ECE 776 - Midterm Spring 2015

Please provide detailed answers.

1. (4 points)

a. Build the Huffman code for a source with pmf $p(1) = 0.35,\ p(2) = 0.3,\ p(3) = 0.25,\ p(4) = 0.1$.

b. Evaluate the average length and compare with the entropy. Explain why the redundancy, i.e., the difference between the average length and the entropy, is not zero.

c. Evaluate the variance of the codeword lengths of the code designed at point a.

d. Consider the alternative Huffman code in which at the second step the symbols to be merged are chosen differently. Evaluate average and variance of the codeword lenghts.

2. (1 point) For Shannon's secrecy system, imposing only the constraint $H(X|Y,K) = 0$ (and not necessarily $I(X;Y) = 0$), prove the inequality $H(K) \geq H(X|Y)$, where K is the key, X is the plaintext and Y is the ciphertext (Hint: You can use the information diagram).

3. (1 point) Remembering that for a geometric random variable N we have $H(N) = E[N]$, find a fixed-to-variable prefix-free scheme that is optimal in terms of average codeword length for the encoding of a geometric random variable.

4. (3 points) Consider a binary test between the hypotheses

$$\mathcal{H}_0 \ : \ X^n, Y^n \sim p(x,y) \text{ i.i.d. (i.e., } p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i))$$

$$\mathcal{H}_1 \ : \ X^n, Y^n \sim p(x)p(y) \text{ i.i.d. (i.e., } p(x^n, y^n) = \prod_{i=1}^{n} p(x_i)p(y_i))$$

a. Write the general form of the optimal Neyman-Pearson test (as a function of the threshold).

b. Describe the asymptotic behavior of the random variable $\frac{1}{n} \log \frac{p(X^n,Y^n)}{p(X^n)p(Y^n)}$ for a pair of random processes X^n, Y^n with i.i.d. elements $(X_i,Y_i) \sim p(x,y)$ for $i = 1, ..., n$ (Hint: Note that $p(X^n) = \prod_{i=1}^{n} p(X_i)$ and $p(Y^n) = \prod_{i=1}^{n} p(Y_i)$).

c. Based on point b., propose a threshold for the test derived in part a. such that the probability of false alarm tends to zero as n goes to infinity.

5. (1 point) Find an upper bound on the number of perfectly random (i.e., i.i.d. Ber(1/2)) bits that can be extracted by means of a deterministic function of a discrete random variable X.