201700080 Information Theory and Statistics
8 April 2021, 9:00 – 12:00

This test consists of 6 problems for a total of 26 points. All answers need to be justified. The use of a non-programmable calculator (not a "GR") is allowed. A handwritten single side A4 cheat sheet is allowed. No additional books or notes may be used.

1. You receive a sample of n observations $x_1, x_2, \ldots, x_n$, that are independent and identically distributed according to a Bernoulli distribution for which the success probability p is unknown. It is known that either $p = p_1$ or $p = p_2$, i.e. the distribution for a single observation is either $P_1(X = x) = (1-p_1)^{1-x} p_1^x$ or $P_2(X = x) = (1-p_2)^{1-x} p_2^x$, $x \in \{0, 1\}$.

   a. [3 pt] Specify a binary hypothesis testing problem for choosing between p_1 and p_2. Derive an optimal decision rule.

   b. [1 pt] Compute $D(P_1 \| P_2)$. *(Hint:* $\mathbb{E}[X] = 1/p$*)*

   c. [2 pt] State the Chernoff-Stein result and discuss the implications for this problem.

2. Let X_1 and X_2 be identically distributed discrete random variables. They are not necessarily independent. Let

$$\rho = 1 - \frac{H(X_2|X_1)}{H(X_1)}.$$

   a. [2 pt] Show that $\rho = \frac{I(X_1;X_2)}{H(X_1)}$.

   b. [1 pt] Show that $0 \le \rho \le 1$.

   c. [1 pt] When is $\rho = 0$?

   d. [2 pt] When is $\rho = 1$?

3. [3 pt] An engineer constructs a Huffman code that achieves expected codeword length $\mathbb{E}[L] = 2.3$. Give an upper and a lower bound on the Shannon entropy of the underlying source. Justify both bounds.

4. [3 pt] A dictionary-based Lempel-Ziv code uses λ to denote the empty string and σ to denote the end-of-message symbol. It gives the following compressed message:

$$(\lambda, a), (1, a), (\lambda, b), (3, c), (3, b), (2, a), (\lambda, \sigma).$$

   Decode and give the original message.

P.T.O. (Please turn over)

5. [3 pt] Consider the binary symmetric channel, i.e. the channel with $\mathcal{X} = \mathcal{Y} = \{0,1\}$ and $P(Y=0|X=0) = P(X=1|Y=1) = 1-p$ and $P(Y=1|X=0) = P(X=0|Y=1) = p$. Derive an expression for the capacity of this channel as a function of p.

6. 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 unbiased estimators of θ based on n observations.