201700080 Information Theory and Statistics
12 April 2018, 8:45 - 11:45

This test consists of 5 problems for a total of 32 points. All answers need to be justified. The use of a non-programmable calculator (not a "GR") is allowed and advised. No books, notes, or other materials may be used.

Formulas you may find useful:

$$D(P \,\|\, Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}$$
$$H(X, Y \mid Z) = H(X \mid Z) + H(Y \mid X, Z)$$
$$I(X\,;\,Y \mid Z) = H(X \mid Z) - H(X \mid Y, Z)$$

1. Consider a source with alphabet $\mathcal{X} = \{a, b, c, d, e, f\}$ and $p(a) = p(b) = 0.25$, $p(c) = p(d) = 0.15$, $p(e) = p(f) = 0.1$.

   a. [1 pt] Consider the code C for which $C(a) = \mathbf{0}$, $C(b) = \mathbf{10}$, $C(c) = \mathbf{111}$, $C(d) = \mathbf{1101}$, $C(e) = \mathbf{11000}$ and $C(f) = \mathbf{11001}$. Is C instantaneously decodable?

   b. [1 pt] Compute the entropy $H(X)$ of the source.

   c. [3 pt] Construct a Huffman code for the source. Denote this code by C' and let L' denote its codewords lengths. Compute $\mathbb{E}[L']$.

   d. [2 pt] Give two different arguments for why it is not possible to construct a prefix code for this source for which the expected codeword length is 2.4.

2. An error correcting code with generator matrix

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

   is used to communicate over an erasure channel.

   a. [1 pt] What are N and K for this code? What is its rate?

P.T.O. (Please turn over)

b. [3 pt] Output [0 0 1 e 0 1] is received, where e denotes an erasure. Use belief propagation / message passing to try and decode the message that was transmitted. Show the intermediate steps of the decoding algorithm. Does belief propagation work in this case? If not, explain why it fails.

c. [1 pt] Continuing from b., find the message by solving the relevant system of equations.

3. Consider the channel with $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, with $P(Y = 1 \mid X = 1) = 1$ and $P(Y = 1 \mid X = 0) = \frac{1}{2}$. Let $p = P(X = 0)$.

   a. [2 pt] Give an expression for $I(X\,;\,Y)$ as function of p.

   b. [1 pt] Give the definition for channel capacity.

   c. [2 pt] Compute the channel capacity for this channel. Give your answer in two decimals precision. *(Hint:* $\frac{d}{dx}H(x; 1 - x) = \log\left(\frac{1-x}{x}\right)$*.)*

   d. [3 pt] State the noisy-channel coding theorem.

4. Alice has implemented a Rock-Paper-Scissors playing machine, in which the machine is supposed to choose between Rock, Paper and Scissors uniformly at random. Alice suspects that she has made a programming error and that the machine is actually using $P(\text{Rock}) = \frac{2}{3}$, $P(\text{Paper}) = \frac{1}{6}$ and $P(\text{Scissors}) = \frac{1}{6}$.

   a. [3 pt] Alice tests the machine by observing a single outcome. Formulate an hypothesis test for deciding if the machine is functioning correctly or not. Make sure that this test is Neyman-Pearson optimal.

   a. [4 pt] Next, Alice repeatedly observes the machine. She wants to make the probability that she decides that the machine is functioning correctly, while it is in fact erronous, smaller than 10^{-4}. Use the Chernoff-Stein Lemma to give an approximation for the number of observations that she needs to make.

5. Let X be discrete random variable for which $p_\theta(k) = P(X = k) = \frac{\theta^k e^{-k}}{k!}$, for $k = 0, 1, 2, \ldots$. Here, $k!$ denotes the factorial, i.e. $0! = 1$ and $k! = \prod_{i=1}^{k} i$ for $k \geq 1$. Random variables of this form satisfy $\mathbb{E}[X] = \mathrm{Var}(X) = \theta$.

   a. [3 pt] Assume that θ is unknown and compute the Fisher information $J(\theta)$. (*Hint: For discrete probability distributions with parameter θ, the Fisher information is given by* $J(\theta) = \mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\ln p_\theta(X)\right)^2\right]$.)

   b. [2 pt] State the Cramer-Rao bound and discuss in your own words the implications of the result from a) on estimators of θ.