MIDTERM

Tuesday, October 28, 2008, 13:15-17:15

This exam has 5 problems and 80 points in total.

Instructions

- You are allowed to use 1 sheet of paper for reference. No mobile phones or calculators are allowed in the exam.
- You can attempt the problems in any order as long as it is clear which problem is being attempted and which solution to the problem you want us to grade.
- If you are stuck in any part of a problem do not dwell on it, try to move on and attempt it later.
- Please solve every problem on **separate paper sheets**.
- It is your responsibility to **number the pages** of your solutions and write on the first sheet the **total number of pages** submitted.

Some Preliminaries

- A sequence of random variables $\{X_n\}$ **converges** toward X **in probability** if

  $$\lim_{n\to\infty} \Pr[|X_n - X| \geq \varepsilon] = 0,$$

  for any $\varepsilon > 0$. For example the Weak Law of Large Numbers implies that if $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables, and $S_n = \frac{1}{n}\sum_{i=1}^{n} X_n$, then

  $$\lim_{n\to\infty} \Pr[|S_n - \mathbb{E}[X]| \geq \varepsilon] = 0.$$

  In other words, S_n converges to $\mathbb{E}[X]$ in probability.
- The following approximations might be useful.

$0\log_2 0 = 0$ $\quad \log_2 3 = 1.58$ $\quad \log_2 5 = 2.32$ $\quad \log_2 6 = 2.58$ $\quad \log_2 7 = 2.81$ $\quad \log_2 23 = 4.52$

Good Luck!

Problem 1 (15 pts)

Let $X_1, X_2, \cdots, X_n, \cdots$ be independent, identically distributed random variables drawn from $\mathcal{X} = \{0, 1, 2, 3, 4, 5\}$ according to the probability distribution $\{8/23, 6/23, 4/23, 2/23, 2/23, 1/23\}$ which is ordered according to $\mathcal{X}$ above.
Define $Y_n = \frac{1}{n} \log p(X_1, X_2, \cdots, X_n)$, and $Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i^2$.

(a) Does Y_n converge in probability? If so, calculate the value Y it converges to. *[5pts]*

(b) Does Z_n converge in probability? If so, calculate the value Z it converges to. *[6pts]*

(c) Compare Z and $(\mathbb{E}[X])^2$ and explain why that relationship holds for an arbitrary choice of $\mathcal{X}$ and $p(x)$. *[4pts]*

Problem 2 (15 pts)

There are 6 bottles of wine, one of which you know has gone bad. You do not know which bottle contains the bad wine, but you know that the probability of each bottle being bad is $(8/23, 6/23, 4/23, 2/23, 2/23, 1/23)$. The bad wine has a distinctive taste. To find the bad wine your friend suggests you to taste a little bit of each wine until you find the bad wine.

(a) To have the least number of tastings on average, what should your strategy be? Which bottle should be tasted first? *[2pts]*

(b) What is the average number of tastings to find the bad wine? *[5pts]*

(c) Calculate the minimum average number of tastings if you are allowed to taste a mixture of different wines and detect a bad wine's taste inside (the distinctive taste is retained even when mixed with other good wines). *[6pts]*

(d) Is the strategy studied in (a) optimal if you are allowed to mix wines? *[2pts]*

Problem 3 (18 pts)

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 $X_i \in \{\mathrm{IN}, \mathrm{CO}, \mathrm{SG}\}$) which we call a random walk in this problem.

