# Day 04: Analyzing experimental data

Erin Rossiter

January 25, 2022

# Today's plan

1. Tidy up
   - Abigail, will you:
     - remind me to take a break at 4:30-4:45
     - jot notes on typos & email after class
   - Note that HW4 will ask for a one-pager on your research topic, question, and hypothesis(es)
2. Lecture part 1 and lab 1 (hypothesis testing)
3. Lecture part 2 and lab 2 (covariates in design)

# Today's plan

1. Tidy up

- Abigail, will you:
  - remind me to take a break at 4:30-4:45
  - jot notes on typos & email after class
- Note that HW4 will ask for a one-pager on your research topic, question, and hypothesis(es)

2. Lecture part 1 and lab 1 (hypothesis testing)
3. Lecture part 2 and lab 2 (covariates in design)

# Today's plan

1. Tidy up
   - Abigail, will you:
     - » remind me to take a break at 4:30-4:45
     - » jot notes on typos & email after class
   - Note that HW4 will ask for a one-pager on your research topic, question, and hypothesis(es)
2. Lecture part 1 and lab 1 (hypothesis testing)
3. Lecture part 2 and lab 2 (covariates in design)

# Today's plan

1. Tidy up

– Abigail, will you:
  » remind me to take a break at 4:30-4:45
  » jot notes on typos & email after class

– Note that HW4 will ask for a one-pager on your research topic, question, and hypothesis(es)

2. Lecture part 1 and lab 1 (hypothesis testing)
3. Lecture part 2 and lab 2 (covariates in design)

# Today's plan

1. Tidy up

- Abigail, will you:
  - » remind me to take a break at 4:30-4:45
  - » jot notes on typos & email after class
- Note that HW4 will ask for a one-pager on your research topic, question, and hypothesis(es)

2. Lecture part 1 and lab 1 (hypothesis testing)
3. Lecture part 2 and lab 2 (covariates in design)

# Recap

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:

  1. Randomization of treatment

     » $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
     » $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
     » We can estimate left-hand terms using our observed data!

  2. Excludability
  3. Noninterference

- What about **uncertainty**?

# ATE

– The average treatment effect (ATE) is often an interesting, important, useful estimand.

– Difference-in-means is an **unbiased** estimator of the ATE assuming:

1. Randomization of treatment
   » $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
   » $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
   » We can estimate left-hand terms using our observed data!
2. Excludability
3. Noninterference

– What about **uncertainty**?

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:
  1. Randomization of treatment
     - $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
     - $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
     - We can estimate left-hand terms using our observed data!
  2. Excludability
  3. Noninterference
- What about **uncertainty**?

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:
  1. Randomization of treatment
     - $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
     - $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
     - We can estimate left-hand terms using our observed data!
  2. Excludability
  3. Noninterference
- What about **uncertainty**?

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:
  1. Randomization of treatment
     - $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
     - $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
     - We can estimate left-hand terms using our observed data!
  2. Excludability
  3. Noninterference
- What about **uncertainty**?

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:
  1. Randomization of treatment
     - $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
     - $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
     - We can estimate left-hand terms using our observed data!
  2. Excludability
  3. Noninterference
- What about **uncertainty**?

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:

  1. Randomization of treatment

     » $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
     » $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
     » We can estimate left-hand terms using our observed data!

  2. Excludability
  3. Noninterference

- What about **uncertainty**?

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:

  1. Randomization of treatment

     » $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
     » $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
     » We can estimate left-hand terms using our observed data!

  2. Excludability
  3. Noninterference

- What about **uncertainty**?

# ATE

- The average treatment effect (ATE) is often an interesting, important, useful estimand.
- Difference-in-means is an **unbiased** estimator of the ATE assuming:
    1. Randomization of treatment
        - $E[Y_i(1)|D_i = 1] = E[Y_i(1)]$
        - $E[Y_i(0)|D_i = 0] = E[Y_i(0)]$
        - We can estimate left-hand terms using our observed data!
    2. Excludability
    3. Noninterference
- What about **uncertainty**?

