Ch 10 and 11

Emilio Horner

2023-02-22

Ch 10

a. Conditional average treatment effect

Its the average treatment effct controlling for certain variables within the group selected. The example the book gives is if you had a group of all women, or people from the same region.

b. Average treatment on the treated

The average treatment effect among people who recieved the treatment variable.

c. Average treatment on the untreated

The average treatment among the people who did not recieve the treatment variable.

2.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

A variable that would be highly heterogeneous could be impact a drug that prevents balding. It would vary a lot based on gender and other genetic factors.

3.

a. the average treatment effect is 6
b. for women it is 4.66
c. for women if only half of them take the treatment it would be ((3 * .5) + 7 * .5) + (4 * .5))/3 = 2.33333 for nonbinary people (ugh) it is 8

so the difference between the average treatment effects for women and non-binary people is 2.33-8= -5.66 d. the average treatment effect for teenagers would be 7, the control aspect is that we only would select teenagers to look at.

4. 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

If you're measuring a self selected group that would choose to get a treatment. For example if its a medical drug that only works on men, the people in the world who have choosen to take the treatment are already self selected to be men. So measuring for thep population as whole would not be as accurate or useful.

5.

Intent to Treat (c)

6.

On weighted treatment effects: a.Describe what a variance-weighted treatment effect is b. Describe what a distribution-weighted treatment effect is c. Under what conditions/research designs would we get each of these?

a. Variance-weighted treatment effect: It is weighted by how representative the observations are if one is choosing a subsample of the data.

b. A distribution-weighted treatment effect is when one has selected a sample where the treated and untreated groups have similar values for variables on back door paths.
c. the variance weighted treatment effect is basically just when some of the treated people actually take the treatment and some don't. so it would be if there was a pill given but some of the men don't take it while all the women do. it weights the treatment for that variance. the distribution weighted variable is similar except it is where the back door is closed during the sample group. So it would be like choosing a sample group of all men or all women to close the backdoor.

7. 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?

Its conditional average treatment effect because there's randomization within a certain group

8.

a. A randomized experiment using a representative sample

b. True randomization within only a certain demographic group

c. Closing back door paths connected to variation in treatment

d. Isolating the part of the variation in treatment variable that is driven by an exogenous variable

e. The control group is comparable to the treatment group, but treatment effects may be different across these groups

f. average treatment effect

g. conditional average treatment effect c, weighted average treatment effect

h. local average treatment effect

i. average treatment on the treated

Ch 11

1. Suppose that you are analyzing the effect of universities and colleges opening during a pandemic on increase in the number of positive cases. Name one strategy that you can use to avoid having to collect data on all types of campus characteristic variables that are constant over time that you may have to control for in your analysis.

a fixed effects model would work for this problem

2. Intuitively, why would a method that isolates front doors allow you to ignore back doors related to unmeasured variables?

You could ignore back doors because they are not relevant to the model. They don't impact the path.

3. On robustness tests:

a. What are robustness tests?

b. What is the purpose of conducting a robustness test?

c. What are placebo tests?

d. Robustness tests test how assumptions in the pathways impact the data.

e. Robustness tests determine either ways to disprove an assumption or robustness tests redo the the analysis in a way that doesn't rely on the assumption.

f. In a placebo the control group just recieves a fake version of the treatment. This reveals whether the treatment actual changed or potentially changed the outcome or whether the act of taking something makes people believe they have been impacted.

4. Suppose you want to study the effect of attending tutoring sessions on grade point averages (GPA). List at least five variables that impact both attendance of tutoring sessions and students' GPA. Is it feasible to measure and control for all of the variables?

intelligence, quality of tutor, rigor of school, whether they have a romantic partner (distraction!), socio-economic status

5. Partial identification is where instead of making strong relational statements we say "up to 10%" improvement in GPA with tutoring because we don't know for sure how much of it is the tutoring vs other variables.

6.

