Lecture 12

Alex Stephenson

10-4-2021

What did we cover last time?

For the last two weeks we have covered the mechanics of regression.

The big takeaways from that discussion are:

1) We can use regression to analyze experiments

2) If we do not have a good design, a regression will give us wonky results.

What are we doing today?

Today we are going to talk about two sided non-compliance

New groups: Always-Takers and Defiers

New assumption: Monotonicity

Notation Review

We are now writing potential outcomes as a function of $d_i(z)$

What does d_i represent?

- Indicates whether a subject was actually treated.

What does z represent?

- Indicates whether a subject was assigned to treatment

Two-Sided Non-complinance

I'll abbreviate this as 2SNC

Classification of Subject Types

We now have four possible groups:

1) Compliers: $d_i(z) = z$

2) Never-Takers: $d_i(1) = 0, d_i(0) = 0$

3) Always-takers: $d_i(1) = 1, d_i(0) = 0$

4) Defiers: $d_i(1) = 0, d_i(0) = 1$

Monotonicity Assumption

Informally, we have no defiers.

More formally:

For all $i \in \{1, \ldots, n\}$, $d_i(1) \geq d_i(0)$.

This rules out defiers because we can characterize their potential outcomes as $d_i(1) < d_i(0)$

ITT, ITT_D and CATE under 2SNC

Fortunately, if we're willing to assume monotonicity nothing changes about the CACE Theorem

We just have to adapt the definition of Compliers to deal with 2SNC.

Compliers: $d_i(1) - d_i(0) = 1$. Using this definition

$$CACE = E[Y_i(d=1) - Y_i(d=0) | d_i(1) - d_i(0) = 1]$$

Example of the Role of Monotonicity

Unit	Y_i(d=0)	Y_i(d=1)	d_i(z=0)	d_i(z=1)	Type	Y_i(z=0)	Y_i(z=1)
1	24	34	0	1	C	24	34
2	18	28	0	1	C	18	28
3	19	32	0	1	C	19	32
4	19	26	0	1	C	19	26
5	18	22	1	0	D	22	18
6	22	28	1	0	D	28	22
7	10	20	1	1	A	20	20
8	11	12	0	0	NT	11	11
9	8	15	0	0	NT	8	8

CACE

What is the CACE here?

ATE for the compliers.

$$\frac{1}{n}\sum_{i}^{n} Y_i(C) = 10$$

In R

```
m %>%
  filter(Type == "C")%>%
  summarise(CACE = mean(`Y_i(d=1)`-`Y_i(d=0)`))%>%
  pull()
```

```
## [1] 10
```

Can we Recover the CACE with the usual formula

First we need the ITT

```
ITT <- mean(m$`Y_i(z=1)` - m$`Y_i(z=0)`)
ITT
```

```
## [1] 3
```

Now get ITT_D

```
ITT_D <- mean(m$`d_i(z=1)`-m$`d_i(z=0)`)
ITT_D
```

```
## [1] 0.2
```

Can we Recover the CACE with the usual formula

Let's remove the defiers and run the calculations again.

```
m %>%
  filter(Type != "D")%>%
  summarise(ITT =mean(`Y_i(z=1)` - `Y_i(z=0)`),
            ITT_D =  mean(`d_i(z=1)`-`d_i(z=0)`),
            CACE = ITT/ITT_D)%>%
  select(CACE)%>%
  pull()
```

```
## [1] 10
```

The Bias of the Defiers

When will Defiers really hurt our estimates?

Rewrite $ITT_D = (\pi_c + \pi_{AT}) - (\pi_D + \pi_{AT}) = \pi_c - \pi_D$

Divide the ITT by difference in proportion:

$$\frac{ITT}{ITT_D} = \frac{(ATE|Compliers)\pi_c - (ATE|Defiers) * \pi_D}{\pi_c - \pi_D}$$

Summarizing 2SNC

1) 2SNC occurs when some units in the control group are treated and some units in treatment are not treated.

2) Comparing average outcomes among units that do and do not receive treatment is a non-experimental strategy prone to bias. We should always avoid doing this.

3) 2SNC changes the interpretation of experimental estimates. We no longer get the ATE, but we can get the ITT (always) and the CACE (under appropriate assumptions)

4) We do not know who the compliers are. This gives rise to caution when generalizing results to other contexts.

5) Monotonicity is required to rule out defiers and obtain an unbiased estimate of the CACE