STAT 320: HW 3 Due: Monday 2/24/14 at 4:40pm

Instructions

- Bring a hard copy of your homework to class on the due date. Half credit for homework turned in after 4:40pm on the due date, no credit for homework turned in after the due date (because solutions will be posted).
- Also bring a hard copy of your R Script to class on the due date. The script should be self-contained, so someone else should be able to run it as is and get your results. The script should also be well-commented, so it is clear which code goes with which question.
- You may discuss these problems with each other verbally, but must write up the answers on your own, and may not share or show your answers to anyone else (this applies to code as well).
- Please be concise!

I've posted sample R code on estimating propensity scores with the Shadish data here. Throughout there are often a couple options - no need to do both options, you can choose whichever you prefer.

Data

In 1992, New Jersey decided to raise it's minimum wage from $4.25 an hour to $5.05 an hour. What was the *causal effect* of this decision on employment in the fast food industry? In general, economists are interested in whether raising the minimum wage lowers employment in low-wage jobs.

To study this, economists from Princeton collected data from fast food restaurants along the New Jersey - Pennsylvania border, with the Pennsylvania restaurants acting as a control group for the New Jersey restaurants. They also collected data on several covariates for the restaurants. The outcome is employment after the minimum wage was raised in New Jersey. For more information, see the NY Times article Supersize My Wage.

The data is available here. The variable `NJ.PA` is an indicator for which state the restaurant is in (1 if NJ, 0 if PA), `EmploymentPre` measures employment for each restaurant before the minimum wage raise in NJ, and `EmploymentPost` measures employment for each restaurant after the minimum wage raise in NJ. `WagePre` measures the hourly wage for each restaurant before the minimum wage raise, and `BurgerKing`, `KFC`, `Roys`, and `Wendys` are indicators for different restaurant chains.

Exercises

1. (2 points) Do you think unconfoundedness is plausible here? Why or why not?

2. (3 points) Plot the initial covariate balance between the treatment groups. (You can do this anyway you like, but one possibility is with the `cov.balance` function in my sample code, which plots the t-statistic for each covariate). Is there noticeable imbalance for any covariates? If so, describe this imbalance.

3. (10 points) Estimate the propensity scores. Suppose economists tell you that all covariates collected are substantively important, and two-way interactions and quadratics (of the quantitative covariates) may or may not be relevant. Include a written description of your process in choosing the propensity score model, the output from your final model, and a plot of the propensity score distributions within each treatment group (`plot.ps` may be useful). (Also don't forget to turn in your R Script as well.)

   Note: In class, we talked about eliminating units without any comparable units in the opposite treatment group. One way to do this is to look at the estimated propensity scores, and eliminate control units with propensity scores below any treatment units and treatment units with propensity scores above any control units. Once these units are eliminated, the propensity score model should be refit This is an iterative process that should be repeated until the propensity scores are overlapping everywhere.

4. (5 points) [**Rubin 2007 paper - unrelated to the data**]
   Read Rubin (2007). The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials, *Statistics in Medicine*, **26**(1): 20-36 and comment on the paper. Comments may take any form (could be your reaction, a summary, thoughts, questions, points that resonated most with you, points of disagreement, a comparison to what you've learned in previous classes, or anything else you want to comment on) as long as they demonstrate that you read and thought about the paper. Comments may be a few sentences to half a page.