# MATH 185 – Homework 3
## Due Wednesday, 04/25, by 11:59 PM

*Send your code here. For Homework 1, write* "MATH 185 - HW 1" *in subject line and nothing else in the body. There should only be one file attached, with the name* hw1-lastname-firstname.R*. Make sure your code is clean, commented and running. Keep your code simple, using packages only if really necessary. If your code does not run, include an explanation of what is going on.*

**Problem 1.** This is about computing bootstrap confidence intervals.

A. Write a function sd.bootCI(x, level = 0.95, B = 2000) that returns a bootstrap pivotal confidence interval for the standard deviation based on a sample, a specified level, and a given number of bootstrap replicates.

B. Write a function sd.bootCI.student(x, level = 0.95, B = 2000) that returns a bootstrap Studentized pivotal confidence interval for the standard deviation. You have (at least) two avenues for estimating for the standard deviation of the sample standard deviation. Choose the first one if you can, otherwise choose the other (which is more computationally demanding):

   (a) Derive a closed form expression for the standard deviation of the sample standard deviation and then use a plug in to obtain an estimate. (No need to show how you derived it. In fact, you can search for that in books or online.)

   (b) Estimate the standard deviation of the sample standard deviation by bootstrap.

C. Perform some simulations to compare the two methods (pivotal and Studentized pivotal). For example, you can generate a sample of size $n \in \{10, 20, 50, 100, 200\}$ from the exponential distribution with rate 1, compute the intervals, see if they include the true standard deviation and record their length, and repeat that $N = 200$ times, and present a summary of these experiments.

**Problem 2.** Consider the `geyser` dataset in the `MASS` package. Focus on the variable `duration`.

A. Use your function from Problem 1 to build a 90% confidence interval for the standard deviation of the duration of Old Faithful Geyser's eruptions.

B. Do the same, but using the functions `boot` and `boot.ci` in the package `boot`.

(Note that this may not be an iid sample of eruptions, since there might be dependencies over time, but we ignore this in this exercise.)

**Problem 3.** Consider the `geyser` dataset in the `MASS` package. Focus on the variable `waiting`. Waiting times are often modeled as exponentially distributed. Apply the Lilliefors test for that purpose. The exponential family is a scale family, and so the Lilliefors approach is appropriate. You can use the package `KScorrect` or implement it yourself as lilliefors.exponential.test(x, M = 2000).