem 5 *(10 points) Bad wine. One is given 6 bottles of wine. It is known that precisely one bottle has gone bad (tastes terrible). From inspection of the bottles it is determined that the probability* p_i *that the* i*th bottle is bad is given by* $(p_1, p_2, \ldots, p_6) = (\frac{8}{23}, \frac{6}{23}, \frac{4}{23}, \frac{2}{23}, \frac{2}{23}, \frac{1}{23})$. *Tasting will determine the bad wine. Suppose you taste the wines one at a time. Choose the order of tasting to minimize the expected number of tastings required to determine the bad bottle. Remember, if the first* 5 *wines pass the test you don't have to taste the last.*

(a) *What is the expected number of tastings required?*

(b) *Which bottle should be tasted first?*

Now you get smart. For the first sample, you mix some of the wines in a fresh glass and sample the mixture. You proceed, mixing and tasting, stopping when the bad bottle has been determined.

(c) *What is the minimum expected number of tastings required to determine the bad wine?*

(d) *What mixture should be tasted first?*