ECE 563 Fall 2016 **Final Exam**

Date Assigned: 7 December 2015.
Date Due: 12 November 2015 by 5pm
Delivery Mechanism: Please turn in the solutions by 5pm to the TA Shaileshh Venkatakrishnan, whose office is in 130 CSL. You can also turn in typed up solutions by email.
Honor Code: The students are expected to work on these problems without consulting either technical material (books or online) or collaborating with fellow-students.

1. **Short Questions**.

   (a) Let X, Y, Z be Bernoulli(0.5) random variables such that $I(X;Y) = I(X;Z) = I(Y;Z) = 0$. Find the minimum value of $H(X, Y, Z)$ over all possible joint distributions of (X, Y, Z). Demonstrate a specific joint distribution that achieves the minimum value.

   (b) Let $\{X_n\}$ be a stationary Markov process. Show that $I(X_1; X_3) + I(X_2; X_4) \leq I(X_1; X_4) + I(X_2; X_3)$.

   (c) A random variable X takes on three distinct values with probabilities 0.6, 0.3 and 0.1.

      i. What are the lengths of the binary Huffman code?

      ii. What are the lengths of the Shannon code?

      iii. What is the smallest integer D such that the expected Shannon codeword length with a D-ary alphabet equals the expected Huffman codeword length with a D-ary alphabet?

   (d) Find the largest differential entropy among non-negative random variables X with mean equal to μ. *Hint*: Using the same idea as the calculation showing that Gaussians have the largest differential entropy among random variables with fixed variance, show that $h(E) - h(X) = D(f_E \| f_X)$ where E is the exponential random variable and f_X and f_E are the pdfs of X and E respectively.

   (e) A fair coin is flipped until the first head occurs. Let X denote the number of flips required.

      i. Find the entropy $H(X)$ in bits.

      ii. Find an "efficient" sequence of yes-no questions of the form, "Is X contained in the set S?". Compare $H(X)$ to the expected number of questions required to determine X.

      iii. Let Y denote the number of flips until the second head appears. Thus, for example, $Y = 5$ if the second head appears on the 5th flip. Argue that $H(Y) = H(X_1 + X_2) < H(X_1, X_2) = 2H(X)$ (by defining X_1, X_2 appropriately), and interpret in words.

   (f) Let X_1, X_2, X_3 be i.i.d. discrete random variables. Which is larger: $H(X_1 | X_1 + X_2 + X_3)$ or $H(X_1 + X_2 | X_1 + X_2 + X_3)$?

2. **Channel Capacity Computation**

   (a) Compute the capacity of the channel with an eight-letter input alphabet and nine-letter output alphabet and with the transition matrix

$$\begin{bmatrix} \frac{2}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \end{bmatrix}.$$

   (b) Consider a binary-input, ternary-output discrete memoryless channel with the following transition matrix

$$\begin{bmatrix} p_1 & p_2 & 1-p_1-p_2 \\ p_1 & 1-p_1-p_2 & p_2 \end{bmatrix}$$

   for some $p_1, p_2 \in (0,1)$. Compute a closed-form expression for the capacity in terms of p_1, p_2.

   (c) A discrete time memoryless channel has an input X constrained to the interval $(-0.5, 0.5)$ and has additive noise Z with uniform probability density over the interval $[-1, 1]$. The output $Y = X + Z$. Find the capacity of the channel and the input distribution that leads to it.

3. State true or false with a *succinct* and *sharp* explanation (without which no points are awarded). Each of the questions below are separate from each other.

   (a) There exists a discrete memoryless channel with a binary input alphabet and a quaternary output alphabet such that its capacity is equal to 1.5 bits per channel use.

   (b) For any two continuous random variables X and Y and any constant c it must be that $h(X+Y) \geq h(X+c)$.

   (c) For any zero-mean X, Y, Z with X and Y being Gaussian and X and Z having the same variance, it must be that $D(Z\|Y) \geq D(X\|Y)$.

   (d) For any random variables X, Y, Z defined on the same probability space, $I(X;Y) = 0$ means that $I(X;Y|Z) = 0$.

   (e) Consider binary random variables X, Y_1, Y_2 such that $I(X;Y_1) = I(X;Y_2) = 0$. Then $I(X;Y_1,Y_2) = 0$.

   (f) Consider binary random variables X, Y_1, Y_2 such that $I(X;Y_1) = I(X;Y_2) = 0$. Then $I(Y_1;Y_2) = 0$.

   (g) $H(X) \leq H(g(X))$ for any function $g(\cdot)$ and any discrete random variable X.

   (h) $h(X) \leq h(g(X))$ for any function $g(\cdot)$ and any continuous random variable X.

4. **Graph Entropy**.

While considering a special kind of communication problem involving data compression with an indistinguishability constraint, Janos Körner introduced a fundamental quantity called the *graph entropy*.

A *probabilistic graph* (G, P) is a graph $G = (V, E)$ with a probability distribution P on its vertices. Let $\mathcal{A}$ denote the collection of *maximal independent sets* of the graph G. (Recall that an *independent set* of a graph is a subset of its vertices no pair of which is connected by an edge; a maximal independent set is one that cannot be increased in size by the addition of another vertex. Note that maximal independent sets can be of different cardinalities.)

The graph entropy $H_G(P)$ of the probabilistic graph (G, P) is defined as follows: it is the minimum of $I(X; Y)$ such that X takes values in V, with distribution P, and Y takes values in $\mathcal{A}$, and $X \in Y$ (yes, this is written correctly! – it means that, conditioned on $Y = a$, X can only take values among the vertices in a).

(a) Show that, if G is the complete graph on V, then $H_G(P) = H(P)$, the usual entropy of the probability distribution P.

(b) What is the graph entropy of the uniform distribution on the vertices of a pentagon?

(c) Let $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ be two graphs on the same vertex set V, let $E = E_1 \cup E_2$, let $G = (V, E)$, and let P be a probability distribution on V. Show that

$$H_G(P) \leq H_{G_1}(P) + H_{G_2}(P).$$