UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: Probability and Random Processes

Problem Set 8
Fall 2018

Issued: Wednesday, October 10, 2018 **Due:** Wednesday, October 17, 2018

Problem 1. **Shannon Code**
Consider the following method for generating a code for a random variable X that takes on m values $\{1, 2, \dots, m\}$ with probabilities $p_1, \dots, p_m$. Assume that the probabilities are ordered so that $p_1 \geq p_2 \geq \cdots \geq p_m > 0$. Define

$$F_i = \sum_{k=1}^{i-1} p_k,$$

to be the sum of the probabilities of all symbols less than i, and $F_1 = 0$. Then, in order to construct the codeword for i, which we denoted by $f(i)$, we consider the binary expansion of $F_i \in [0, 1)$, and round it off to l_i bits, where $l_i = \lceil \log_2 \frac{1}{p_i} \rceil$. Here, we do not allow the binary expansions to end with infinitely many ones, e.g. we write $1/2$ in binary as 0.1 not $0.0111\dots$.

1. Construct the code for the probability distribution $(0.5, 0.25, 0.125, 0.125)$.

2. Show that this code is prefix–free, that is, if $i \neq j$ are two different symbols, then their corresponding codewords are not prefix of each other, i.e. $f(i)$ is not a prefix of $f(j)$.

   Hint: show that if $u, v \in [0, 1)$ are such that $|u - v| \geq 2^{-l}$, then the first l bits of the binary representation of u and v can not be the same.

3. If L denotes the average codeword length, show that

$$H(X) \leq L < H(X) + 1.$$

4. Assume that $X_1, X_2, \dots$ are i.i.d. copies of X. Note that for each $n \geq 1$, we can treat the block $X_1, \dots, X_n$ as one random variable taking value in the set $\{1, \dots, m\}^n$ and use the above scheme to encode it. Let L_n denote the average codeword length for this coding scheme and show that

$$\lim_{n \to \infty} \frac{1}{n} L_n = H(X).$$

Problem 2. **Mutual Information**
The **mutual information** of X and Y is defined as

$$I(X;Y) := H(X) - H(X \mid Y).$$

Here, $H(X \mid Y)$ denotes the **conditional entropy** of X given Y, which is defined as: $H(X \mid Y) = -\sum_{y \in \mathcal{Y}} p_Y(y) \sum_{x \in \mathcal{X}} p_{X|Y}(x \mid y) \log_2 p_{X|Y}(x \mid y)$. The interpretation of conditional entropy is the average amount of uncertainty remaining in the random variable X after observing Y. The interpretation of mutual information is therefore the amount of information about X gained by observing Y.

1. Show that $H(X,Y) = H(Y) + H(X \mid Y) = H(X) + H(Y \mid X)$. This is often called the **Chain Rule**. Interpret this rule.

2. Show that $I(X;Y) = H(X) + H(Y) - H(X,Y)$. Note that this shows that $I(X;Y) = I(Y;X)$, i.e., mutual information is symmetric.

3. Consider the **noisy typewriter** in Figure 1. Let X be the input to the

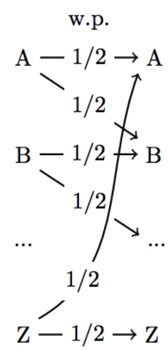

Figure 1: Noisy typewriter: Each symbol gets sent to one of the adjacent symbols with probability 1/2.

noisy typewriter, and let Y be the output (X is a random variable that takes values in the English alphabet). What is the distribution of X that maximizes $I(X;Y)$?

It turns out that $I(X;Y) \geq 0$ with equality if and only if X and Y are independent. The mutual information is an important quantity for channel coding.

Problem 3. **Random Multiplication**
Let X be uniformly distributed in the set $\{0, 1, 2, \ldots, 6\}$ and Z be uniformly distributed in the set $\{1, 2, \ldots, 6\}$. Also, define $Y = XZ (\text{mod } 7)$. Find $H(X|Y)$.

Problem 4. **Isolated Vertices**
Consider a Erdös-Renyi random graph $\mathcal{G}(n, p(n))$, where n is the number of vertices and $p(n)$ is the probability that a specific edge appears in the graph. Let X_n be the

number of isolated vertices in $\mathcal{G}(n, p(n))$. Show that

$$\mathbb{E}[X_n] \xrightarrow{n\to\infty} \begin{cases} \infty, & p(n) \ll \dfrac{\ln n}{n}, \\ \exp(-c), & p(n) = \dfrac{\ln n + c}{n}, \\ 0, & p(n) \gg \dfrac{\ln n}{n}, \end{cases}$$

where the notation $p(n) \ll f(n)$ means that $p(n)/f(n) \to 0$ as $n \to \infty$, and $p(n) \gg f(n)$ means $p(n)/f(n) \to \infty$ as $n \to \infty$. Show also that in the third case, $p(n) \gg (\ln n)/n$, we have $X_n \to 0$ in probability as well.

Problem 5. **Random Bipartite Graph**
Consider a random bipartite graph with, K left nodes and M right nodes. Each of the $K \cdot M$ possible edges of this graph is present with probability p independently.

1. Find the distribution of the degree of a particular right node.
2. We call a right node with degree one a *singleton*. What is the average number of singletons in a random bipartite graph?
3. Find the average number of left nodes that are connected to at least one singleton.

Problem 6. **[Bonus] Connected Random Graph**
We start with the empty graph on n vertices, and iteratively we keep on adding undirected edges $\{u, v\}$ uniformly at random from the edges that are not so far present in the graph, until the graph is connected. Let X be a random variable which is equal to the total number of edges of the graph. Show that $\mathbb{E}[X] = O(n \log n)$.
Hint: consider the random variable X_k which is equal to the number of edges added while there are k connected components, until there are $k-1$ connected components. Don't try to calculate $\mathbb{E}[X_k]$, an upper bound is enough.