E2 201: Information Theory (2020)
Homework 1

Instructor: Himanshu Tyagi

Reading Assignment

- Read chapter 3 of Cover and Thomas book.

Homework Questions Questions marked $*$ are more difficult.

Q1 (Chebyshev inequality)
Use Chebyshev's inequality to give bounds for the number of heads that will be seen when an unbiased coin is tossed 10^8 times.

Q2 (Markov inequality)
Let $X_1, ..., X_n$ be independent and identically distributed random variables with $\texttt{Exponential}(1)$ distribution. Let $Y = \sum_{i=1}^n X_i$. Use Markov's inequality to find a lower bound for the least n required for Y to cross 10^6 with probability 2/3.

Q3 * (Birthday paradox)
Consider n independent draws $X_1, ..., X_n$ from a uniform distribution on [10000]. Using Chebyshev's inequality, find estimates for the number of distinct pairs (X_i, X_j) that are identical. Use this estimate to find an estimate for the least n such that the one such pair exists with probability at least 2/3.

Q4 The total variation distance is defined as

$$d(\mathrm{P}_0, \mathrm{P}_1) = \sup_{A \subset \mathcal{X}} \mathrm{P}_0(A) - \mathrm{P}_1(A).$$

Prove the following equivalent forms of the total variation distance for discrete distributions P_0 and P_1:

$$\begin{aligned} d(\mathrm{P}_0, \mathrm{P}_1) &= \sup_{A \subset \mathcal{X}} \mathrm{P}_1(A) - \mathrm{P}_0(A) \\ &= \sup_{A \subset \mathcal{X}} |\mathrm{P}_0(A) - \mathrm{P}_1(A)| \\ &= \sum_{x \in \mathcal{X}: \mathrm{P}_1(x) \geq \mathrm{P}_0(x)} \mathrm{P}_1(x) - \mathrm{P}_0(x) \\ &= \sum_{x \in \mathcal{X}: \mathrm{P}_0(x) \geq \mathrm{P}_1(x)} \mathrm{P}_0(x) - \mathrm{P}_1(x) \\ &= \frac{1}{2} \sum_{x \in \mathcal{X}} |\mathrm{P}_0(x) - \mathrm{P}_1(x)|. \end{aligned}$$

Q5 Prove the following properties of the total variation distance:

(i) $0 \leq d(\mathrm{P}_0, \mathrm{P}_1) \leq 1$.

(ii) $d(\mathrm{P}_0, \mathrm{P}_1) = 0$ if and only if $\mathrm{P}_0 = \mathrm{P}_1$.

(iii) $d(\mathrm{P}_0, \mathrm{P}_1) = 1$ if and only if P_0 and P_1 have disjoint supports.

Q6 For the set $\mathcal{T}_\lambda^{(1)} = \{x : -\log P(x) \leq \lambda\}$, show that $|\mathcal{T}| \leq 2^\lambda$. Provide example of a distribution and a λ for which is bound is tight.

Q7 Consider the set $\mathcal{T}_\lambda^{(2)} = \{x : -\log P(x) > \lambda\}$. For a distribution P on $\mathcal{X}$, suppose that $P(\mathcal{T}_\lambda^{(2)}) \geq 1 - \varepsilon$. Show that

$$|\mathcal{T}_\lambda^{(2)}| > 2^\lambda (1 - \varepsilon).$$

Q8 * For a distribution P on $\mathcal{X}$, define

$$L_\varepsilon(P) := \min\{\lceil \log |A| \rceil : \exists\, A \subset \mathcal{X} \text{such that} P(A) \geq 1 - \varepsilon\},$$
$$\overline{L}(P) := \min\{\mathbb{E}_P\left[|e(X)|\right] : e : \mathcal{X} \to \{0,1\}^* \text{ is one-to-one}\},$$

where $|b|$ denotes the length of a binary vector $b \in \{0,1\}^*$. Show that

$$\frac{\overline{L}(P)}{1-\varepsilon} - \frac{\varepsilon \lceil \log \mathcal{X} \rceil}{1-\varepsilon} \leq L_\varepsilon(P) \leq \frac{\overline{L}(P)}{\varepsilon} + 1.$$