Gov 50: 19. Estimation: Experiments

Matthew Blackwell

Harvard University

Fall 2018

1. Today's agenda

2. Treatment effects with binary outcomes

3. Treatment effects with non-binary outcomes

1/ Today's agenda

Logistics

- Congrats on Midterm 2!
- Final project:
    - Paragraph discussing data, research question due tomorrow 11/21.
    - Draft analyses and results due Friday, 11/30.
    - Final report due 12/10.

Where are we? Where are we going?

- Last time: estimation and inference for surveys.
  - How far will the sample mean be from the population mean?
- Now: estimation and inference for comparisons between groups.

2/ Treatment effects with binary outcomes

Comparison between groups

- More interesting to compare across groups.
    - Differences in public opinion across groups
    - Difference between treatment and control groups.
- Bedrock of causal inference!

Social pressure experiment

- Back to the Social Pressure Mailer GOTV example.
  - Primary election in MI 2006
- Treatment group: postcards showing their own and their neighbors' voting records.
  - Sample size of treated group, $n_T = 360$
- Control group: received nothing.
  - Sample size of the control group, $n_C = 1890$

Outcomes

- Outcome: $X_i = 1$ if i voted, 0 otherwise.
- Turnout rate (sample mean) in treated group, $\overline{X}_T = 0.37$
- Turnout rate (sample mean) in control group, $\overline{X}_C = 0.30$
- Estimated **average treatment effect**

$$\widehat{\text{ATE}} = \overline{X}_T - \overline{X}_C = 0.07$$

Inference for the difference

- Parameter: **population ATE** $\mu_T - \mu_C$
  - μ_T: Turnout rate in the population if everyone received treatment.
  - μ_C: Turnout rate in the population if everyone received control.
- Estimator: $\widehat{\text{ATE}} = \overline{X}_T - \overline{X}_C$

- $\overline{X}_T$ is a r.v. with mean $\mathbb{E}[\overline{X}_T] = \mu_T$
- $\overline{X}_C$ is a r.v. with mean $\mathbb{E}[\overline{X}_C] = \mu_C$
- $\rightsquigarrow$ $\overline{X}_T - \overline{X}_C$ is a r.v. with mean $\mu_T - \mu_C$
  - Sample difference in means is on average equal to the population difference in means.

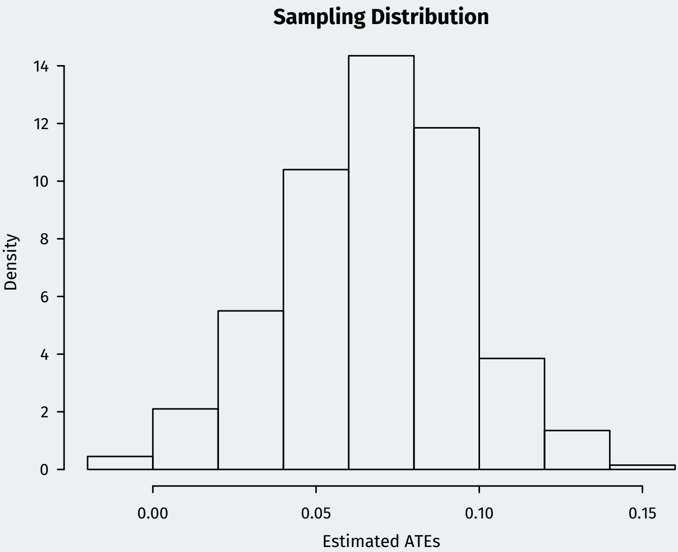
Simulation

- What if these were the true population means? We would still expect some **variation** in our estimates:

```
xt.sims <- rbinom(1000, size = 360, prob = 0.37) / 360
xc.sims <- rbinom(1000, size = 1890, prob = 0.30) / 1890

hist(xt.sims - xc.sims, freq = FALSE, xlab = "Estimated ATEs",
     main = "Sampling Distribution")
```

Simulations



Standard error

- Is an $\widehat{\text{ATE}} = 0.07$ big?
- How much variation would we expect in the difference in means across repeated samples?
- **Variance** of our estimates:

$$\mathbb{V}\left(\widehat{\text{ATE}}\right) = \mathbb{V}\left(\overline{X}_T - \overline{X}_C\right) = \mathbb{V}(\overline{X}_{\text{T}}) + \mathbb{V}(\overline{X}_{\text{C}})$$
$$= \frac{\mu_T(1-\mu_T)}{n_T} + \frac{\mu_C(1-\mu_C)}{n_C}$$

- **Standard error** is the square root of this variance:

$$\widehat{\text{SE}}_{\widehat{\text{ATE}}} = \sqrt{\frac{\overline{X}_T(1-\overline{X}_T)}{n_T} + \frac{\overline{X}_C(1-\overline{X}_C)}{n_C}} = 0.028$$

- SE represents how far, on average, $\overline{X}_T - \overline{X}_C$ will be from $\mu_T - \mu_C$.

