POLI 30 D: Political Inquiry

Professor Umberto Mignozzetti
(Based on DSS Materials)

Lecture 05 | Does Social Pressure Increase Probability of Turning Out to Vote?

Before we start

Announcements:

- Quizzes and Participation: On Canvas.
- Github page: https://github.com/umbertomig/POLI30Dpublic
- Piazza forum: https://piazza.com/ucsd/winter2023/17221

Before we start

Recap:

- We learned the definitions of Theory, Scientific Theory, and Hypotheses.
- Data, datasets, variables, and how to compute means.
- Causal effect, treatments, outcomes, and randomization.

Great job!

- Do you have any questions about these contents?

Plan for Today

- Percentage Points
- Review: Unit of Measurement of Means
- Unit of Measurement of Diffs-in-Means

- In-Class Exercise: Does Social Pressure Increase the Probability of Turning Out To Vote?

What is a percentage point?

Unit of measurement for the arithmetic difference between two percentages:

% - % = p.p.

- Example: if a candidate's vote share increases from 50% to 60%, the vote share increases by

  $$\triangle \text{vshare} = \text{vshare}_{\text{final}} - \text{vshare}_{\text{initial}} = 60\% - 50\% = 10 \text{ p.p.}$$

- Why not 10%?
  - What is 10% of 50%? $0.10 \times 50 = 5$ p.p.

    $$\text{vshare}_{\text{final}} = \text{vshare}_{\text{initial}} + \triangle \text{vshare} = 50\% + 5 \text{ p.p.} = 55\%$$

Unit of Measurement of the Diffs-in-Means Estimator

unit of measurement of the difference-in-means estimator

if outcome variable is non-binary:	if outcome variable is binary:
in the same unit of measurement as the outcome variable	in percentage points (after multipying the result by 100)

Does Social Pressure Affect Turnout?

(Based on Alan S. Gerber, Donald P. Green, and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review*, 102 (1): 33-48.)

Does Social Pressure Affect Turnout?

- We will answer by analyzing data from an experiment where registered voters were randomly assigned to either
  - (a) receive a message designed to induce social pressure to vote, or
  - (b) receive nothing
- The message told that after the election, their neighbors would be informed about their voting decision:
  - What do you think the effect might be?

Does Social Pressure Affect Turnout?

- What do we need to calculate the ATE of receiving the message on the probability of turning out to vote?
  - the difference-in-means estimator
- Why does the difference-in-means estimator provide a *valid* estimate of the ATE?
  - the data come from a randomized experiment (where treatment was randomly assigned)
  - as a result, treatment and control groups are comparable

Does Social Pressure Affect Turnout?

- In this case, the difference-in-means estimator is:

$$\overline{voted}_{\text{treatment group}} - \overline{voted}_{\text{control group}}$$

- $\overline{voted}_{\text{treatment group}}$: proportion of registered voters who voted among those who received the message
- $\overline{voted}_{\text{control group}}$: proportion of registered voters who voted among those who did not receive the message

In-Class Exercise

1. Open RStudio (RStudio will open R)
2. Open **inclass01.R** from within RStudio. Find it on Canvas > Modules.
   - RStudio: File >> Open File
3. Let's Run the code step-by-step.

Does Social Pressure Affect Turnout?

```r
## STEP 1. Load the dataset
voting <- read.csv("https://raw.githubusercontent.com/umber
```

```r
## STEP 2. Look at the data
head(voting, 3) # shows the first three observations
```

```
##   birth message voted
## 1  1981      no     0
## 2  1959      no     1
## 3  1956      no     1
```

```r
## what does each observation represent?
## what is the outcome variable?
## what is the treatment variable?
```

STEP 3. Create a binary treatment variable

- First, we need to learn how to use `==`and `ifelse()`
- The operator `==` tests whether the observations of a variable are equal to a particular value
  - values should be in quotes if text but without quotes, if numbers
  - examples:
    - `data$variable==1`
    - `data$variable=="yes"`

STEP 3. Create a binary treatment variable

The function `ifelse()`

- creates the contents of a new variable based on the values of an existing one
- requires three arguments, separated by commas, in the following order:
  (1) logical test (using `==`),
  (2) return value if logical test is true,
  (3) return value if the logical test is false
- Example: `ifelse(data$variable=="yes", 1, 0)`

