

Assignment #3 contains problems about probability. Problems which require the use of R have the symbol **(R)**.

Please include **your name** and **the discussion section (day/time) that you attend** on your homework.

This assignment is worth 50 points in total, and points for each problem are indicated.

If you feel challenged by these problems, I encourage you to do additional problems on your own. Many problems have answers in the back of the textbook.

Your assignment must be turned in during lecture or to your TA's mailbox by 5pm on the due date. We will not grade late homework. If there are special circumstances, please speak to Professor Larget, preferably in advance, for consideration.

1. **(R)** [10 points] For this problem, I just ask you to turn in the graphs.

This problem teaches you to use R to estimate a probability through simulation. The example we will use is based on Exercise 3.3 on page 83 of the text book. The idea is to find the probability that five randomly selected digits would all be unique. Determining this theoretically is not too difficult. The probability is $1 \times 0.9 \times 0.8 \times 0.7 \times 0.6 = 0.3024$. In the simulation, we will look to see how rapidly the sample relative frequency approaches the actual probability.

The following code demonstrates how to take a single sample of five random digits.

```
> sample(0:9,5,replace=T)
```

The first argument to sample says that you are sampling from the digits 0 through 9, the second argument says the sample size is 5, and the final argument specifies the sampling to be done with replacement. Try this code yourself to see what happens. Are all of the digits the same or not?

The course web page has several R functions in a file named `hw3.R` that you can use to make plots similar to those on page 82 of your textbook. The first step is to load these functions into R. If you are using a computer connected to the internet, the following command should work.

```
source("http://www.stat.wisc.edu/courses/st371-larget/hw3.R")
```

Otherwise, you can download the file onto your computer by right-clicking on the link to `hw3.R` from the schedule on the course homepage. You source this data into R from an option from the File menu (Source File... on a Mac, Source R code... on a Windows machine).

After loading the code, you carry out the simulations with the following commands. Lines with a '#' are comments.

```
# show a single sample
> sampleDigits(1,5)
# keep 100 random samples
> x = sampleDigits(100,5)
# determine which samples are all unique numbers
> res100 = getResult(x)
# compute the proportion
```

```
> sum(res100)/100
# compare to theoretical value
> p = 0.6*0.7*0.8*0.9
> p
# plot the results
> plotResults(res100,p)
# repeat for 10,000 samples
> x = sampleDigits(10000,5)
> res10000 = getResults(x)
> sum(res10000)/10000
> plotResults(res10000,p)
```

2. [5 points] Do Exercise 3.7
3. [5 points] Do Exercise 3.9
4. [5 points] Do Exercise 3.11
5. [10 points] For the data in Exercise 3.21, find (a) $\Pr(Y \geq 2)$; (b) $\Pr(Y \leq 2)$; (c) $E(Y)$; and (d) $\text{Var}(Y)$.
6. [5 points] Do Exercise 3.27
7. [5 points] Do Exercise 3.30
8. [5 points] Do Exercise 3.34