ECE 534: Elements of Information Theory, Fall 2013
Homework 5 - BONUS
Out: September 25, September 30, 2013.
Solutions

1. PROBLEM 5.8. *Huffman coding.* 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 code length for this encoding.

   (c) Find a ternary Huffman code for X.

2. PROBLEM 5.12. *Shannon codes and Huffman codes.* Consider a random variable X that takes on four values with probabilities $(\frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{12})$.

   (a) Construct a Huffman code for this random variable.

   (b) Show that there exist two different sets of optimal lengths for the codewords; namely, show that codeword length assignments $(1, 2, 3, 3)$ and $(2, 2, 2, 2)$ are both optimal.

   (c) Conclude that there are optimal codes with codeword lengths for some symbols that exceed the Shannon code length $\lceil \log 2 \frac{1}{p(x)} \rceil$

3. PROBLEM 5.17. *Data compression.* Find an optimal set of binary codeword lengths $l_1, l_2, \ldots$ (minimizing $\sum p_i l_i$) for an instantaneous code for each of the following probability mass functions:

   (a) $p = (10/41, 9/41, 8/41, 7/41, 7/41)$

   (b) $p = (\frac{9}{10}, (\frac{9}{10})(\frac{1}{10})^1, (\frac{9}{10})(\frac{1}{10})^2, (\frac{9}{10})(\frac{1}{10})^3, \ldots)$

4. PROBLEM 5.30. *Relative entropy is cost of miscoding.* Let the random variable X have five possible outcomes $\{1, 2, 3, 4, 5\}$. Consider two distributions $p(x)$ and $q(x)$ on this random variable.

   (a) Calculate $H(p)$, $H(q)$, $D(p||q)$, and $D(q||p)$.

   (b) The last two columns represent codes for the random variable. Verify that the average length of C_1 under p is equal to the entropy $H(p)$. Thus, C_1 is optimal

Symbol	$p(x)$	$q(x)$	$C_1(x)$	$C_2(x)$
1	1/2	1/2	0	0
2	1/4	1/8	10	100
3	1/8	1/8	110	101
4	1/16	1/8	1110	110
5	1/16	1/8	1111	111

Table 1: Problem 5.30

for p. Verify that C_2 is optimal for q.

(c) Now assume that we use code C_2 when the distribution is p. What is the average length of the codewords. By how much does it exceed the entropy p?

(d) What is the loss if we use code C_1 when the distribution is q?

5. PROBLEM 5.32. *Bad wine.* One is given six 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 $(p_1, p_2, ..., p_6) = (\frac{8}{23}, \frac{6}{23}, \frac{4}{23}, \frac{2}{23}, \frac{2}{23}, \frac{1}{23})$. Tasting will determine the bad wine. Suppose that 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 five wines pass the test, you dont 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.

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

(b) What mixture should be tasted first?