# Standard error

To be clear:

- a **sampling distribution** is a distribution of a statistic (e.g., $\hat{ATE}$)
- a **standard deviation** is a single value quantifying variability for any variable $X$
- the **standard error** is the standard deviation of the sampling distribution

# Standard error

To be clear:

- a **sampling distribution** is a distribution of a statistic (e.g., $A\hat{T}E$)
- a **standard deviation** is a single value quantifying variability for any variable $X$
- the **standard error** is the standard deviation of the sampling distribution

# Standard error

To be clear:

- a **sampling distribution** is a distribution of a statistic (e.g., $A\hat{T}E$)
- a **standard deviation** is a single value quantifying variability for any variable $X$
- the **standard error** is the standard deviation of the sampling distribution

# Standard error

To be clear:

- a **sampling distribution** is a distribution of a statistic (e.g., $A\hat{T}E$)
- a **standard deviation** is a single value quantifying variability for any variable $X$
- the **standard error** is the standard deviation of the sampling distribution

# Standard error

To be clear:

- a **sampling distribution** is a distribution of a statistic (e.g., $A\hat{T}E$)
- a **standard deviation** is a single value quantifying variability for any variable $X$
- the **standard error** is the standard deviation of the sampling distribution

# Standard error

There's a true standard error: $SE(\hat{ATE})$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(\hat{ATE})$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Standard error

There's a true standard error: $SE(A\hat{T}E)$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(A\hat{T}E)$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

## Standard error

There's a true standard error: $SE(\hat{ATE})$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(\hat{ATE})$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Standard error

There's a true standard error: $SE(\hat{ATE})$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(\hat{ATE})$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Standard error

There's a true standard error: $SE(A\hat{T}E)$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(A\hat{T}E)$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Standard error

There's a true standard error: $SE(A\hat{T}E)$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{S}E(A\hat{T}E)$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Standard error

There's a true standard error: $SE(\hat{ATE})$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(\hat{ATE})$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Standard error

There's a true standard error: $SE(\hat{ATE})$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(\hat{ATE})$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Standard error

There's a true standard error: $SE(\hat{ATE})$

- How much might our estimate vary across all hypothetical random assignments?
- Our design choices influence this!
  - $N$ size
  - Potential outcome variability
  - Randomization routine (blocking, clustering)
- Beneficial to reduce standard error in design phase.

Of course, we must estimate the standard error: $\hat{SE}(\hat{ATE})$

- DeclareDesign helps understand both true standard error and how we plan to estimate it
  - Is our estimation strategy getting a standard error that's too big? Meaning, we'd fail to reject a false null? No stars when there should be?!

# Today

✓ Estimating the $ATE$ with DIM

✓ Uncertainty associated with the $\hat{ATE}$

Today → Hypothesis testing

1. (Typical) null hypotheses of no *average* effect and hypothesis testing with approximate p-values based on assumptions about sampling distributions
2. Sharp null hypotheses and randomization inference hypothesis testing with exact p-values

Today → More talk about covariates

# Today

- ✓ Estimating the $ATE$ with DIM
- ✓ Uncertainty associated with the $\hat{ATE}$

Today → Hypothesis testing

1. (Typical) null hypotheses of no *average* effect and hypothesis testing with approximate p-values based on assumptions about sampling distributions
2. Sharp null hypotheses and randomization inference hypothesis testing with exact p-values

Today → More talk about covariates

# Today

✓ Estimating the $ATE$ with DIM

✓ Uncertainty associated with the $\hat{ATE}$

Today $\rightarrow$ Hypothesis testing

1. (Typical) null hypotheses of no *average* effect and hypothesis testing with approximate p-values based on assumptions about sampling distributions
2. Sharp null hypotheses and randomization inference hypothesis testing with exact p-values

Today $\rightarrow$ More talk about covariates

# Today

✓ Estimating the $ATE$ with DIM

✓ Uncertainty associated with the $A\hat{T}E$

Today → Hypothesis testing

1. (Typical) null hypotheses of no *average* effect and hypothesis testing with approximate p-values based on assumptions about sampling distributions
2. Sharp null hypotheses and randomization inference hypothesis testing with exact p-values

Today → More talk about covariates

# Today

✓ Estimating the $ATE$ with DIM

✓ Uncertainty associated with the $\hat{ATE}$

Today $\rightarrow$ Hypothesis testing

1. (Typical) null hypotheses of no *average* effect and hypothesis testing with approximate p-values based on assumptions about sampling distributions
2. Sharp null hypotheses and randomization inference hypothesis testing with exact p-values

