Gov 51: Inference for Experiments

Matthew Blackwell

Harvard University

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.

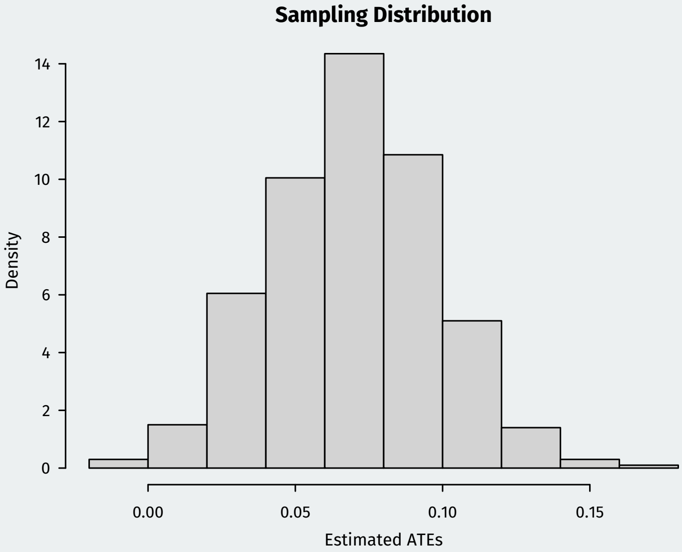
Simulation

- 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:

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

- **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.