EEE 551 Information Theory (Spring 2016)

Homework 6

Due: April 27, 2022

---

1. *Exact sizes of binary types.* Consider sequences of length $n = 20$ on a binary alphabet $\mathcal{X} = \{0, 1\}$.
   (a) List all the types for this alphabet and sequence length.
   (b) For each type, determine the exact size of the type class.
   (c) For each type, evaluate the upper and lower bounds that we derived on the type class (using the entropy of the type). Confirm that the sizes you determined in part (b) are between the two bounds.

2. Problem 11.3 from Cover-Thomas: *Error exponent for universal codes.* A universal source code of rate R achieves a probability of error $P_e^{(n)} \doteq 2^{-nD(P^*\|Q)}$, where Q is the true distribution and P^* achieves $\min D(P\|Q)$ over all P such that $H(P) \geq R$.
   (a) Find P^* in terms of Q and R.
   (b) Now let X be binary. Find the region of source probabilities $Q(x)$, $x \in \{0, 1\}$, for which rate R is sufficient for the universal source code to achieve $P_e^{(n)} \to 0$.

3. Problem 11.5 from Cover-Thomas: *Counting.* Let $\mathcal{X} = \{1, 2, \ldots, m\}$. Show that the number of sequences $x^n \in \mathcal{X}^n$ satisfying $\frac{1}{n}\sum_{i=1}^n g(x_i) \geq \alpha$ is approximately equal to 2^{nH^*}, to first order exponent in the exponent, for n sufficiently large, where

$$H^* = \max_{P: \sum_{i=1}^m P(i)g(i) \geq \alpha} H(P).$$

4. *Six sided-dice probabilities.* Consider an experiment in which a fair 6-sided die is rolled n times. (Fair means that each of the 6 outcomes has equal probability.) Find each of the following probabilities, to first order exponent. That is, find $\lim_{n\to\infty} -\frac{1}{n}\log P_n$ where P_n is the probability for n rolls.
   (a) The probability that less than $n/12$ of the rolls are 1, *and* more than $n/3$ of the rolls are 2.
   (b) The probability that *any* number occurs more than $n/2$ times.

5. *Chernoff bounds.*
   (a) Prove that, for any random variable X with expectation μ, and any $\epsilon > 0$,

$$\Pr\{X > \mu + \epsilon\} \leq \min_{t>0} e^{-t(\mu+\epsilon)}\mathbb{E}[e^{tX}].$$

   (b) For each of the following distributions, explicitly evaluate the right-hand side of the bound from part (a), by optimizing over t:
      i. X is Gaussian with expectation μ and variance σ^2.
      ii. X is exponential with expectation μ.

6. *Hypothesis testing for binary sources.* Let $X_1, X_2, \ldots, X_n$ be i.i.d. binary random variables drawn from $Q(x)$. Consider the hypothesis testing problem $H_0 : Q(x) = \text{Bern}(p_0)$ vs. $H_1 : Q(x) = \text{Bern}(p_1)$.
   (a) Find the error exponent for $\Pr\{\text{Decide } H_1|H_0\}$ in the best hypothesis test for H_0 vs. H_1 subject to $\Pr\{\text{Decide } H_0|H_1\} \leq \frac{1}{2}$.
   (b) Now suppose we are in a Bayesian setting in which each hypothesis has prior probability $\frac{1}{2}$. Find the error exponent for the probability of error of the best hypothesis test in this scenario.

7. *Example of Cramér-Rao lower bound.* Consider the following probability model with unknown parameter θ:

$$f(x;\theta) = \begin{cases} \frac{x^{\frac{1}{\theta}-1}}{\theta}, & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that $T(X) = -\ln X$ is an unbiased estimator for θ.

(b) Find the Fisher information $J(\theta)$.

(c) Find the mean square error for $T(X)$ as defined in part (a), and compare it to the Cramér-Rao lower bound.