ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 9 Introduction to Communication Systems
Homework 5 October 19, 2010, SG1 – 15:15-17:00

Problem 1. Let $X_1, X_2, \cdots$ be i.i.d. random variables drawn according to the probability distribution $p(x)$, $x \in \mathcal{X}$, i.e., $p(x_1, \cdots, x_n) = \prod_{i=1}^n p(x_i)$.

(a) What does $[p(x_1, \cdots, x_n)]^{\frac{1}{n}}$ converge to (in probability)?

(b) Let $f(x)$ be a function from $\mathcal{X}$ to the interval $(0, 1]$. What does $[\prod_{i=1}^n f(x_i)]^{\frac{1}{n}}$ converge to (in probability)? What about $[\prod_{i=1}^n f(x_i)]^{\frac{1}{n+\log(n)}}$?

(c) How does $\mathbb{E}(\prod_{i=1}^n f(x_i))^{\frac{1}{n}}$ compare to $\mathbb{E}f(X_1)$?

Problem 2. A discrete stochastic process $X_1, X_2, \cdots$ is said to be a Markov chain or a Markov process if for $n \in \mathbb{N}$

$$p_{X_{n+1}|X_n,\ldots,X_1}(x_{n+1} \mid x_n, x_{n-1}, \cdots, x_1) = p_{X_{n+1}|X_n}(x_{n+1} \mid x_n)$$

for all $x_1, x_2, \cdots, x_n, x_{n+1} \in \mathcal{X}$. A Markov chain is said to be *time-invariant* if

$$p_{X_{n+1}|X_n}(x_{n+1} \mid x_n) = p(x_{n+1} \mid x_n)$$

for some *transition probability* $p(\cdot \mid \cdot)$. In the sequel we only consider time-invariant Markov chains. Let $p(x_1)$ denote the marginal distribution of X_1.

(a) Show that if for each $x \in \mathcal{X}$,

$$p(x) = \sum_{y \in \mathcal{X}} p(y)p(x|y),$$

then $X_1, X_2, \cdots$ is a *stationary.*

(b) Find a (simpled) expression for the entropy rate of a stationary Markov source.

(c) One of the new students got lost at EPFL the day he arrived and for the whole day he walked around in EPFL. As he didn't know where he was going, he decided to choose one of the possible doors (illustrated in Figure 1) leading out of each building uniformly at random and follows the path out of the current building to the connecting building (regardless of the door he entered in to the current building). To make it simple, let's assume that EPFL's plan is as illustrated in Figure 1, and the points of our interest are only IN building, CO building and SG building. The sequence of the buildings he passed in his walk $(X_1, X_2, \cdots, X_i, \cdots)$ forms a stochastic process (where $X_i \in \{\text{IN}, \text{CO}, \text{SG}\}$) which we call a random walk in this problem.

With what probabilities will the student be in each of the aforementioned buildings at the end of the day, *i.e.,* what is the stationary distribution of the random walk?

(d) What is the entropy rate, call it $\mathcal{H}$, of this random walk?

(d) Compare the entropy rate of this random walk with the entropy of the stationary distribution (i.e., compare $H(X)$ with $\mathcal{H}$) and explain why that relationship holds.

Figure 1: EPFL plan and the new student paths.

Problem 3. Consider a tree with M leaves $n_1, \cdots, n_M$ with probabilities $P(n_1), \cdots, P(n_M)$. Each intermediate node n of the tree is then assigned a probability $P(n)$ which is equal to the sum of the probabilities of the leaves that descend from it. Label each branch of the tree with the label of the node that is on that end of the branch further away from the root. Let $d(n)$ be a distance associated with the branch labeled n. The distance to a leaf is the sum of the branch distances on the path to from root to leaf.

(a) Show that the expected distance to a leaf is given by $\sum_n P(n)d(n)$ where the sum is over all nodes other than the root. Recall that we showed this in the class for $d(n) = 1$.

(b) Let $Q(n) = P(n)/P(n')$ where n' is the parent of n, and define the entropy of an intermediate node n' as

$$H_{n'} = \sum_{n:\ n \text{ is a child of } n'} -Q_n \log Q_n$$

Show that the entropy of the leaves

$$H(\text{leaves}) = \sum_j P(n_j) \log P(n_j)$$

is equal to $\sum_{n \in I} P(n) H_n$ where the sum is over all intermediate nodes including the root. Hint: use part (a) with $d(n) = \log Q(n)$.

(c) Let X be a memoryless source with entropy H. Consider some valid prefix-free dictionary for this source and consider the tree where leaf nodes corresponds to dictionary words. Show that $H_n = H$ for each intermediate node in the tree, and show that

$$H(\text{leaves}) = \mathbb{E}[L] H,$$

where $\mathbb{E}[L]$ is the expected word length of the dictionary. Note that we proved this result in class by a different technique.