Confidence intervals

- We can construct confidence intervals based on the CLT like last time.

$$
\begin{aligned}
CI_{95} &= \widehat{\text{ATE}} \pm 1.96 \times \widehat{\text{SE}}_{\widehat{\text{ATE}}} \\
&= 0.07 \pm 1.96 \times 0.028 \\
&= 0.07 \pm 0.054 \\
&= [0.016, 0.124]
\end{aligned}
$$

- Range of possibilities taking into account plausible chance errors.
- 0 not included in this CI $\rightsquigarrow$ chance error as big as the estimated effect unlikely.

3/ Treatment effects with non-binary outcomes

Minimum wage study revisited

- Social pressure experiment had binary outcomes $\rightsquigarrow$ special rules.
- What about general outcomes? (continuous, other discrete)
- Setting: study of how minimum wage increase in New Jersey affected employment, using Pennsylvania as a comparison group.

```
minwage <- read.csv("data/minwage.csv")

# proportion of those fully employed before and after
# the increase in the minimum wage
minwage$fullPropBefore <- minwage$fullBefore /
    (minwage$fullBefore + minwage$partBefore)
minwage$fullPropAfter <- minwage$fullAfter /
    (minwage$fullAfter + minwage$partAfter)

# separate NJ and PA
minwageNJ <- subset(minwage, subset = (location != "PA"))
minwagePA <- subset(minwage, subset = (location == "PA"))
```

Cross-section comparison

- Assume no confounders between NJ and PA
- Estimate: $\widehat{\text{ATE}} = \overline{X}_{\text{NJ}} - \overline{X}_{\text{PA}}$

```
est <- mean(minwageNJ$fullPropAfter) -
    mean(minwagePA$fullPropAfter)
est
```

```
## [1] 0.0481
```

Standard error

- Standard error of a general difference-in-means of **independent** samples is

$$\mathbb{V}\left(\widehat{\text{ATE}}\right) = \mathbb{V}\left(\overline{X}_{\text{NJ}} - \overline{X}_{\text{PA}}\right) = \mathbb{V}(\overline{X}_{\text{NJ}}) + \mathbb{V}(\overline{X}_{\text{PA}})$$

- Use this to estimate the SE:

$$\widehat{\text{SE}}_{\widehat{\text{ATE}}} = \sqrt{\widehat{\mathbb{V}}\left(\widehat{\text{ATE}}\right)} = \sqrt{\frac{\widehat{\mathbb{V}}(X_{\text{NJ}})}{n_{\text{NJ}}} + \frac{\widehat{\mathbb{V}}(X_{\text{PA}})}{n_{\text{PA}}}}$$

```
nNJ <- nrow(minwageNJ)
nPA <- nrow(minwagePA)
se <- sqrt(var(minwageNJ$fullPropAfter) / nNJ +
           var(minwagePA$fullPropAfter) / nPA)
se
```

```
## [1] 0.0336
```

Quick aside on CIs

- Confidence intervals based on CLT:

$$\widehat{\text{ATE}} \pm z_{\alpha/2} \times \widehat{\text{SE}}_{\widehat{\text{ATE}}}$$

- How do we calculate $z_{\alpha/2}$ for any possible CI?
- Plug $1 - \alpha/2$ into `qnorm()` function:
- Example: 92% CI $\leadsto \alpha = 0.08 \leadsto 1 - \alpha/2 = 0.96$

```r
# z-values for 92% CI
qnorm(0.96)
```

```
## [1] 1.75
```

Confidence intervals

- Confidence intervals based on CLT:

$$\widehat{\text{ATE}} \pm z_{\alpha/2} \times \widehat{\text{SE}}_{\widehat{\text{ATE}}}$$

```
# 90%
c(est - se * qnorm(0.95), est + se * qnorm(0.95))
```

```
## [1] -0.00715  0.10338
```

```
# 95%
c(est - se * qnorm(0.975), est + se * qnorm(0.975))
```

```
## [1] -0.0177  0.1140
```

```
# 99%
c(est - se * qnorm(0.995), est + se * qnorm(0.995))
```

```
## [1] -0.0384  0.1347
```

- These are large-sample approximations!

Before-and-after comparison

- Assumption: only change over time is the treatment
- Average changes in employment in each store before and after MW change
  - Let $Z_i = X_{i,\text{after}} - X_{i,\text{before}}$
- Estimate: $\widehat{\text{ATE}} = \frac{1}{n_{NJ}} \sum_{i=1}^{N} Z_i$

```
diffs <- minwageNJ$fullPropAfter - minwageNJ$fullPropBefore
est <- mean(diffs)
est
```

```
## [1] 0.0239
```

Standard errors for before-and-after

- Standard error: $\widehat{SE}_{\widehat{ATE}} = \sqrt{\widehat{\mathbb{V}}\left(\widehat{ATE}\right)} = \sqrt{\frac{\widehat{\mathbb{V}}(Z_i)}{n_{NJ}}}$

```
se <- sqrt(var(diffs) / nNJ)
se
```

```
## [1] 0.0176
```

- 95% confidence interval:

```
c(est - se * qnorm(0.975), est + se * qnorm(0.975))
```

```
## [1] -0.0107  0.0585
```

Next steps

- Next week: hypothesis testing
- Then regression estimation.