Today $\rightarrow$ More talk about covariates

# Today

✓ Estimating the $ATE$ with DIM

✓ Uncertainty associated with the $\hat{ATE}$

Today $\rightarrow$ Hypothesis testing

1. (Typical) null hypotheses of no *average* effect and hypothesis testing with approximate p-values based on assumptions about sampling distributions
2. Sharp null hypotheses and randomization inference hypothesis testing with exact p-values

Today $\rightarrow$ More talk about covariates

# Today

✓ Estimating the $ATE$ with DIM

✓ Uncertainty associated with the $\hat{ATE}$

Today $\rightarrow$ Hypothesis testing

1. (Typical) null hypotheses of no *average* effect and hypothesis testing with approximate p-values based on assumptions about sampling distributions
2. Sharp null hypotheses and randomization inference hypothesis testing with exact p-values

Today $\rightarrow$ More talk about covariates

# Hypothesis testing

## “Traditional” hypothesis testing

# Refresher on hypothesis testing

1. We state a null hypothesis
   - e.g., $H_0$: $ATE = 0$
   - No average treatment effect
   - The claim we'd like to reject
2. We choose a test statistic
   - e.g., $t$-value
3. Determine the distribution of the test statistic under the null
   - A thought experiment. If the null is true, what data would we expect to see? (board)
   - e.g., Student's $t$
4. Calculate the probability ($p$-value) of our test statistic under the null
   - Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis
   - e.g., $H_0$: $ATE = 0$
   - No average treatment effect
   - The claim we'd like to reject
2. We choose a test statistic
   - e.g., $t$-value
3. Determine the distribution of the test statistic under the null
   - A thought experiment. If the null is true, what data would we expect to see? (board)
   - e.g., Student's $t$
4. Calculate the probability ($p$-value) of our test statistic under the null
   - Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

   - e.g., $H_0$: $ATE = 0$
   - No average treatment effect
   - The claim we'd like to reject

2. We choose a test statistic

   - e.g., $t$-value

3. Determine the distribution of the test statistic under the null

   - A thought experiment. If the null is true, what data would we expect to see? (board)
   - e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

   - Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

   – e.g., $H_0$: $ATE = 0$
   – No average treatment effect
   – The claim we'd like to reject

2. We choose a test statistic

   – e.g., $t$-value

3. Determine the distribution of the test statistic under the null

   – A thought experiment. If the null is true, what data would we expect to see? (board)
   – e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

   – Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis
   - e.g., $H_0$: $ATE = 0$
   - No average treatment effect
   - The claim we'd like to reject
2. We choose a test statistic
   - e.g., $t$-value
3. Determine the distribution of the test statistic under the null
   - A thought experiment. If the null is true, what data would we expect to see? (board)
   - e.g., Student's $t$
4. Calculate the probability ($p$-value) of our test statistic under the null
   - Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

   - e.g., $H_0$: $ATE = 0$
   - No average treatment effect
   - The claim we'd like to reject

2. We choose a test statistic

   - e.g., $t$-value

3. Determine the distribution of the test statistic under the null

   - A thought experiment. If the null is true, what data would we expect to see? (board)
   - e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

   - Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

- e.g., $H_0$: $ATE = 0$
- No average treatment effect
- The claim we'd like to reject

2. We choose a test statistic

- e.g., $t$-value

3. Determine the distribution of the test statistic under the null

- A thought experiment. If the null is true, what data would we expect to see? (board)
- e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

- Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

- e.g., $H_0$: $ATE = 0$
- No average treatment effect
- The claim we'd like to reject

2. We choose a test statistic

- e.g., $t$-value

3. Determine the distribution of the test statistic under the null

- A thought experiment. If the null is true, what data would we expect to see? (board)
- e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

- Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

- e.g., $H_0$: $ATE = 0$
- No average treatment effect
- The claim we'd like to reject

2. We choose a test statistic

- e.g., $t$-value

3. Determine the distribution of the test statistic under the null

- A thought experiment. If the null is true, what data would we expect to see? (board)
- e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

- Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

- e.g., $H_0$: $ATE = 0$
- No average treatment effect
- The claim we'd like to reject

2. We choose a test statistic

- e.g., $t$-value

3. Determine the distribution of the test statistic under the null

- A thought experiment. If the null is true, what data would we expect to see? (board)
- e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

- Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis

- e.g., $H_0$: $ATE = 0$
- No average treatment effect
- The claim we'd like to reject

2. We choose a test statistic

- e.g., $t$-value

3. Determine the distribution of the test statistic under the null

- A thought experiment. If the null is true, what data would we expect to see? (board)
- e.g., Student's $t$

4. Calculate the probability ($p$-value) of our test statistic under the null

- Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Refresher on hypothesis testing

1. We state a null hypothesis
   - e.g., $H_0$: $ATE = 0$
   - No average treatment effect
   - The claim we'd like to reject
2. We choose a test statistic
   - e.g., $t$-value
3. Determine the distribution of the test statistic under the null
   - A thought experiment. If the null is true, what data would we expect to see? (board)
   - e.g., Student's $t$
4. Calculate the probability ($p$-value) of our test statistic under the null
   - Continuing our thought experiment. If the null is true, how surprising is our data? (board)

# Example

- Data from page 65 of GG
- Treatment is encouragement to make a charitable donation
- Outcome is how much money donated

```
df <- data.frame("Donation" = c(500, 100, 100, 50,
                               25, 25, 0, 0, 0, 0,
                               25, 20, 15, 15, 10,
                               5, 5, 5, 0, 0),
                 "Treatment" = c(rep(1,10),
                                 rep(0, 10)))
head(df)
```

```
##   Donation Treatment
## 1      500         1
## 2      100         1
## 3      100         1
## 4       50         1
## 5       25         1
## 6       25         1
```

# Example

- Data from page 65 of GG
- Treatment is encouragement to make a charitable donation
- Outcome is how much money donated

```
df <- data.frame("Donation" = c(500, 100, 100, 50,
                                 25, 25, 0, 0, 0, 0,
                                 25, 20, 15, 15, 10,
                                 5, 5, 5, 0, 0),
                 "Treatment" = c(rep(1,10),
                                 rep(0, 10)))
head(df)
```

```
##   Donation Treatment
## 1      500         1
## 2      100         1
## 3      100         1
## 4       50         1
## 5       25         1
## 6       25         1
```

# Example

- Data from page 65 of GG
- Treatment is encouragement to make a charitable donation
- Outcome is how much money donated

```
df <- data.frame("Donation" = c(500, 100, 100, 50,
                                25, 25, 0, 0, 0, 0,
                                25, 20, 15, 15, 10,
                                5, 5, 5, 0, 0),
                 "Treatment" = c(rep(1,10),
                                rep(0, 10)))
head(df)
```

```
##   Donation Treatment
## 1      500         1
## 2      100         1
## 3      100         1
## 4       50         1
## 5       25         1
## 6       25         1
```

# Example

- Data from page 65 of GG
- Treatment is encouragement to make a charitable donation
- Outcome is how much money donated

```r
df <- data.frame("Donation" = c(500, 100, 100, 50,
                                25, 25, 0, 0, 0, 0,
                                25, 20, 15, 15, 10,
                                5, 5, 5, 0, 0),
                 "Treatment" = c(rep(1,10),
                                 rep(0, 10)))
head(df)
```

```
##   Donation Treatment
## 1      500         1
## 2      100         1
## 3      100         1
## 4       50         1
## 5       25         1
## 6       25         1
```

# t-test

1. State null hypothesis $H_0 : ATE < 0$
2. Choose the $t$ value as our test statistic
3. Assume sampling distribution of $t$ follows Student's $t$ under the null
   - *We're assuming sampling dist of DIM estimator has a certain shape*
4. Calculate probability of observing our test statistic (from #2) given assumed sampling distribution under the null (from #3)

# t-test

1. State null hypothesis $H_0 : ATE < 0$
2. Choose the $t$ value as our test statistic
3. Assume sampling distribution of $t$ follows Student's $t$ under the null
   - *We're assuming sampling dist of DIM estimator has a certain shape*
4. Calculate probability of observing our test statistic (from #2) given assumed sampling distribution under the null (from #3)

# t-test

1. State null hypothesis $H_0 : ATE < 0$
2. Choose the $t$ value as our test statistic
3. Assume sampling distribution of $t$ follows Student's $t$ under the null

