Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
Department of Mechanical Engineering

6.050J/2.110J **Information and Entropy** **Spring 2006**

Issued: March 14, 2006 **Problem Set 6** **Due: March 17, 2006**

Problem 1: The Registrar's Worst Nightmare

To reinforce Bayes' Theorem, which was covered last week, here's another probability problem.

There are two ways for MIT students to register for 2.110/6.050. Either put it on your Registration Form on Registration Day (call this event R) or use an Add Form after Registration Day (event A). (Assume you cannot do both.) Those who for some reason don't want to take 2.110/6.050 do neither (call this event N). For MIT students, events R, A, and N form a partition. After processing by the Registrar, students may actually be enrolled in 2.110/6.050 (call it event E) or unenrolled (event U). Events U and E also form a partition. (Students can discover whether they are enrolled by checking their status of registration.)

You would probably want things to work so that all Forms are handled properly, i.e., $p(E \mid R) = 1$, $p(E \mid A) = 1$, and $p(U \mid N) = 1$. Unfortunately, one year, 15% of the Add Forms got lost in the mail, 5% of the 2.110/6.050 entries on Registration Forms were illegible because of ink smears, and errors in typing enrolled 1% of the people who didn't want to take the course, by mistake. Knowledge you gain in one part of the problem may be applied to all subsequent parts of the problem.

a. If your friend Alice submitted an Add Form for 2.110/6.050, what is the probability $p(E \mid A)$ that she is enrolled?

b. A study showed that of those that were enrolled in 2.110/6.050, 20% did so by using an Add Form. If your friend Bob is enrolled, what is the probability $p(A \mid E)$ that he submitted an Add Form?

c. Suppose that 2% of the students are enrolled in 2.110/6.050, so $p(E) = 0.02$. Write symbolically and find numerically the probability that a random freshman, say your friend Carol, BOTH submitted an Add Form AND is enrolled.

d. What is the probability that Carol submitted an Add Form for 2.110/6.050?

e. You have just learned that Carol is actually enrolled in 2.110/6.050, a fact which you did not know in parts c. and d. Now what is the probability that she submitted an Add Form?

f. What is the probability that a student, David, selected at random from the whole of the student body, did not put 2.110/6.050 on his Registration Form and did not use an Add Form?

g. What is the probability that David is enrolled in 2.110/6.050?

h. What is the probability that Ed, who did not put 2.110/6.050 on the Registration form and who also did not use an Add Form, is enrolled in 2.110/6.050?

Problem 2: Communicate

a. . A noisy symmetric binary channel is able to transmit 24,000 bits per second, but has a 1% probability of introducing an error into each bit transmitted. What is its channel capacity in bits per second?

One application proposed for this channel requires an error rate of less than 0.1% but the bit rate need not be as large as 24 Kbps. Consider counteracting the noise by using the [7,4] Hamming code. Define a new channel, consisting of the old channel with the attached coder and decoder.

b. What is the bit rate of this newly defined channel?

c. What is the approximate error rate of the channel?

d. Find the channel capacity.

e. Repeat parts (b) through (d) using a triple redundancy code instead of the Hamming code and compare its capacity with capacities that you have found in (a) and (d). What can you infer?

Problem 3: Dealing with bad detectors

Alyssa P. Hacker, our friend at Little-Nibble.com, is designing a novel fiber optical communication system. In her system, bit 1 is transmitted by sending a very weak optical pulse while bit 0 is represented by sending nothing in the fiber. The challenge that she faces is detection of these weak pulses at the receiver.

Alyssa finds a company that produces detectors that seem to be sensitive enough. When she buys their detectors and tests of them, she realizes that although they sometimes detect the weak pulses, they often fail. More precisely, when there is no signal in the fiber (the bit is 0), no detection occurs, but when there is a signal (the bit is 1) it is only detected 50% of the time.

Alyssa wonders about the performance of her system if these detectors are used. She decides to model the system composed of both the optical fiber and the detector as a binary channel, knowing that with these detectors it will not be symmetric. She assumes that the optical fiber in the system is noiseless and lossless.

a. Denote the input of the channel as A and the output as B and give the transition probability matrix with elements c_{00}, c_{01}, c_{10}, c_{11}.

b. Alyssa is planning to use this nonsymmetric channel with a nonsymmetric source coder that produces a sequence of bits of which only one fifth, on average, are 1. That is, $P(A_0) = 0.8$ and $P(A_1) = 0.2$. What is the entropy of the output of this source coder?

c. What are the probabilities of the channel outputs B_0 and B_1 ?

d. First Alyssa determines what the channel capacity would be if the detectors were ideal, i.e., always detected the weak pulses properly. What is it? (Hint: read the definition of channel capacity carefully, and be sure to pay attention to the proper role, if any, of the fact that the source coder is not symmetric.)

e. What is the mutual information between input and output of the channel with the noisy detectors?

f. What is the maximum mutual information over all possible input probability distributions? What do you think about Alyssas source coder?

g. What is the Channel capacity of this channel if the bit rate is 10^9 bits per second?

Turning in Your Solutions

If you used MATLAB for this problem set, you may have some M-files and a diary. You may turn in this problem set by e-mailing your M-files and diary to 6.050-submit@mit.edu. Do this either by attaching them to the e-mail as text files, or by pasting their content directly into the body of the e-mail (if you do the latter, please indicate clearly where each file begins and ends). Alternatively, you may turn in your solutions on paper in room 38-344. The deadline for submission is the same no matter which option you choose.

Your solutions are due 5:00 PM on Friday, March 17, 2006. Later that day, solutions will be posted on the course website.