Problem Set 1

MGMT 737

Spring 2024

You should have gotten this homework assignment from the Github classroom environment (https://classroom.github.com/classrooms/192971645-yale-mgmt-737-spring-2025-classroom). In submitting your problem set, you should have two files in your Github repository:

1. `homework1.R`, which contains your code,
2. `homework1.md`, which contains your writeup.

and `homework1.R` or `homework1.py`. The first file should be a PDF file that contains your answers to the problem set. The second file should contain the code you used to generate your answers. You can use any other files you want, but they should not be in the Github repository.

This problem set is expected to be coded in either `R` or `Python` (`R` is recommended).

1. Randomization. This analysis will use the Dehijia and Wahba sample from the Lalonde dataset of the NSW experiment. The dataset is `lalonde_nsw.csv`. The outcome variable is `re78` (real earnings in 1978). The treatment indicator is `treat`. The remaining variables are potential covariates. Assume for the purposes of this problem set that `treat` is completely randomly assigned.

   (a) Calculate the average treatment effect of the policy $E(\tau_i)$ using a simple difference in means.

   (b) Calculate the average treatment effect on the treated of the policy $E(\tau_i|\text{treat} = 1)$. How does it compare to part a?

   (c) Test the null of $\tau_i = 0$ for all i using a randomization test.
   N.B. Hold fixed the number of treated and control (e.g. assume the treatment count would be held fixed) and permute the labels randomly 1000 times – you do not need to fully do every permutation (there would be too many). Report the quantile that your estimate from the previous question falls.

   (d) Run a regression using robust standard errors (you may use canned software) of the outcome on the treatment dummy, and compare the p-values from this test to the previous answer.