Causal Inference - Problem Set 7

Author: Gustavo Baroni

Professor Marcos Ross

TA: Richard Martins da Cunha

Due to: 17/11

Question 1

Consider the model $Yi = D_i\beta + \epsilon_i$. The interest lies on β, which gauges the causal effect of the treatment on Y_i. D_i is defined as the treatment dummy, which is equal to 1 when the individual i is treated and zero, otherwise. Suppose that

$$D_i = \begin{cases} 1\ if\ x_i \geq x \\ 0\ if\ x_i < x \end{cases}$$

(a) Rewrite the model using potential outcomes framework and list all assumptions that are necessary for consistency, giving the intuition of each one of them.

Rewriting the model using potential outcomes framework:

$$Y_i = D_i Y_i + (1 - D_i) Y_i$$

Besides that, it is required for consistency the Stable Unit-Treatment-Value Assumption (SUTVA), in which Donald Rubin argue that the treatment for one individual shall not impact another's potential outcome. There is also the assumption in which the potential outcome shall be continuous and that there is a discontinuity on cutoff of the running variable.

(b) Show that on a neighborhood around the discontinuity, the estimator of RDD is equal to the causal effect of treatment on Y, i.e, equal β.

Given that

$$E[Y_i|x_i = x](1) \rightarrow E[\alpha x_i + \beta + \epsilon_i | x_i = x]$$
$$E[Y_i|x_i = x](0) \rightarrow E[\alpha x_i + \epsilon_i | x_i = x]$$

We can affirm that the causal effect is equal to β, as supported by the following math operation:

$$E[Y_i|x_i = x](1) - E[Y_i|x_i = x](0) = \beta$$

(c) Does this propriety is valid for all values of x? What is the implication of this?

The propriety is not valid for all values. Once one determines the x, the result is just valid to that specific cutoff.

Question 2

Consider the RDD approach where Y_{is} is the score in a national math exam of individual i in the school s. The treatment is one robotics class per week. This is, all treated schools have a robotics class and the non-treated schools don't have it. However, the assignment for the treat was not random: All schools with less than 100 students are treated. Define D_i as the treatment dummy, i.e. is equal 1 for the schools that has a robotics class and zero, otherwise.

(a) If we are interested in the causal effect of the robotics class on the student's math skills, what's the problem in comparing the mean of the treatment group and the non-treatment group?

Contrasting the means of the treated and controlled groups would lead to a biased result. Since the assignment mechanism was the quantity of students in each school, instead of be random, the means are not comparable.

(b) We can give a better estimator for the causal effect of the robotic class on the student's math skills? What's assumptions are necessary? Is probable that all of these assumptions hold? Why?

It is possible to apply a Regression Discontinuity Design (RDD) method to mitigate the bias in the previous exercise. If we choose a threshold of 100, we could then compare the results, once this is the quantity of students that decides whether a school would get robotics.

Question 3

Still thinking about RDDs, explain the concept of "bandwidth." How does using a larger bandwidth impact your estimate of treatment effects?

"Bandwidth" is, basically, the width of the bin in an analysis of some data. As David Lee and Thomas Lemieux argued[1], "this is typically done informally by trying to pick a bandwidth that makes the graphs look informative in the sense that bins are wide enough to reduce the amount of noise, but narrow enough to compare observations "close enough" on both sides of the cutoff point." To sum up, the higher the value of the bandwidth, the higher the variance, and the lowest the bias.

[1] David S Lee and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature* 48, no. 2 (June 1, 2010): 308–9, https://doi.org/10.1257/jel.48.2.281.