COSC 6222 :: Test #1

Attempt all questions. All questions have equal weight.

In class part: Time limit 80 minutes; open book, calculators permitted.
Take home part: Due Friday, October 17, 5 PM. Consulting with others is prohibited.

1. Let p(x,y) represent a joint probability of the binary random variables x and y, where p(0,0) = 1/2, p(1,0) = 1/4, p(0,1) = 1/4, and p(x,y) = 0 otherwise. Calculate:
   a. H(X)
   b. H(X|Y)
   c. H(X,Y)
   d. The Huffman code for the pair (X,Y).

2. Let f(x,y) and g(x,y) represent joint distributions on x and y. Let f(y|x), f(x|y), and f(y) represent the appropriate conditional and marginal densities obtained from f(x,y); and similarly for g(y|x),g(x|y), and g(y). Show that:

$$D(f(y \mid x) \parallel g(y \mid x)) \le D(f(x \mid y) \parallel g(x \mid y)) + D(f(y) \parallel g(y)).$$

3. Let {A, B} represent a set of symbols for a source X, where $Pr(A) = 0.8$ and $Pr(B) = 0.2$.
   a. For this source, what is the minimum average length of any uniquely decodable code using an encoding alphabet with k symbols?
   b. Calculate the quantity in part a for k = 2 (i.e., binary) and k = 3 (i.e., ternary).
   c. Give the binary Shannon-Fano-Elias code for the two-symbol extension of X (i.e., encoding AA, AB, BA, BB). Calculate the average code length.
   d. Repeat part c for a ternary Shannon-Fano-Elias code.

4. Let X = (X_1, X_2, ..., X_n) represent a sequence of iid random variables. Let C represent the set of all sequences X such that the probability of X is greater than p^*, i.e.,

$$C = \{X : p(X) \ge p^*\}.$$

   a. Show that $|C| \le 1/p^*$.
   b. For what values of p^* is it true that $Pr(X \in C) \rightarrow 1$ as n→∞?

5. Ahmed and Bob are playing a tennis match. The winner of a tennis match is the first player to win three "sets". If A and B represent sets won by Ahmed and Bob, respectively, let X represent the sequence of set victories, and let Y represent the number of sets in the match (e.g., we could have X = AABA and Y = 4, where Ahmed won all sets except the third). Ahmed is the better tennis player and wins sets with probability 3/4. Find:
   a. H(X)
   b. H(Y)
   c. H(Y|X)