ECE6605
Information Theory

HW #4 : Due October 15, 2003

1) Huffman Coding. Find the binary Huffman code for an discrete memoryless source with $P = \left[\frac{5}{24},\frac{5}{24},\frac{4}{24},\frac{4}{24},\frac{3}{24},\frac{3}{24}\right]$ for the given number of consecutive source outputs. How close is the code to H(X) ? You may want to write a short program to do this.

a) n=1
b) n=2
c) n=3

2) Design a code using the Shannon-Fano-Elias coding scheme for a source with alphabet A = {a,b,c} where p(a) = 1/12, p(b) = 1/3, p(c) = 7/12. Compute the average codeword length and compare it to the entropy. (You will probably have to write a short C or Matlab program to get the binary expansion of the probabilities. Notice that according to the algorithm you only need a finite precision estimate of the fraction).

3) Can L=(1,2,2) be the word lengths of a binary Huffman code? How about (2,2,3,3)? Do the codewords of a binary Huffman code always meet the Kraft inequality with equality? If so, why? If not, give a counterexample?

4) Let X be a random variable with alphabet {1,2,3} and p(1)=0.5, p(2)=0.25, and p(3)=0.25. Suppose a Huffman code has codewords C(1) = 0, C(2)=10, C(3)=11. Let $X_1,X_2,...,X_n$ be an iid sequence according to this distribution and let $Z_1,Z_2,...,Z_n = C(X_1)C(X_2),...,C(X_n)$ be the string og binary symbols resulting from the concatenation of corresponding codewords. For example 133 becomes 01111.

a) Find the entropy rate H(X) and H(Z) in bits/symbol. Show that the sequence $Z_1,Z_2,...,Z_n$ is not compressible further.

b) Let the code from above be modified to be C(1)=00, C(2)=1, C(3)=01. Find the entropy rate H(Z). Explain the differences between the results in a) and the results in b)