POLS0012 Causal Analysis: Tutorial Exercise 2

Question 1

In this question we will demonstrate that randomised experiments work by creating an imaginary experiment, using the dataset "a" contained in the file "experiment.Rda." It includes the potential outcome under control ($y0$) and the potential outcome under treatment ($y1$) for 100 units that form the sample for our experiment. Note that this is a purely hypothetical scenario. In reality, we never observe potential outcomes under both treatment and control for the same units: we only observe one of them (the fundamental problem of causal inference). By creating a randomised experiment with this dataset, we'll demonstrate that experiments overcome the fundamental problem.

a) Find the true Average Treatment Effect for all units, using $y0$ and $y1$

b) Now, we'll randomly assign half of the units to treatment and half to control by creating a new variable indicating treatment status. As described in the lecture, we can do this using the following steps:[1]

   i) Assign each unit a random number from 1 to 100: create a new column in the dataset named *rand*, using the `sample()` command and a vector of the numbers 1 to 100

   ii) Re-order the dataset from lowest to highest value of *rand* using the code `a <- a[order(a$rand),]`

   iii) Create a treatment variable named *tr* that equals 1 for the first 50 units and 0 for the second 50 using the code `c(rep(1,50),rep(0,50))`

c) Conduct a test to assess whether the treatment and control groups have the same average potential outcomes under control ($y0$). Has randomisation succeeded in creating treatment and control groups with equivalent potential outcomes under control?

d) Find the Average Treatment Effect from the experiment. How similar is it to the true Average Treatment Effect?

e) Now, let's see how our experimental procedure performs over repeated randomisations, using a simulation:

   i) First, create a function that takes in the dataset a, carries out a randomised experiment, and reports the ATE estimated in (d). You can do this by combining code from (b) and (d), without using `set.seed`
   Code Hint: Remember that a function in R has the following format:
   `function.name <- function(input) { function commands }`

[1] First, input the command `set.seed(1)` so that your results are the same as the solutions

ii) Second, find the results of 10,000 randomised experiments by running the function 10,000 times and storing the results in a variable.
Code Hint: You can do this using the `replicate()` command, with two arguments: the first is the number of simulations and the second is the function being evaluated

What is the mean ATE from your 10,000 experiments? Does this suggest that the experimental procedure is unbiased?

f) Finally, repeat (e), calculating the mean difference in potential outcomes under control (y0) between the treatment and control groups instead of the ATE. What is the mean difference from your 10,000 experiments?

Question 2

Why do people bother to vote? One hypothesis is adherence to social norms. Voting is widely regarded as a civic duty and people worry that others will think badly of them if they fail to participate. According to this theory, voters may receive two different types of utility from voting; (a) the intrinsic rewards from performing this duty and (b) the extrinsic rewards received when others observe them doing so. To gauge the effects of priming intrinsic motives and applying varying degrees of extrinsic pressure on voting behaviour, Gerber, Green, and Larimer conducted a famous field experiment in Michigan prior to the August 2006 primary election.[2] The sample for the experiment was 344,084 voters. They were randomly assigned to either the control group or one of four treatment groups.

We'll practice analysing experiments by focusing on two of their treatments. The first treatment, "civic duty", incolved sending a letter to the voter carrying the message "DO YOUR CIVIC DUTY - VOTE!". The second treatment, "Neighbors" sent the same letter, but also informed the voter that who votes is public information (which is the case by law in the USA). It listed the recent voting record of each registered voter in the household and the voting records of those living nearby, and stated that a follow-up mailing after the election would report back to the household and to their neighbours on who had voted and who had not. The idea was to see whether priming extrinsic motivations would encourage this treatment group to turn out more than the control group. The control group received no letter. For this question we'll use the original data of Gerber et al, contained in the file gerber.Rda. Below is a list of the variable definitions:

- *sex* - gender (1 if female, 0 if male)
- *yob* - year of birth
- *p2004* = 1 if Respondent voted in the 2004 Primary Election, 0 otherwise

[2]For the original paper, see: `//isps.yale.edu/sites/default/files/publication/2012/12/ISPS08-001.pdf`

- *voting* = 1 if Respondent voted in the 2006 Primary Election, 0 otherwise [the outcome variable]
- *control* = 1 if Respondent is assigned to the control group, 0 otherwise
- *civicduty* = 1 if Respondent is assigned to the "Civic Duty" group, 0 otherwise
- *neighbors* = 1 if Respondent is assigned to the "Neighbors" group, 0 otherwise)

a) For both treatments, calculate the average treatment effect and test whether it is statistically significant. Interpret the results, giving a precise explanation of the magnitude of the treatment effects. What do they suggest about the motivations that people have for voting?

b) For both treatment groups, compare the mean values of *yob*, *sex* and *p2004* to the control group. Do the results suggest that randomisation was successful? Is selection bias likely to be a problem in this experiment?

c) Calculate the ATE for the the *neighbors* treatment using:

   i) A regression containing only *neighbors*

   ii) A regression containing *neighbors* and the three background characteristics

   Note that you will need to subset your data appropriately in order to obtain the correct control group. Are there any big differences in the estimated ATE between the two specifications? Or between these two estimates and the result from part (a)? Why or why not?