(a) 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? *[6pts]*

(b) What is the entropy rate ($\mathcal{H}$) of this random walk? *[8pts]*

(c) 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. *[4pts]*

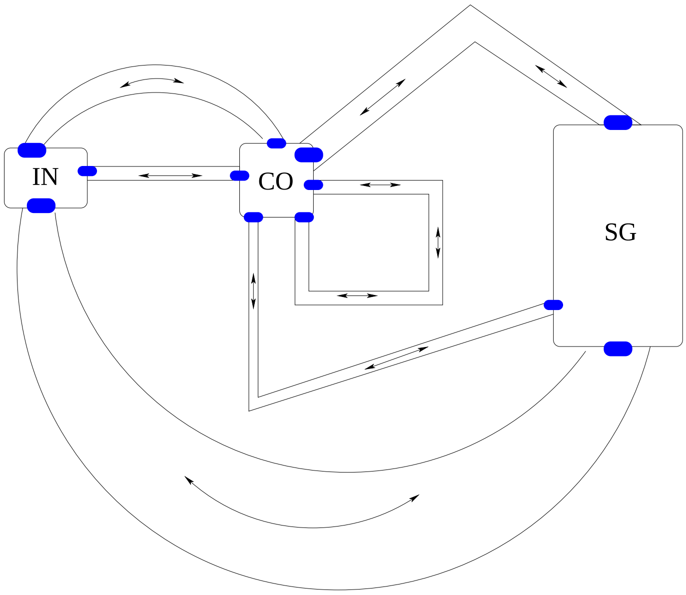

Figure 1: EPFL plan and the new student paths.

Problem 4 (20 pts)

A source is to be encoded but we do not know the distribution of the source symbols exactly. What we know is that with probability λ, the source produces alphabet symbols (A, B, C, D, E, F) according to model 1 with probabilities $P_1 : (1/2, 1/4, 1/16, 1/16, 1/16, 1/16)$, and with probability $1 - \lambda$, the source produces alphabet symbols (A, B, C, D, E, F) according to model 2 with probabilities $P_2 : (1/2, 1/4, 0, 0, 1/8, 1/8)$. What is the optimal encoding strategy and the average length of the code in the following scenarios

(a) Neither the encoder, nor the decoder knows which model produces the source. *[8pts]*
Hint: You might need to consider different regimes of $\lambda \in [0, 1]$. *You may need to consider the cases* $\lambda = 0$ *and* $\lambda = 1$*, separately as well.*

(b) Both encoder and decoder know the model producing the source. *[4pts]*

(c) Let the symbols come from model 2, *i.e.,* $\lambda = 0$. Assume a (liar) genie tells both the encoder and the decoder that the symbols come from model 1, *i.e.,* that $\lambda = 1$, and the encoder designs an optimal code based on this information. Find the average length of the code. What is the penalty of this mis-information, *i.e.,* calculate the difference between the average length of optimal code for the true model and the designed code. How is this difference related to the Kullback-Leibler divergence, $D(\cdot||\cdot)$, between the two distributions. *[8pts]*

Problem 5 (12 pts)

Let the random variable X be a message we want to send to a receiver (receiver 1) and a good approximation $\hat{X}$ is required at that receiver. Consider another receiver, named receiver 2, which has access to a random variable Z, where Z and X are from a joint distribution $p(x, z)$. Receiver 2 is interested in another approximation of X, denoted by $\check{X}$. Imagine the encoding strategy is as follows. The encoder describes X, by $S = f_1(X)$, where S is such that we can find $\hat{X}$ as a function of S ($\hat{X} = f_2(S)$). Imagine that T is constructed from S ($T = f_3(S)$), such that $\breve{X}$ is then found as a function of Z and T ($\breve{X} = f_4(T, Z)$). This system is shown in Fig. 2. Show that

$$I(X; \breve{X}) \leq I(X; \hat{X}) + I(X; SZ|\hat{X}).$$

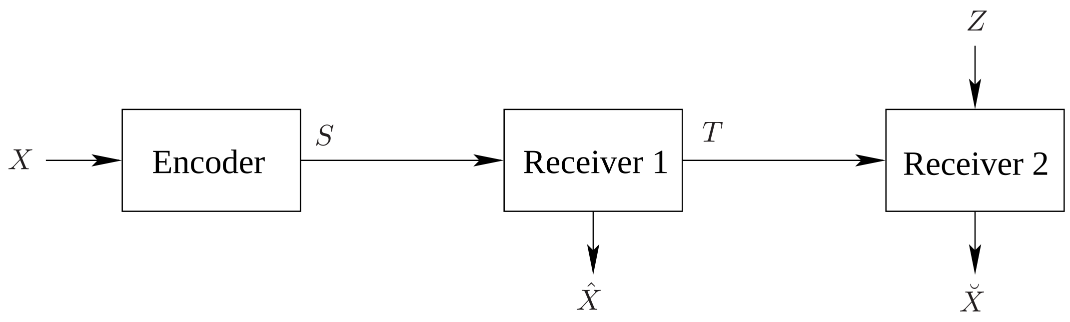

Figure 2: Transmission system in Problem 5.