Samuel Snelson

SOC 723 Spring 2023

24 February 2023

**Homework for Chapter 10: Treatment Effects**

1. Define *in your own words* (i.e., don’t just copy down what’s written in the glossary) each of the following terms:
   1. Conditional average treatment effect
      1. First, the average treatment effect is the average difference in the outcome between treatment and control levels (or some main independent variable and a reference point of that variable [e.g., average treatment effect of each $1,000 of income on well-being)]. So, a conditional average treatment effect (CATE) is an ATE on a subset of observations (usually corresponding to some variable of interest).

* 1. Average treatment on the treated
     1. An average treatment effect on the treated (ATT) is a particular type of conditional ATE where we are interested in the average effect of the treatment only among those who were treated. Note that this doesn’t mean that only the data on the treated are used. A common approach to computing this estimate is to simulate counterfactual treatment data representing the outcome of the untreated group if they had received treatment.
  2. Average treatment on the untreated
     1. Conversely, an average treatment effect on the untreated (personal preference with average treatment effect on control, ATC) is an estimate of the treatment effect if the untreated were to receive the treatment. Again, we can compute this estimate by simulating counterfactual control data representing the outcome of the treatment group if they hadn’t received treatment.

1. Provide an example of a treatment effect that you would expect to be highly heterogeneous and explain why you think it is likely to be heterogeneous.
   1. Consider the relationship between placement into foster care and the probability of graduating high school. On average, being placed into foster care is associated with a decreased probability of graduating high school. Yet, there is still considerable variation in the outcome across ethnoracial groups, and depending on the type of placement, the extent of support the youth receives from the family, school, and other services, as well as the stability of the youth’s placement (e.g., moving many times detracts from educational achievement).
      1. In this situation, I would be interested in several conditional ATEs (given the aforementioned variables) and potentially some weighted ATEs.

1. Consider the data in the table below that shows the hypothetical treatment effect of cognitive behavioral therapy on depression for six participants. For the sake of this example, the six participants represent the population of interest.

|  |  |  |  |
| --- | --- | --- | --- |
| Case | Age | Gender | Effect |
| A | 15 | Man | 7 |
| B | 40 | Woman | 3 |
| C | 30 | Woman | 7 |
| D | 20 | Non-binary | 8 |
| E | 15 | Man | 7 |
| F | 25 | Woman | 4 |

* 1. What is the overall average treatment effect for the population?
     + 1. Note that this estimate is unconditional – that is, it is averaged over all sampled observations. As NHK has provided the data, however, we have the individual level effects for all cases. So we can simply take the average of all effects.
       2. The average treatment effect of CBT on depression is 6 (whatever units which aren’t implied).
  2. What is the average treatment effect for Women?
     + 1. This computation is the same as the previous ATE with the condition (haha) that we are looking only at women. Similar to above, we have all case-level effects, so we can average across womens’ effects.
       2. The average treatment effect of CBT on depression among women is about 4.6 (units).
  3. If nearly all Non-binary people get treated, and about half of all Women get treated, and we control for the differences between Women and Non-binary people, what kind of treatment effect average will we get, and what can we say about the numerical estimate we’ll get?
     1. If we accounted for the proportion of each group who is in treatment versus control levels and statistically adjust for “differences between Women and Non-Binary people”, then this would give us a variance-weighted average treatment effect of CBT on depression. Accounting for group proportions in the treatment weighs each groups’ treatment effect by the group’s variance () and the statistical adjustment weighs more highly backdoor paths (e.g., the exogenous source of variation in treatment) other than the path closed by the adjustment.
        1. As for the numerical estimate, as described above, it will regard as more influential high-variance groups and backdoor paths other than the variable adjusted for.
  4. If we assume that, in the absence of treatment, everyone would have had the same outcome, and also only teenagers (19 or younger) ever receive treatment, and we compare treated people to control people, what kind of treatment effect average will we get, and what can we say about the numerical estimate we’ll get?
     1. If we believe that if no one got CBT depression scores would be the same for all individuals, we are then interested in the effect of the treatment on the treated (ATT; because depression scores would only change if there was a treatment). I am not totally sure what NHK means by the second part of this but there are two possibilities.
        1. If only teenagers are included in the experiment, then the treatment effect estimate would be conditional on being a teenager in the sample in addition to the ATT.
        2. If teenagers and adults were included in the experiment and both were assigned to treatment but only teenagers ultimately received treatment, then the estimate could either reflect an intent-to-treat effect (if the total number of cases of interest are included in the denominator) or a local average treatment effect on the treated (LATT; if the denominator accounts for the weighting in treatment received [where adults would get a weight of 0 in the denominator]). On second thought, it sounds like the intent-to-treat and LATT are the same thing, but I’m not confident.

