Lab. 2: Chapter 3, R and distributional theory

R and distributions.

Objectives:

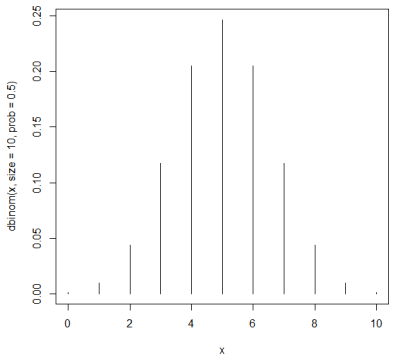
- Learn discrete distributions
- Calculate probabilities and create samples
- Make plots of various kinds using the base and ggplot2 packages

Output:

- Make an Rmd document for this lab
- At the conclusion render into the 3 file types
- Upload the 4 files to the server.

Tasks:

1. The Binomial distribution. Throw a coin "n" times and the number of successes is "x". Where p = probability of a success
    a. Write the formula for $p(X = x|n, p)$ using Latex
    b. Using the answer to the above write your own r function dmybin() to calculate $p(X = x|n, p)$ record this in Rmd
    c. Now calculate $p(X = 4|n = 10, p = 0.5)$ using your function
    d. Use the built in R function dbinom() to calculate the same probability.
    e. What if we wish to calculate the cumulative probability $p(X \leq x|n, p)$, we would need to sum individual probabilities. Make a function called pmybin() that would do the job.
    f. Use the function to calculate $p(X \leq 5|n = 10, p = 0.5)$
    g. Use the built in R function pbinom() to do the same and see whether the answers are the same.
    h. Make the following plot, where n=10,p=0.5 except put your name on the title:



2. Learn how to use the four basic distributional functions

   dstem, pstem, rstem, qstem

   a. Suppose that $X \sim Pois(\lambda)$. Use R and the above function types to answer the following.
      i. Find $P(X = 4|\lambda = 3)$
      ii. Find $P(X \leq 4|\lambda = 3)$
      iii. Find $P(X > 4|\lambda = 3)$
      iv. Find x so that $P(X \leq x|\lambda = 3)$= 0.9997077
      v. Create a sample of size 100 from a Poisson distribution that has parameter $\lambda = 3$. Store in an object.
      vi. Make a second sample of size 100 from a Poisson that has parameter $\lambda = 6$, store in an object
   b. Make boxplots of the random samples you made above.
      i. We will make a data frame of the data. Call the first group "Fst" and the second group "Snd". All data in the first group have to have "Fst" associated with them etc.
      ii. See Laboratory2.R for some exemplar code using ggplot
   c. Make violin plots of the same using ggplot