MATH 185 – Homework 5
Due Wednesday, 05/16, 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 problem is about bootstrap tests:

A. Write a function bootBetweenSumSquares(y, g, B = 1e4) which implements the bootstrap test for comparing the means of multiple groups of observations based on the between sum of squares as described in the notes. Try your function on the smokers dataset.

B. Write a function bootWelchF(y, g, B = 1e4) which implements the bootstrap test for comparing the means of multiple groups of observations based on the Welch F-statistic. (The formula for this statistic is not on the notes, but it can be computed using the function oneway.test.). Try your function on the smokers dataset.

C. Perform some simulations to compare the power of these two tests. First, consider $k = 2$ groups, each of size $n \in \{10, 100\}$. Assume the underlying distributions are normal, $\mathcal{N}(0, 1)$ for the first group, and $\mathcal{N}(\mu, \sigma^2)$ for the second group, where $\sigma \in \{1, 3\}$. As usual, in each of these 4 choices of (n, σ), choose a grid of μ's (about 10 of them) and plot the power of each test (calibrated at level 5%) as a different curve. Repeat each setting (given by a triplet (n, σ, μ)) $N = 200$ times. The result should therefore be four plots, with two curves in each plot (one for each of the two tests). Briefly comment on your findings.

D. Repeat with $k = 4$ groups. Design the numerical experiment on your own.

Problem 2. Propose a test that generalizes the (two-sample) Kolmogorov-Smirnov test to more than 2 groups. Describe the main idea in a couple of sentences as a comment. Then implement your test as a function called oneway.ks.test(y, g, B = 1e4). Try your function on the smokers dataset. [It is acceptable to search for existing tests that fit the description. If you find such a test, provide its name. You still have to implement it yourself.]