E2 201: Information Theory (2020)
Homework 2

Instructor: Himanshu Tyagi

Reading Assignment

- Read chapter 2 of Cover and Thomas book. From Te Sun Han's book, try to read Lemma 1.3.1 (similar to the theorem on page 9 of Unit 2) and Lemma 2.1.1 (related to theorems in Section C, D of Unit 3)

Homework Questions Questions marked $*$ are more difficult.

Q1 Calculate the entropy of the following rvs:

(i) X is the output of n independent tosses of a coin which shows heads with probability p;

(ii) $X \sim \texttt{Geometric}(p)$;

(iii) Consider a channel $(\mathcal{X}, W, \mathcal{Y})$ with input $\mathcal{X} = \mathbb{R}$ and output $\mathcal{Y} = \mathbb{N}$ such that for every input $x \in \mathcal{X}$, the output Y is distributed as $\texttt{Poisson}(x)$. Calculate the entropy of rv Y which is the output corresponding to a random input X distributed as $\texttt{exponential}(\lambda)$.

Q2 For distributions in part (i) and (ii) of Q1, write a computer program to find the following quantity (as a function of ε):

$$\lambda_\varepsilon := \min\{\lambda : P(\{x : \lceil -\log P(x) \rceil \leq \lambda\}) \geq 1 - \varepsilon\}.$$

Plot $\lambda_\varepsilon / H(P)$ as a function of ε for different values of the parameter p of the distributions (you can fix $n = 1000$ in part (i)).

Q3 Let $L_\varepsilon(P)$ be given by

$$L_\varepsilon(P) = \min\{\ell \in \mathbb{N} : \log |S| \leq \ell, P(S) \geq 1 - \varepsilon, S \subset \mathcal{X}\}.$$

In the class, we showed that

$$L_\varepsilon(P) \leq H(P)/\varepsilon. \tag{1}$$

Suppose now we are given weights $w_x > 0$, $x \in \mathcal{X}$. Consider the quantity obtained by replacing $|S|$ by $\sum_{x \in S} w_x$ in the definition of $L_\varepsilon(P)$. For this quantity, establish a counterpart of (1).

Q4 Consider the following two procedures for generating random hash $F : \mathcal{X} \to \{0,1\}^\ell$. First, F is chosen uniformly over all functions $f : \mathcal{X} \to \{0,1\}^\ell$. Second, $F(x)$ is generated uniformly over $\{0,1\}^\ell$, independently for all x.
Show that these two processes yield the same distribution of F.

Q5 Consider a distribution P and a set $A \subset \mathcal{X}$ such that

$$\min_{x \in A} - \log P(x) > \lambda.$$

Show that for every $N > 1$ we can find a partition $A_1, ..., A_M, A_{M+1}$ of A such that

$$\begin{aligned} N2^{-\lambda} \leq & P(A_i) \leq (N+1)2^{-\lambda}, \qquad 1 \leq i \leq M, \\ & P(A_{M+1}) \leq N2^{-\lambda}, \\ \text{and } & P(A) \cdot \frac{2^\lambda}{N+1} - \frac{N}{N+1} \leq M \leq \frac{2^\lambda}{N}. \end{aligned}$$

Q6 Consider the following distribution P on $\{0,1\}^{10}$:

- with probability $1/2^\ell$, the first ℓ bits of the sequence are 0, and all sequences starting with ℓ 0s appear with equal probability, for $1 \leq \ell \leq 10$;
- all the remaining sequences appear with equal probability.

What is the minimum number of random bits needed to generate a sample from $\hat{P}$ such that $d(\hat{P}, P) \leq 2^{-5}$? Provide a concrete algorithm. You need not formally prove that this algorithm uses the minimum number of bits; a heuristic argument will suffice.

Q7 Calculate the KL divergence between the following distributions.

(i) $\mathrm{P} \equiv \mathcal{N}(\mu_1, \sigma^2)$ and $\mathrm{Q} \equiv \mathcal{N}(\mu_2, \sigma^2)$;

(ii) $\mathrm{P} \equiv \texttt{Poisson}(\lambda_1)$ and $\mathrm{Q} \equiv \texttt{Poisson}(\lambda_2)$;

(iii) $\mathrm{P} \equiv \texttt{Geometric}(p)$ and $\mathrm{Q} \equiv \texttt{Geometric}(q)$.

Q8 * This question concerns total variation distance and KL divergence between two Bernoulli distributions.

(i) Calculate the total variation distance between $\mathrm{P} \equiv \texttt{Bernoulli}(p)$ and $\mathrm{Q} \equiv \texttt{Bernoulli}(q)$.

(ii)* For P and Q above, show that $D(\mathrm{P}\|\mathrm{Q}) \geq \frac{2}{\ln 2}(p-q)^2$.

Q9 * Consider the binary hypothesis testing problem where one out of two coins with biases 1/2 and $1/2 + \varepsilon$ is chosen randomly and tossed multiple times. Let $n^*(\varepsilon)$ denote the minimum number of tosses needed to find out which coin was chosen with average probability of error less than 1/3. Show that $c/\varepsilon^2 \leq n^*(\varepsilon) \leq C/\varepsilon^2$ for some constants c and C.

Answer the same question when biases are 0 and ε.