IEOR W4150 Introduction to Probability and Statistics Fall 2018
Dr. A. B. Dieker

Homework 09-R
due on Thursday November 8, 2018, 9:00am EST

1. Create a group with three members. You sign up your group through this link. If you do not have three members yet, please sign up on the right-hand side of this sheet and meet other incomplete teams after class on Wed 11/31.

   (a) Install R, Rstudio, and Shiny.

   (b) Choose 10 ticker symbols and load the daily stock price information for at least one year into R (daily stock price information is readily available, do a web search to find it). Make sure your data includes both the open and close prices.

   Although you may for instance select mutual funds as well, I will be referring to stocks in all the assignments. (You can change stock symbols later in these R assignments and the project, but you'll have to repeat these steps.)

   (c) Create a histogram of the close prices for one of your stock symbols. Explain how you have chosen the bin size.

   (d) Using R, calculate the log-returns for your stocks for each stock symbol and each day in your data set. The log-return is defined as

$$\log\left(\frac{S_t^{\text{close}}}{S_t^{\text{open}}}\right)$$

   where S_t^{close} and S_t^{open} are the close and open prices, respectively, on day t. (Here log is natural log.) Produce histograms for the log-returns of each of your 10 stocks. Experiment with the bin sizes.

What should we turn in for this homework?

You must turn in exactly one page for this homework for your whole group. It needs to be submitted on Gradescope. Only one person should submit per team, *but that person has to enter the names of the other team members in Gradescope.* For (c), write a one-paragraph description of your findings and supplement this description with a well-chosen selection of histograms. For (d), submit only a well-chosen selection of histograms. For each of these, make sure that you explain what the data represents in each histogram (the TA has to understand the histograms without having access to your data set).