1. Give an example where the average treatment effect on the treated would be more useful to consider than the overall average treatment effect and explain why.
   1. One example I am thinking of is the example I used above. The relationship between placement in foster care and the probability of graduating from high school would present a useful case of the ATT because, although one could be interested in differences between those in and not in foster care, I would be particularly interested in the difference in the probability of graduating between those in foster care and those not placed *if* they were. What is interesting about this is that the differences between those actually placed and those placed by simulation are the contextual factors surrounding placement and experiences within placement.
2. Which of the following describes the average treatment effect of assigning treatment, whether or not treatment is actually received?
   1. Local average treatment effect
   2. Average treatment on the treated
   3. Intent-to-treat
   4. Variance-weighted average treatment effect
3. On weighted treatment effects:
   1. Describe what a variance-weighted treatment effect is.
      1. Variance-weighted treatment effects refer to ATEs where the treatment effect of some defined group is weighted by the variance in those groups’ treatment distributions, assuming the treatment has been isolated as a source of exogenous variation. The more variance a treatment group has, the more highly weighted their treatment value is in the computation of the average effect.
         1. I’m not quite sure how to represent this in the same expected value notation as above, but NHK calculates it with the following:
   2. Describe what a distribution-weighted treatment effect is.
      1. Distribution-weighted treatment effects refer to ATEs where the desired counterfactual is obtained weighting more highly cases with more similar values on selected covariates. This is reflected in various matching methods where members of the treatment group are compared to members of the control groups with similar values on some set of covariates of interest. In principle, this approach seeks to construct a counterfactual based on covariate similarity.
   3. Under what conditions/research designs would we get each of these?
      1. Variance-weighted ATEs arise when one wished to account for variability in the distribution of treatment across groups of interest. If 10% of men received treatment and 75% of women received treatment in some natural experimental or instrumental variable context, we may be interested in weighting effect estimates by variance.
      2. Distribution-weighted ATEs arise when the research either samples or matches treatment and control groups on covariate similarity. In both cases, these reflect a way of adjusting (more or less effectively) for the covariates being made strategically similar or the same. In the latter case, matching often is used when researchers don’t have access to experimental data but wish to estimate treatment effects (in a similar manner as an experimental treatment would allow).
4. Suppose you are conducting an experiment to see whether pricing cookies at $1.99 versus $2 affects the decision to purchase the cookies. The population of interest is all adults in the United States. You recruit people from your university to participate and randomize them to either see cookies priced as $1.99 or $2, then write down whether they purchased cookies. What kind of average treatment effect can you identify from this experiment?
   1. If I am seeking to generalize to “all adults in the United States” and I exclusively sample “people from [my] university,” then this implies treatment effects are conditional on being from that university.
   2. If we are looking at the difference in the (conditional) average treatment effect based on data for those who got the $1.99 cookie (treatment) and $2 cookie (control), then we could get the ATE from this?
   3. If we were interested in the average treatment effect of getting the $1.99 cookie option relative to the control group if they had instead gotten the $1.99 option, then this would identify the conditional ATT. Conversely, if we were interested in the average treatment effect of getting the $1.99 option if the treatment group had instead gotten the $2 option, this would reflect the conditional ATC. This would be relevant if we were interested in what the effect would be if we were to extent the treatment to the control group (i.e., predicting what the outcome would be for the treatment group if they had gotten the $2 option). This would reflect any effect the control group would get from treatment.
5. For each of the following identification strategies, what kind of treatment effect(s) is most likely to be identified?
   1. A randomized experiment using a representative sample
      1. ATE
   2. True randomization within only a certain demographic group
      1. CATE (if the effect is isolated on that group). If the whole sample if used but with partial randomization, then the backdoor corresponding to the non-random process of selection into the treatment would be closed (ideally) giving a variance-weighted ATE.
   3. Closing back door paths connected to variation in treatment
      1. (variance-) weighted-ATE
   4. Isolating the part of the variation in treatment variable that is driven by an exogenous variable
      1. LATE
   5. The control group is comparable to the treatment group, but treatment effects may be different across these groups
      1. ATT/ATC. If we are interested in the average treatment effects between the treatment and treatment-counterfactual