201700080 Information Theory and Statistics
23 April 2024, 13:45 – 16:45

This test consists of 6 problems for a total of 33 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. Consider a Pareto distribution with density function

$$f_\theta(x) = \theta x_0^\theta x^{-\theta-1}, \quad x \geq x_0, \quad \theta > 1.$$

Assume that $x_0 > 0$ is given, but θ is unknown.

a. [3 pt] Compute the Fisher information $J(\theta)$.

b. [2pt] 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.

2. Every evening, the University canteen serves either pasta, curry or fish & chips. The cook claims he picks one of the three recipes uniformly at random each night. However, students have long suspected that he is actually selecting recipes according to the distribution $P(\text{pasta}) = \frac{2}{3}$, $P(\text{curry}) = \frac{1}{6}$ and $P(\text{fish \& chips}) = \frac{1}{6}$.

a. [2pt] Give the log-likelihood function of the presumed distribution and the distribution claimed by the cook.

b. [2pt] Philip decides to test the hypothesis by eating in the canteen for n evenings in a row and recording the served meals. Formulate a binary hypothesis testing problem for choosing between the two distributions.

c. [3pt] Verify Pinsker's inequality for the two distributions.

3. [3pt] Let X, Y and Z be jointly distributed random variables. Prove the following inequality and find conditions for equality:

$$H(X,Y,Z) - H(X,Y) \leq H(X,Z) - H(X).$$

4. Consider a source with alphabet $\mathcal{X} = \{a, b, c, d, e\}$ and $p(a) = 0.3$, $p(b) = 0.2$, $p(c) = 0.15$, $p(d) = 0.2$ and $p(e) = 0.15$. Consider the code C for which $C(a) = \texttt{00}$, $C(b) = \texttt{01}$, $C(c) = \texttt{101}$, $C(d) = \texttt{11}$ and $C(e) = \texttt{1001}$.

a. [1 pt] Use code C to encode the message *aacade* and give the resulting bitstring.

b. [2 pt] Is C a prefix code? Is it a uniquely decodable code? Justify your answers.

c. [1 pt] Compute the expected codeword length $\mathbb{E}[L]$ for C.

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

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

f. [1 pt] Is it possible to construct a prefix code for the source for which the expected codword length is 2.2? Explain.

5. Consider the channel $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, with $P(Y = 1|X = 1) = P(Y = 0|X = 0) = 1 - e$.

   a. [2 pt] Give the definition of channel capacity and state the noisy channel coding theorem.

   b. [3pt] Find the channel capacity for this channel as a function of e. Express your answer using the binary entropy function.

6. A message from a source with $\mathcal{X} = \{\mathtt{a}, \mathtt{b}, ..., \mathtt{z}\}$ is encoded using the sliding-window based Lempel-Ziv code. The message is

   `aabacabracadaaa`

   a. [2 pt] Encode the message and give the resulting sequence of (o, ℓ, c) triples (no need to give the bit-level representation).

   b. [2 pt] If we encode a sequence of length n and consider $n \to \infty$, how many bits per source symbol will the Lempel-Ziv algorithm produce?