Stat 135, Fall 2006 A. Adhikari
HOMEWORK 3 (due Friday 9/22)

1. Read Section 8.2 and Example A of Section 8.4. There's some redundancy there, but it doesn't hurt to re-read stuff about the Poisson. Then do Problem 8.10.

2. 8.2. You've already done the necessary reading.

3. 8.4a,b.

4. This problem asks you to establish three facts stated in lecture. Recall that the gamma density with parameters $\alpha > 0$ and $\lambda > 0$ is defined by

$$\frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \qquad t > 0$$

The denominator in the constant of integration is

$$\Gamma(\alpha) \;=\; \int_0^\infty t^{\alpha-1} e^{-t} dt$$

a) Use integration by parts to show that $\Gamma(\alpha+1) \;=\; \alpha\Gamma(\alpha)$ for all $\alpha > 0$.
b) Let Z have the standard normal density. Show that Z^2 has the gamma density with parameters $\alpha = 1/2$ and $\lambda = 1/2$ (a.k.a. the chi-squared density with 1 degree of freedom). How can you use your calculation to show that $\Gamma(1/2) \;=\; \sqrt{\pi}$?

In Problems 6 to 10, please use R wherever possible. These problems are intended to give you a warmup with the package.

6. Enter the digits of your SID number as a dataset in R. So for example if your SID number is 19987121, then your dataset is 1, 9, 9, 8, 7, 1, 2, and 1. Call this your population. Draw the histogram of the population. Find its mean and SD (careful: the denominator should be 8, not 7).

7. Find the expectation and standard error of the mean of 400 draws with replacment from your population. (Do you have to draw 400 times first?) Plot the normal curve with this expectation and SE.

8. Now take 100 samples, each of size 400 with replacement, from your population. Compute all the sample means and draw the histogram of these 100 means. On the histogram, superpose the normal curve from Problem 7 (i.e., draw the curve again, over the histogram).

9. From each of your 100 samples, construct a 95%-confidence interval for the mean of your population. Say how many of your intervals contain the population mean, and provide the endpoints of the intervals which **don't** contain the population mean.

10. Take a break from your little population for a bit, and go back to generalities. Suppose you have a population which contains a parameter of interest, and suppose 100 samples are drawn from the population, independently of each other. Suppose that from each sample we can construct a 95% confidence for the parameter. Let G be the number of "good" confidence intervals, that is, those that contain the parameter. What is the probability distribution of G? Draw its histogram. Find $E(G)$ and $SE(G)$.

On the horizontal axis of the histogram, mark the number of "good" intervals you got in Problem 9. Based on the values of $E(G)$ and $SE(G)$, is your mark in a surprising place? Explain briefly.