# STAT 200B: Assignment 2

due Tuesday February 10th 2009 in class

## Homework

From the book: Problems 3.1.5, 3.1.8

## Lab

1. **Distribution Review** Match the most appropriate distribution to the random variable, X, described in each scenario. Give values for the parameters and also state any additional assumptions that are required. Hint: Wikipedia is your friend.

   Choose from: normal, poisson, binomial, multinomial, exponential, gamma or uniform.

   (a) You buy a packet of M&M's for an afternoon snack. Each packet has 25 pieces that can be one of 6 different colours (red, green, yellow, blue, brown and orange). X = (number of red, number of green, number of yellow, number of blue, number of brown)

   (b) When you arrive at your professor's office hours there are 5 students ahead of you. The professor takes on average 6 mins to talk to one student. X = time in minutes until you see the professor.

   (c) You are having a morning coffee in a cafe. At this time of day on average 40 people order coffees every hour. X = the number of people you see order a coffee in the 30 mins you are at the cafe.

   (d) You are going home and arrive at the bus stop just as a bus pulls away. You know there are 5 buses every hour. X = time in minutes until next bus arrives.

   (e) 8% of US males have red-green colour blindness. You randomly sample 100 men. X = number of men in your sample who have red-green colour blindness.

2. **Central Limit Theorem**

   Let $X_1, \ldots, X_n$ be an iid sample from a continuous uniform distribution on the interval [0,1].

(a) Do the $X_i$ satisfy the conditions for the Central Limit Theorem?

(b) According to the Central Limit Theorem what is the asymptotic distribution of $\overline{X}$?

(c) Generate 1000 samples of $X_1, \ldots, X_5$ (a sample of size 5). Plot a histogram of the sample means and overlay the density of the asymptotic distribution.

Hint: It's easiest to make a matrix of your samples and use the `rowMeans` functions to calculate the sample means.

```
# create 1000 samples (in the rows) of 5 iid unif(0, 1) r.v.'s
samples <- matrix(runif(5 * 1000), ncol = 5)
hist(rowMeans(samples), prob = TRUE, breaks = 40)
```

(d) Repeat for samples of sizes of 10, 25, 50 and 100. Plot all on the same plot region with the same x-axis limits. Comment on your findings.

Hint: Use `par(mfrow = c(1, 5))` before your plotting commands to get five plots on one page. The argument `xlim` to `hist` will constrain the x-axis of the histogram. For example, `xlim = c(0, 1)` will constrain the x-axis to $(0, 1)$.

(e) Repeat steps 1-4 for $X_i$ an iid sample from a poisson with mean 1. Comment on the differences between the two distributions.