UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Problem Set 10
Fall 2020

1. **Capacity of a Binary Erasure Channel**

Recall that the *entropy* of a discrete random variable X is

$$H(X) \triangleq -\sum_x p(x) \log p(x) = -\mathbb{E}[\log p(X)],$$

where $p(\cdot)$ is the PMF of X. The logarithm is taken with base 2 (the unit of entropy is bits).

(a) Prove that $H(X) \geq 0$.

(b) Entropy is often described as the average information content of a random variable. If $H(X) = 0$, then no new information is given by observing X. On the other hand, if $H(X) = m$, then observing the value of X gives you m bits of information on average.

Let X be a Bernoulli random variable with $\mathbb{P}(X = 1) = p$. Would you intuitively expect $H(X)$ to be greater when $p = 1/2$ or when $p = 1/3$ (and why)? Calculate $H(X)$ in both of these cases and verify your answer.

(c) We now consider a **binary erasure channel** (BEC).

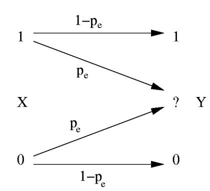


Figure 1: The channel model for the BEC showing a mapping from channel input X to channel output Y. The probability of erasure is p_e.

The input X is a Bernoulli random variable with $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = 1/2$. Each time that we use the channel the input X will either get get erased with probability p_e, or it will get transmitted correctly with probability $1 - p_e$. Using the character "?" to denote erasures, the output Y of the channel can be written as

$$Y = \begin{cases} X, & \text{with probability } 1 - p_e \\ ?, & \text{with probability } p_e. \end{cases}$$

Compute $H(Y)$.

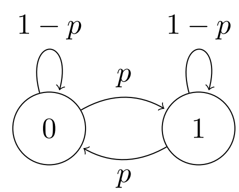
(d) We defined the entropy of a single random variable as a measure of the uncertainty inherent in the distribution of the random variable. We now extend this definition for a pair of random variables (X, Y), but there is nothing really new in this definition because the pair (X, Y) can be considered to be a single vector-valued random variable. Define the *joint entropy* of a pair of discrete random variables (X, Y) to be

$$H(X, Y) \triangleq -\mathbb{E}[\log p(X, Y)],$$

where $p(\cdot, \cdot)$ is the joint PMF and the expectation is also taken over the joint distribution of X and Y.
Compute $H(X, Y)$, for the BEC.

2. **Compression of a Markov Chain**

Consider an irreducible Markov chain $(X_n)_{n \in \mathbb{N}}$ as shown below.



Suppose $X_0 \sim B(\frac{1}{2})$. Roughly how many bits are needed to represent $(X_0, X_1, \ldots, X_n)$?

3. **Huffman Questions**

Consider a set of n objects. Let $X_i = 1$ or 0 accordingly as the i-th object is good or defective. Let $X_1, X_2, \ldots, X_n$ be independent with $\mathbb{P}(X_i = 1) = p_i$; and $p_1 > p_2 > \cdots > p_n > 1/2$. We are asked to determine the set of all defective objects. Any yes-no question you can think of is admissible.

(a) Propose an algorithm based on Huffman coding in order to identify all defective objects.

(b) Suppose the worst case scenario happens and we have to ask the maximimum number of questions. What (in words) is the last question we should ask? And what two sets are we distinguishing with this question?

Note: This problem is related to the 'Entropy and Information Content' section of the Huffman Coding Lab.

4. **Statistical Estimation**

Given $X \in \{0, 1\}$, the random variable Y is exponentially distributed with rate $3X + 1$.

(a) Assume $\mathbb{P}(X = 1) = p \in (0, 1)$ and $\mathbb{P}(X = 0) = 1 - p$. Find the MAP estimate of X given Y.

(b) Find the MLE of X given Y.

(c) Solve the hypothesis testing problem of X given Y with a probability of false alarm at most 0.1. That is, find $\hat{X}$ as a function of Y that maximizes $\mathbb{P}(\hat{X} = 1 \mid X = 1)$ subject to $\mathbb{P}(\hat{X} = 1 \mid X = 0) \leq 0.1$.

(d) For what value of p does one have the same solution for (a) and (c)?

5. **Gaussian Hypothesis Testing**

Consider a hypothesis testing problem that if $X = 0$, you observe a sample of $\mathcal{N}(\mu_0, \sigma^2)$, and if $X = 1$, you observe a sample of $\mathcal{N}(\mu_1, \sigma^2)$, where $\mu_0, \mu_1 \in \mathbb{R}$, $\sigma^2 > 0$. Find the Neyman-Pearson test for false alarm $\alpha \in (0, 1)$, that is, $\mathbb{P}(\hat{X} = 1 \mid X = 0) \leq \alpha$.

6. Flipping Coins and Hypothesizing

You flip a coin until you see heads. Let

$$X = \begin{cases} 1 & \text{if the bias of the coin is } q > p. \\ 0 & \text{if the bias of the coin is } p. \end{cases}$$

Find a decision rule $\hat{X}(Y)$ that maximizes $\mathbb{P}[\hat{X} = 1 \mid X = 1]$ subject to $\mathbb{P}[\hat{X} = 1 \mid X = 0] \leq \beta$ for $\beta \in [0, 1]$. Remember to calculate the randomization constant γ.