Problem Set 2

Your Name

This problem set is due at 8:30 am on 9/19

Please upload both Rmd and PDF files on Sakai

PDF file should contain both codes and outputs

For this question we will use a dataset from a randomised experiment conducted by Marianne Bertrand and Sendhil Mullainathan, who sent 4,870 fictious resumes out to employers in response to job adverts in Boston and Chicago in 2001. The resumes differ in various attributes including the names of the applicants, and different resumes were randomly allocated to job openings. Some of the names are distinctly white sounding and some distinctly black sounding. The researchers collecting these data were interested to learn whether black sounding names obtain fewer callbacks for interviews than white names.

Download the data set `bm.dta` from Sakai.

1. The data set contains two dummy variables (0-1 variables) for female (`female`) and whether the applicant has computer skills (`computerskills`). Tabulate these variables by `black`. Do gender and computer skills look balanced across race groups? (1 point)
2. Do a similar tabulation for education and the number of jobs previous held (`ofjobs`). These variables take on 5 and 7 different values, respectively. Does education and the number of previous jobs look balanced across race groups? (2 points)
3. Use `summarize()`, `group_by()`, and other commands to look at the mean and standard deviation for the variable for years of experience (`yearsexp`) separately for black and whites. Does this variable look similar by race? (1 point)
4. What do you make of the overall results on resume characteristics? Why do we care about whether these variables look similar across the race groups? (1 point)
5. The variable of interest on the data set is the variable `call`, which indicates a call back for an interview. (2 points)
    - Manually perform a two-sample t-test comparing applicants with black sounding names and white sounding names. (using the commands we went over in class)
    - Double check your results using the built-in `t.test` command in R. Do you find differences in call back rates by race?
6. Again, manually perform a two-sample t-test comparing of the number of previous jobs (`ofjobs`) between applicants with black sounding names and white sounding names. Double check your results using the built-in `t.test` command in R. Do you find differences in call back rates by race? (2 points)
7. What do you conclude from the results of the Bertand and Mullainathan experiment? (1 point)