Problem 1

Huffman Coding

Consider the chance variable $X$ that takes on the values $(x_1, x_2, \ldots, x_7)$ with respective probabilities $(0.49, 0.26, 0.12, 0.04, 0.04, 0.03, 0.02)$.

a) Find a binary Huffman code for $X$.

b) Find the expected code length for this encoding.

c) Find a ternary Huffman code for $X$.

Problem 2

Bad Codes

Which of these codes cannot be Huffman codes for any probability assignment?

a) $\{0, 10, 11\}$.

b) $\{00, 01, 10, 110\}$.

c) $\{01, 10\}$.

Problem 3

Optimal Codeword Lengths

Although the codeword lengths of an optimal variable length code are complicated functions of the message probabilities $\{p_1, p_2, \ldots, p_m\}$, it can be said that less probable symbols are encoded into longer codewords. Suppose that the message probabilities are given in decreasing order $p_1 > p_2 \geq \ldots \geq p_m$.

a) Prove that for any binary Huffman code, if the most probable message symbol has probability $p_1 > \frac{2}{5}$, then that symbol must be assigned a codeword of length 1.

b) Prove that for any binary Huffman code, if the most probable message symbol has probability $p_1 < \frac{1}{4}$, then that symbol must be assigned a codeword of length at least 2.
Problem 4

Let $X_1, X_2, \ldots$ be drawn IID according to the probability mass function $p(x), x \in \{1, 2, \ldots, m\}$. Thus, $p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p(x_i)$. We know that

$$-\frac{1}{n} \log p(X_1, X_2, \ldots, X_n) \to H(X)$$

in probability. Let $q(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} q(x_i)$, where $q$ is another probability mass function on $\{1, 2, \ldots, m\}$.

a) Evaluate $\lim_{n \to \infty} \left(-\frac{1}{n} \log q(X_1, X_2, \ldots, X_n)\right)$, where $X_1, X_2, \ldots$ are IID $\sim p(x)$, and express it as a function of relative entropies and entropies.

b) Now evaluate the limit of the log likelihood ratio $\frac{1}{n} \log \frac{p(X_1, X_2, \ldots, X_n)}{q(X_1, X_2, \ldots, X_n)}$ when $X_1, X_2, \ldots$ are IID $\sim p(x)$.

Problem 5

Proof of Theorem 3.3.1 in Cover & Thomas: High-Probability Sets and the Typical Set

Let $X_1, \ldots, X_n$ be a random sequence chosen IID $\sim P_X(\cdot)$. Fix some $\epsilon < \frac{1}{2}$, let $A^{(n)}_\epsilon$ denote the set of weakly typical sequences, and let $B^{(n)}_\delta$ be an arbitrary set of length-$n$ sequences such that

$$\Pr(B^{(n)}_\delta) > 1 - \delta.$$

a) For some $0 < \epsilon_1, \epsilon_2 < \frac{1}{2}$ and given any two sets $A, B$ such that $\Pr(A) > 1 - \epsilon_1$ and $\Pr(B) > 1 - \epsilon_2$ show that

$$\Pr(A \cap B) > 1 - \epsilon_1 - \epsilon_2.$$

Show that from this then follows that

$$\Pr(A^{(n)}_\epsilon \cap B^{(n)}_\delta) > 1 - \epsilon - \delta.$$

b) Justify each step in the chain of equalities/inequalities below:

$$1 - \epsilon - \delta < \Pr(A^{(n)}_\epsilon \cap B^{(n)}_\delta)$$

$$= \sum_{x \in A^{(n)}_\epsilon \cap B^{(n)}_\delta} P_X(x)$$

$$\leq \sum_{x \in A^{(n)}_\epsilon \cap B^{(n)}_\delta} 2^{-n(H(X)-\epsilon)}$$

$$= |A^{(n)}_\epsilon \cap B^{(n)}_\delta| \cdot 2^{-n(H(X)-\epsilon)}$$

$$\leq |B^{(n)}_\delta| \cdot 2^{-n(H(X)-\epsilon)}.$$

c) Complete the proof of Theorem 3.3.1 in Cover & Thomas (p. 63), i.e., show that for $0 < \delta < \frac{1}{2}$ and $\delta' > 0$

$$\frac{1}{n} \log |B^{(n)}_\delta| > H(X) - \delta'$$

for $n$ sufficiently large.
Problem 6

From AEP to Kraft’s Inequality

Recall the strong converse for the source coding theorem:

**Theorem 1.** Consider a sequence of source codes where the rate for encoding $n$ source symbols at a time is $\rho_n$. If

$$\lim_{n \to \infty} \rho_n < H(X),$$

(1)

where $H(X)$ is the entropy of the source, then the probability of successful decoding for these codes tends to zero as $n$ tends to infinity.

Use Theorem 1 to show that if $\ell_1, \ldots, \ell_d$ are the codeword lengths of a prefix-free one-to-variable code, then

$$\sum_{i=1}^{d} 2^{-\ell_i} \leq 1.$$  

(2)

**Hint:** Prove by contradiction: Assume that there exists a prefix-free code $C$ whose codeword lengths do not satisfy (2). Show that then there exists a source for which the expected codeword length of $C$ is smaller than the entropy of the source. Then use the Law of Large Numbers to show that, by encoding $n$ independent such source symbols at a time using extensions of $C$, we can satisfy (1) while making the probability of successful decoding tend to one as $n$ tends to infinity.