a.

```
library(broom)
library(tibble)
library(ggdag)
```

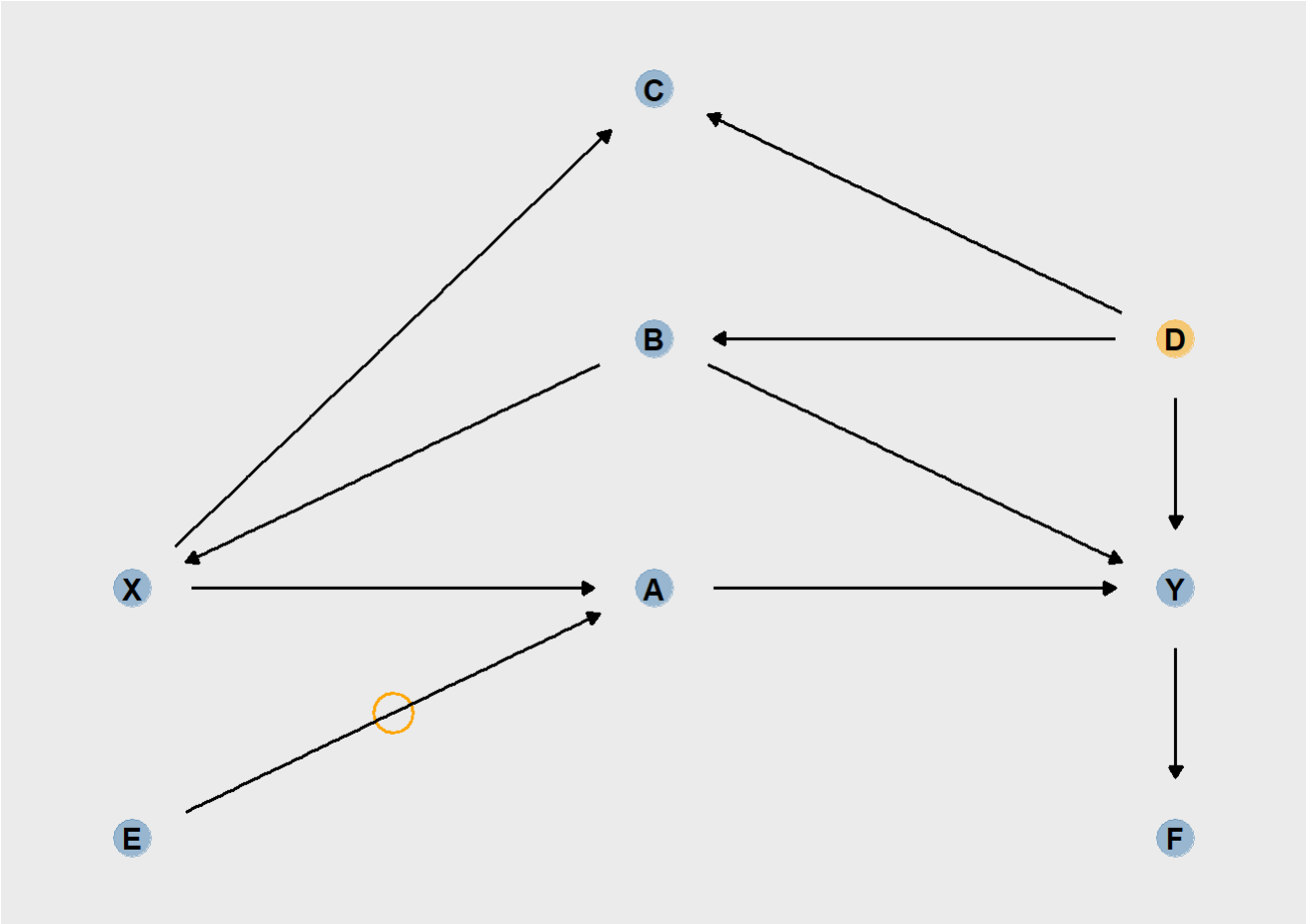
```
##
## Attaching package: 'ggdag'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(ggplot2)
library(dagitty)
```

```
library(ggdag)
dag_coords <-
 tibble(name = c("E", "X", "A", "B", "C", "D", "Y", "F"),
 x = c(1, 1, 3, 3, 3, 5, 5, 5),
 y = c(1, 3, 3, 5, 7, 5, 3, 1))
dagify(A ~ E,
 A ~ X,
 C ~ X,
 C ~ D,
 B ~ D,
 X ~ B,
 Y ~ B,
 Y ~ A,
 Y ~ D,
 F ~ Y,
 coords = dag_coords) %>%

 ggplot(aes(x = x, y = y, xend = xend, yend = yend)) +
 geom_dag_point(aes(color = name == "D"),
 alpha = 1/2, size = 6.5, show.legend = F) +
 geom_point(x = 2, y = 2,
 size = 6.5, shape = 1, stroke = 1, color = "orange") +
 geom_dag_text(color = "black") +
 geom_dag_edges() +
 scale_color_manual(values = c("steelblue", "orange")) +
 scale_x_continuous(NULL, breaks = NULL, expand = c(.1, .1)) +
 scale_y_continuous(NULL, breaks = NULL, expand = c(.1, .1))
```



b. X <— B —-> Y
c. Control for B
d. That there is a relationship or that its a placebo effect.

7. What does it mean to say that the effect of financial deregulation on the rate at which firms go bankrupt is “bounded from above” at 2 percentage points?

it means that the effect is up to 2% positive. its bounded so thats the outer most limit. (e)