553.283 Introduction to R
Homework 5

Note 1: If a question asks you for a numerical answer, your submission for that question must consist of the R command that produces that answer followed immediately by the output.

Note 2: Please label all axes on any plots you create.

Note 3: When writing functions, please indent each nested block of code further away from the margin than the block in which it is nested. It makes your code much easier for me to read.

Note 4: Properly comment your code where ambiguity may arise. Comments in R are preceded by the # sign.

1. A Bernoulli random variable X with success probability p has the probability mass function

$$P(X = k) = \begin{cases} 1 - p, & k = 0 \\ p, & k = 1 \\ 0, & \text{otherwise} \end{cases}.$$

   Without using any of the R functions for the binomial distribution, write *d*-, *p*-, and *r*- functions for the Bernoulli distribution. Verify that they work as intended.

2. Some days Joshua finds that he has nothing better to do than sit on a porch and watch planes from his apartment window. On average, he sees about 8 planes pass in a single day. What is the probability he observes (strictly) more than 12 planes in a day? Joshua believes that the number of planes follows a Poisson distribution.

   Hint: Recall which random variable mentioned during class is concerned with the distribution of occurrences of an event over a period of time.

3. Simulate a 30-day month of Joshua's plane-watching pastime, create a histogram, and compute the mean and standard deviation.

4. Simulate an entire year of Joshua's plane-watching pastime (assume every month is 30 days long), compute the mean for each month, and generate a histogram of the means. Does the distribution of the means look normal? Why or why not?

5. Simulate two decades' worth of Joshua's watching planes (no leap-years), compute the mean for each year, and generate a histogram of these means. Does the distribution look normal? Why or why not?

6. Write a function *simulatePlanes()* that takes in as input a number of days and a logical value determining whether you want tothe plot or not. If you do not want to see the plot, have it return a vector of simulated plane counts for each day. If you would like to see the plot, have it return a list, the first element of which is the vector of simulated plane counts for each day, and the second element of which is a histogram.

7. **Bonus Question: (10 Points)** In the function above, if you would like to see a plot, have it return a ggplot2 histogram object with the main title being "Simulated Plane Counts for n Days," only the variable n is replaced with the input. Please include an overlaid density estimate as well.