STAT 515

Homework #8, due Friday, Mar. 30 at 2:30pm

This homework must be submitted electronically to ANGEL. I strongly encourage the use of LaTeX.

Please make every assignment easier to grade by neatly organizing your writeup and clearly labeling your final answers when appropriate.

1. The matrix exponential function is defined as

$$\exp\{M\} = \sum_{i=0}^{\infty} \frac{M^i}{i!}.$$

However, this definition does not provide a suitable method for calculating $\exp\{M\}$ for a given M. One simple alternative is to use one of the two formulas

$$\exp\{M\} = \lim_{n\to\infty} \left(I + \frac{M}{n}\right)^n = \lim_{n\to\infty} \left[\left(I - \frac{M}{n}\right)^{-1}\right]^n.$$

For the post office problem on HW #7, the rate matrix is given by

$$R = \begin{bmatrix} -8 & 8 & 0 & 0 & 0 \\ 12 & -20 & 8 & 0 & 0 \\ 0 & 12 & -20 & 8 & 0 \\ 0 & 0 & 12 & -20 & 8 \\ 0 & 0 & 0 & 12 & -12 \end{bmatrix},$$

where rates are in hours. Approximate the value of $\exp\{.5R\}$, which is the transition probability matrix for a time step of 30 minutes, using two methods:

   (a) For successively larger powers of 2, i.e., $n = 2, 4, 8, \ldots$, find the value of $(I + .5R/n)^n$. Continue until the change in each entry is smaller than 10^{-5}. Report your final value of n and your final approximation of $\exp\{.5R\}$.

   (b) Repeat the same procedure as in part (a) but use $[(I - .5R/n)^{-1}]^n$ instead.

   (c) Use the `expm` function in R or Matlab to evaluate $\exp\{.5R\}$ and compare with the two approximations you obtained.
   In R, you will have to install and load the package called `Matrix`*. Do this using* `install.packages("Matrix")` *and then* `library(Matrix)`*.*

2. Problem 3 in homework #7 described two video game machines at an amusement park. For video game i, each period when it is being used is exponentially distributed with mean $1/\alpha_i$ hours and each period when it is not being used is exponentially distributed with mean $1/\beta_i$ hours, independent of the other machine. Furthermore, $\alpha_1 = 2$, $\alpha_2 = 3$, $\beta_1 = 5$, and $\beta_2 = 6$.

   (a) If neither machine is in use when the park opens at 8:00am, find the probability that both machines are in use at 9:30am.

   (b) Simulate 10,000 realizations of the Markov chain and give a 95% confidence interval for the probability in part (a) based on your simulation. Does your empirical estimate agree with the theoretical value?

3. Suppose that X_1, X_2, X_3, X_4 are i.i.d. from a uniform$(0, 1)$ distribution. Let $S = X_1 + X_2 + X_3 + X_4$.

   (a) Find $P(S < 1)$ exactly using a four-dimensional integral. (Hint: This is not too difficult.)

   (b) Now consider $P(S < 1.5)$. This is much more difficult to find analytically. Instead, use Monte Carlo simulation to approximate this probability. Give a 99% confidence interval for the true probability, and use a large enough sample so that your interval is no wider than 0.01. Report the sample size you used in addition to the interval.

   (c) The central limit theorem approximation to $P(S < 1.5)$ is $P(Y < 1.5)$, where Y is a normal random variable with the same mean and variance as S. Based on your answer to part (b), how good does the central limit theorem approximation appear in this case?