CSE 569 Homework #2

Total 3 points

Due on the Blackboard on Friday, October 7 **by 11:59pm**.

Note:

1. Submission of homework must be electronic. Most of problems can be solved by hand. The plots of Problem 1 can be approximately done by hand or by a computer program. Then you can either type your work or take pictures of your hand-written sheets and the upload your work onto the Blackboard.
2. If you have any question on the homework problems, you should post your question on the Blackboard discussion board (under Homework 2 Q & A), instead of sending emails to the instructor or TAs. The instructor will answer your question there. This will help to avoid repeated questions, and also help the entire class to stay on the same page whenever any clarification/correction is made.

Problem 1. Let X have an exponential density

$$p(x \mid \theta) = \begin{cases} \theta e^{-\theta x}, & if\ x \geq 0 \\ 0, & otherwise \end{cases}$$

(1) Plot $p(x|\theta)$ versus x for θ =1 (i.e., $p(x|\theta)$ is viewed as a function of x).

(2) Plot $p(x|\theta)$ versus θ , $0 \leq \theta \leq 5$, for x =2 ((i.e., $p(x|\theta)$ is viewed as a function of θ).

(3) Given a training set of n samples D ={x_1, x_2, …, x_n}(i.i.d. samples drawn from the above distribution), find the MLE for θ.

Problem 2. You are given a probability density function (pdf) $p(x)$ for a random variable X as $p(x) = ce^{-|2x-\mu|}$, where $|2x-\mu|$ is the absolute value of $(2x-\mu)$, and μ is a parameter of the density.

(a) What is the value for the constant c ?

(b) Assume that we have the following training data set D ={x_1, x_2, x_3}. Find the Maximum Likelihood Estimate (MLE) of μ given D.

Problem 3. A random variable X can take values 0, 1, 2, …., and it has the following distribution:

$$P(X = n \mid \lambda) = \frac{\lambda^n}{n!} e^{-\lambda}$$

where λ is the parameter of the distribution. We are also given a training set D ={X_1}, i.e., only one sample drawn from the above distribution.

(1) Find the MLE of λ.

(2) In this part, we further assume λ has the following prior distribution,

$$f(\lambda) = \begin{cases} e^{-\lambda}, & if\ \lambda > 0 \\ 0, & otherwise \end{cases}$$

Use the Bayesian Estimation technique to find an estimate for λ by finding the mean of the posterior density $f(\lambda|D)$.

Problem 4. (Optional) This is a problem on HMM. The attached txt files *Observations.txt* and *States.txt* are for this problem.

[**Note**: For this option, you may (and are encouraged to) use Matlab's built-in functions such as, hmmtrain(), etc.]

There are two unbalanced coins, labeled 1 and 2, respectively. They are unbalanced in the sense that when each of them is tossed, the probability of showing a head (i.e., P(H)) is not 0.5. Coin 1 has a P(H) different from that of Coin 2.

The following experiment was repeated 30 times to generate the txt files. (Note: all the probabilities mentioned in this problem are assumed to remain unchanged across different repetitions of the experiment).

A person standing behind a curtain picks one of the two coins with certain probability and tosses it. After a coined is tossed, it will be re-tossed with certain probability (this is to say, another coin may be chosen for tossing with 1 minus this probability). After each toss, the label of the coin used (1 or 2) and the outcome of the toss (head or tail, or H or T) are noted down. Twenty (20) tosses are performed.

The experiments result in the two data files, both contains 30 lines by 20 columns of data points. Each line is from one experiment, and on each line, (i) *States.txt* gives the recorded coin labels for the 20 tosses, and (ii) *Observations.txt* gives corresponding outcomes (H or T).

You are asked to use an HMM to model the experiments and implement algorithms for estimating the model parameters in the following two situations:

(1) You have access to both *Observations.txt* and the corresponding *States.txt* (knowing *States.txt* is equivalent to your seeing which coin was picked in each toss behind the curtain).

(2) You have access only to the *Observations.txt* file (so you don't know which coin was picked behind the curtain for each toss, and you are given only the outcome, an H or a T). In this case, in addition to estimating the HMM parameters, you are also required to estimate the most likely state sequence for each of the observation sequences under the estimated model, i.e., producing an estimate for the file *States.txt*.