Homework 3 – Intro. to Computational Statistics

For all problems, please show all your work. As described in the Homework Guidelines, use RMarkdown to write up your work as a .Rmd file, "knit" the result to a PDF file, and submit that PDF file to Blackboard. (You can also knit to an HTML or Word document and save that as a PDF, as decribed in the Homework Guidelines.) Be sure to use R code for all your calculations, and the latex equation format to write up any math. See the Homework Guidelines in Course Resources on Blackboard for more formatting details.

For any of the following calculations, feel free to mix calculations by hand with R calculations. If you do something "by hand," be sure to write out step by step in neatly formatted math (using latex) how you calculated your result. Usually the fastest route is to write it out by hand on paper, and then when you've got a solution you're happy with, type it up in RMarkdown; with a little practice, the write-up should go quite quickly.

1.

a. What's the chance of getting a sequential pair on two rolls of a die (eg, a 3 then a 4 counts, but a 4 then a 3 does not). (Hint: you can calculate this manually if you like, by counting up the sample space and finding the fraction of that sample space that consists of ordered pairs.)

b. Given a dartboard with a inner circle that is 2/3 of the total area, and a bulls-eye that is 5% of the total area (and entirely within the inner circle): if you are throwing a random dart (that is guaranteed to hit somewhere on the board, but everywhere inside is equally likely), what is the chance of hitting the bulls-eye conditional on knowing your dart is somewhere inside the innner circle?

c. You take a test for a scary disease, and get a positive result. The disease is quite rare – 1 in 1000 in the general population. The test has a accuracy (sensitivity) of 95%, and a false positive rate of only 5%. What is the chance you have the disease?

d. What is the chance you have the disease if everything remains the same, but the disease is even rarer, 1 in 10,000?

e. What does this tell you about the dangers of tests for rare diseases?

2.

a. You have a 20-side die. Using `sample`, roll it 1000 times and count the number of rolls that are 10 or less.

b. Generate a histogram using ggplot of 10,000 draws from a uniform distribution between 2 and 7.

c. Try to write down the equation for this probability density function.

d. What is the probability that a draw from this distribution will be between 1.5 and 3.2?

3.

a. Using R's cdf for the binomial, what is the probability of getting 500 or fewer "20"s when rolling your 20-sided die 10,000 times. Looking back at 2a, what proportion of your rolls were actually 20s?

b. Using rbinom, roll a 100-sided die 100 times and report the total number of 7s you get.

c. You are a klutz, and the average number of times you drop your pencil in a day is 1. Using the poisson functions in R, what's the chance of dropping your pencil two or more times in a day? (Hint: calculate the chance of dropping it one or fewer times, and then take 1 minus that.)

d. Because he is lazy, your teacher has assigned grades for an exam at random, and to help hide his deception he has given the fake grades a normal distribution with a mean of 70 and a standard deviation of 10. What is the chance your exam got a score of 85 or above? What is the chance you got a score between 50 and 60?