

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
Department of Electrical Engineering and Computer Science  
Department of Mechanical Engineering

6.050J/2.110J

Information and Entropy

Spring 2003

Issued: March 30, 2003

## Problem Set 7

Due: April 4, 2003

**Note:** The **quiz** will be held Thursday, April 24, 2003, 12:00 noon - 1:00 PM, in Room 37-212. The quiz will be **closed book** except that you may bring one sheet of 8 1/2 x 11 inch paper with notes on both sides. Calculators will not be necessary but you may bring one if you wish. Material through the end of Problem Set 8 may be covered on the quiz.

### Problem 1: Uncertain Employment

You have graduated from MIT with high honors, and have moved on to the field of politics, your passion. After many years of hard work, you eventually become advisor to the President of the United States on employment policy. According to the classification system you have devised there are three possibilities for the employment status of a person:

- Fully employed (100% of capacity)
- Partially employed (50% of capacity)
- Unemployed (0% capacity)

The President is trying to set economic policy so that the employment rate of the country is kept to 95%, as measured with your system (i.e., each partially-employed person counts as half a fully-employed person). There is some confusion, however, about whether this average is to be the goal each month (averaging over all the citizens) or the goal for each citizen (averaging over all months in the President's term in office).

Initially, the President says the average is to be met each month. You do not know the exact statistics yet of the economic data, and so you express what knowledge you do have in terms of probabilities – the probabilities  $p(E)$  (fully-employed),  $p(F)$  (partially-employed, or as an engineer would say, fractionally-employed), and  $p(U)$  (unemployed) of the employment status of any particular citizen. Since each citizen falls into exactly one of the categories, you know that the events  $E$ ,  $F$ , and  $U$  form a partition, so

$$p(E) + p(F) + p(U) = 1 \quad (7-1)$$

Furthermore, you know that when the policy goal is achieved, the average employment will be 0.95:

$$1.0p(E) + 0.5p(F) + 0.0p(U) = 0.95 \quad (7-2)$$

Noticing that there are many possible probability distributions consistent with these two constraints, you decide to find the distribution that uses only the information you have. In other words, you use the principle of maximum entropy. (This approach is consistent with your experience that usually politicians reveal the minimum amount of information about the the success of their policies.)

**Note:** To do this problem, you may want to use MATLAB. If you do and then hand in your assignment on paper, please write the MATLAB statement you used. Remember to show all work. Note that  $\log_2(\mathbf{x})$  is not implemented correctly on all MATLAB versions, so use  $\log(\mathbf{x})/\log(2)$  instead. The commands `solve` and `diff` can be used to find where a function's derivative is zero; to see the help files for these operations, type `help sym/solve` or `help sym/diff`.

- a. Recall that all probabilities must lie in the interval between 0 and 1.

- i. What range of  $p(E)$ , the probability that any one citizen is fully-employed, is consistent with these constraints?
  - ii. Plot the entropy of the probability distribution over this range as a function of  $p(E)$ . (If you turn in the problem set on paper, be sure to include this graph!)
  - iii. Find the maximum entropy, and give the value of all three probabilities at this maximum.
- b. The President feels that most of the citizens would rather be fully-employed than otherwise, so he asks for another probability distribution that is consistent with the constraints but that maximizes  $p(E)$ . Before calculating the entropy in this case, say whether it could be less than, equal to, or greater than the entropy calculated in part (a-iii) above.
- c. Now find the other two probabilities, and the entropy.
- d. Another economic advisor points out that the constraint on the average could be fulfilled by averaging over the entire term, so there is no need to adhere to it the very first month. He suggests a wider range of options in assigning the probabilities. What is the minimum employment (averaged over all citizens) that the country can have in the first month? Describe in words how this situation is achieved. What are the three probabilities in this case? What is the entropy? In this section and the next you may assume that the government has fine enough control over the economy to adjust the number of fully employed, partially employed, or unemployed persons to whatever number it so desires.
- e. What is the maximum employment (averaged over all citizens) the country can have in any given month? Describe in words how this situation is achieved. What are the three probabilities in this case? What is the entropy?
- 

## Problem 2: Candy Raffle

You have entered a raffle at the Cambridgeside Galleria wherein you must guess certain properties of a jar filled with Tootsie Rolls, M&M's, and Hershey's Kisses. Normally you would guess only how many pieces of candy were in the jar, but the Galleria is wise to the ways of MIT students, and so has decided to make the raffle slightly more difficult than normal by asking for properties other than the number of pieces of candy. You know that M&M's weigh 0.9 grams, Tootsie Rolls weigh 2.1 grams, and a Hershey's Kiss weighs 1.5 grams. You also know that Tootsie Rolls are 85% sugar and 15% other ingredients, M&M's are 45% sugar and 55% other ingredients, and Hershey's Kisses are 25% sugar and 75% other ingredients. The candy in the jar has a total mass of 2300 grams, but you don't know how many of each type are in the jar. To win the raffle, you are asked to provide the answers to the following questions:

- a. What ranges of numbers for each type of candy in the jar are compatible with this total mass?
  - b. You are told there are 1100 pieces of candy in the jar. Make no additional assumptions. Determine those numbers of each type of candy that maximize your remaining uncertainty. Don't worry if you get fractional pieces of candy. (The MATLAB solve command will not work on this problem.)
  - c. Given the numbers you calculated, what do you expect the percentage by mass of sugar inside the jar to be?
  - d. Can you have 600 Tootsie Rolls? Is this consistent with the answer in part a? Explain.
-

## Turning in Your Solutions

If you used MATLAB for this problem set, you may have some M-files and a diary. Name the M-files with names like `ps7p1.m`, `ps7p2.m`, and name the diary `ps7diary`. You may turn in this problem set by e-mailing your written solutions, 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). If you have figures or diagrams you may include them as graphics files (GIF, JPG or PDF preferred) attached to your email. 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, April 4, 2003. Later that day, solutions will be posted on the course website.