STEP 3. Create a binary treatment variable

```
## STEP 3. Create a binary treatment variable
## create variable pressure inside dataframe voting
voting$pressure <- # stores return values in new variable
  ifelse(voting$message=="yes", # logical test
         1, # return value if the logical test is true
         0) # return value if the logical test is false
```

- You should run the code all at once (not line by line)

STEP 3. Create a binary treatment variable

- Whenever we create a new variable, we should make sure it was created correctly by looking at it:

```
head(voting, 4) # shows first observations
```

```
##   birth message voted pressure
## 1  1981      no     0        0
## 2  1959      no     1        0
## 3  1956      no     1        0
## 4  1939     yes     1        1
```

- Note that when *message* equals "yes", *pressure* equals 1; and when *message* equals "no", *pressure* equals 0

STEP 4. Compute the difference-in-means estimator

$$\overline{Y}_{\text{treatment group}} - \overline{Y}_{\text{control group}}$$

$\overline{Y}_{\text{treatment group}}$: average outcome for the treatment group
$\overline{Y}_{\text{control group}}$: average outcome for the control group

- In the voting experiment:

$$\overline{voted}_{\text{treatment group}} - \overline{voted}_{\text{control group}}$$

- $\overline{voted}_{\text{treatment group}}$: mean of *voted* for treatment group
- $\overline{voted}_{\text{control group}}$: mean of *voted* for control group

STEP 4. Compute the difference-in-means estimator

- Let's start by practicing computing and interpreting means

```
mean(voting$voted) # calculates the mean of voted
```

```
## [1] 0.3101759
```

- Interpretation?
  - 31% of *all* the registered voters who were part of the experiment voted
- Why in %?
  - Because *voted* is binary

STEP 4. Compute the difference-in-means estimator

- `mean(voting$voted)` computes the mean of *voted* for all the observations in the dataset
- To compute the difference-in-means estimator, we need to calculate the mean of *voted* for subsets of observations
- Specifically, we need to compute:
  - the mean of *voted* for the treatment group (for which *pressure* equals 1)
  - the mean of *voted* for the control group (for which *pressure* equals 0)
- To do this, we need to learn how to use the `[]` operator

STEP 4. Compute the difference-in-means estimator

- Operator `[]`:
  - extracts a selection of observations from a variable
  - to its left, we specify the variable we want to subset
  - inside the square brackets, we specify the criteria of selection
  - example: `data$var1[data$var2==1]`
    # extracts the observations of the variable *var1* for which the variable *var2* equals 1

STEP 4. Compute the difference-in-means estimator

```
mean(voting$voted[voting$pressure == 1]) # treatment
## [1] 0.3779482
mean(voting$voted[voting$pressure == 0]) # control
## [1] 0.2966383
```

- Interpretation of the first mean?
  - 38% of the registered voters who received the message voted (38x100=38%)
- Interpretation of the second mean?
  - 30% of the registered voters who did *not* receive the message voted (30x100=30%)

STEP 4. Compute the difference-in-means estimator

```
mean(voting$voted[voting$pressure==1]) -
  mean(voting$voted[voting$pressure==0])
## [1] 0.08130991
```

What are the effect's direction, size, and unit of measurement?

STEP 5. Write conclusion statement

- What assumptions are we making when estimating the average causal effect?
  - registered voters who received the message are comparable to registered voters who did not
- Why is this a reasonable assumption?
  - data come from a randomized experiment

STEP 5. Write conclusion statement

- What's the treatment?
  - receiving the message inducing social pressure
- What's the outcome?
  - probability of voting
- What's the average causal effect's direction, size, and unit of measurement?
  - an increase of 8 percentage points, on average

CONCLUSION STATEMENT

Assuming that [the treatment and control groups are comparable] (a reasonable assumption because ...),
we estimate that [the treatment] [increases/decreases] [the outcome] by [size and unit of measurement of the effect], on average.

- Assuming that registered voters who received the message are comparable to the registered voters who did not (a reasonable assumption because the data come from a randomized experiment), we estimate that receiving the message inducing social pressure increases the probability of voting by 8 percentage points, on average.

Summary

- **Today's Class:**
  - Units of Measurement of Means and Diffs-in-Means
  - In-Class Exercise: Does Social Pressure Increase the Probability of Turning Out To Vote?
- Next class:
  - Surveys

Questions?

See you in the next class!