Assignment #4 contains many probability problems. Problems not in the textbook ask you to graph probability distributions to learn how the distributions change with changes in the value of the parameters.

It is possible to do all of the probability calculations with your calculator, but I encourage you to use R as well. The graphing problems would be extraordinarily difficult without the use of R. You do not need to turn in graphs for this assignment. I encourage you to install R onto your own computer, but you may complete the assignment by using R already on one of the public computer labs on campus, such as in Union South or the CALS computer lab. If you use R to do a problem, it is not necessary to include the R code you used to solve it with your solution.

Here are the assigned problems from the textbook.

From chapter 3, problems 3.11, 3.27, 3.30, and 3.34.
From chapter 4, problems 4.1, 4.5, and 4.8.

In addition, please complete these problems that require the use of R. You will need to download the file `prob.R` from the course R Help web page. After starting R, source in this code using `Source R code...` from the `File` menu. (Alternatively, there is a `source` command in R you can type.)

Problem 8: Use R to plot the five binomial distributions with $n = 9$ fixed and for $p = 0.1$, 0.3, 0.5, 0.7, and 0.9.

(a) How does the center of the distribution change as p increases? (Does it stay the same, increase, or decrease? If it changes, how?)

(b) For which values of p is the variance largest and for which is it smallest?

(c) For which values of p is the distribution skewed to the right and for which is it skewed to the left?

(d) For which values of p is the skewness the strongest?

Problem 9: Use R to plot the six binomial distributions with $n = 10, 20, 40, 80, 160$ and for $p = 0.1$ fixed.

(a) How does the center of the distribution change as n increases? (Does it stay the same, increase, or decrease? If it changes, how?)

(b) How does the variance change as n increases?

(c) How does the skewness change as n increases?

See the additional problems on the next page.

Problem 10: Answer these questions based on the plots from the previous problem.

(a) Create a table with the following headers.

n	(A) range of outcomes with visible probabilities	(B) range of all possible outcomes	ratio (A/B)

For example, on my computer, when I plot the binomial(10,0.1) distribution with the command `gbinom(10,0.1,scale=T)`, the probabilities are visible from 0 to 5, so the range of outcomes with visible probabilities is $5 - 0 = 5$ and the range of possible outcomes is $10 - 0 = 10$. The ratio of these values is $5/10 = 0.5$. Repeat this for $n = 20, 40, 80, 160$ and for $p = 0.1$ fixed.

(b) You can think of the range of outcomes with visible probabilities as an informal estimate of the size of the interval where most of the probability is located. How does this change as n increases? Find an approximate formula for this as a function of n.

(c) You can think of the ratio as an informal estimate of the size of the proportion of possible outcomes with nonnegligible probabilities. How does this change as n increases? Find an approximate formula for this as a function of n.

Problem 11: A model for a single site in a DNA sequence is that the number of nucleotide substitutions at the site in an evolutionary lineage follows a Poisson process with a rate of 0.05 substitutions per million years. Let X be the random number of substitutions in a ten million year time period.

(a) What is $E(X)$, the expected value of X?

(b) What $\Pr\{X = 0\}$, the probability of no substitutions at this site?

(c) What is $\Pr\{X \geq 2\}$, the probability of two or more substitutions at the site?

There are 11 problems, so this assignment is worth 55 HW points. If you feel challenged by these problems, I encourage you to do additional problems on your own. I suggest choosing extra problems with answers in the back of the textbook.