   – *We're assuming sampling dist of DIM estimator has a certain shape*

4. Calculate probability of observing our test statistic (from #2) given assumed sampling distribution under the null (from #3)

# t-test

1. State null hypothesis $H_0 : ATE < 0$
2. Choose the $t$ value as our test statistic
3. Assume sampling distribution of $t$ follows Student's $t$ under the null

   – *We're assuming sampling dist of DIM estimator has a certain shape*

4. Calculate probability of observing our test statistic (from #2) given assumed sampling distribution under the null (from #3)

# t-test

1. State null hypothesis $H_0 : ATE < 0$
2. Choose the $t$ value as our test statistic
3. Assume sampling distribution of $t$ follows Student's $t$ under the null

   – *We're assuming sampling dist of DIM estimator has a certain shape*

4. Calculate probability of observing our test statistic (from #2) given assumed sampling distribution under the null (from #3)

# t-test by hand

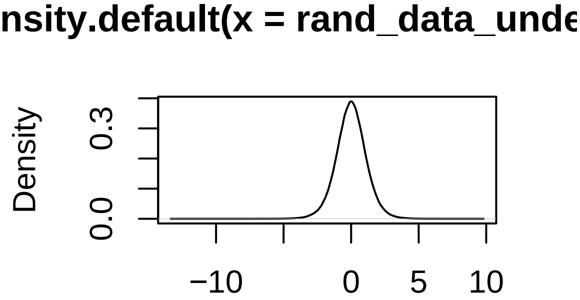
```
# 1. state null
h_not <- 0

# 2. choose test statistic (t-test with unequal variances)
s1 <- sd(df$Donation[df$Treatment == 0])^2
s2 <- sd(df$Donation[df$Treatment == 1])^2
t <- (70-h_not)/sqrt(s1/10 + s2/10)
t
```

```
## [1] 1.44776
```

# t-test by hand

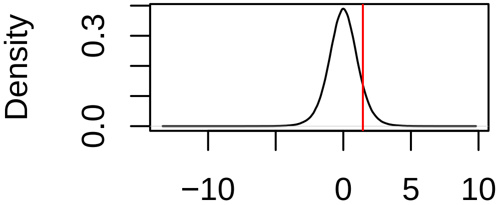
```
# 3. think about data we'd observe if the null is true
# (Notice the assumption about the sampling
# distribution we're making here!)
rand_data_under_null <- rt(n = 1000000, df = 9.0558)
plot(density(rand_data_under_null),
     xlab = "", cex = .25)
```



# t-test by hand

```
# (from step 3)
plot(density(rand_data_under_null), main = "", xlab = "", cex =

# 4. how odd is our data (or something bigger) given the null?
abline(v = t, col = "red")
```




```
pt(t, df = 9.0558, lower.tail = F) #or...
```

```
## [1] 0.09070126
```

```
sum(rand_data_under_null > t)/1000000
```

```
## [1] 0.090729
```

# t-test in R using `t.test()`

```
t.test(df$Donation[df$Treatment == 1],
       df$Donation[df$Treatment == 0],
       var.equal = F, alternative = "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  df$Donation[df$Treatment == 1] and df$Donation[df$Trea
## t = 1.4478, df = 9.0558, p-value = 0.0907
## alternative hypothesis: true difference in means is greater t
## 95 percent confidence interval:
##  -18.56994       Inf
## sample estimates:
## mean of x mean of y
##        80        10
```

# t-test in R using `difference_in_means()`

Identical, but

- no one-sided tests
- allows for easy design-based estimation when you use blocking and clustering (see homework)
- *I encourage this package since it's built with experiments in mind*

```
options(width = 60)
estimatr::difference_in_means(Donation ~ Treatment,
                              data = df)
```

```
## Design:  Standard
##           Estimate Std. Error t value  Pr(>|t|)  CI Lower
## Treatment       70   48.35057 1.44776 0.1814026 -39.27398
##           CI Upper      DF
## Treatment  179.274 9.05578
```

# t-test in R using `difference_in_means()`

Identical, but

- no one-sided tests
- allows for easy design-based estimation when you use blocking and clustering (see homework)
- *I encourage this package since it's built with experiments in mind*

```
options(width = 60)
estimatr::difference_in_means(Donation ~ Treatment,
                              data = df)
```

```
## Design:  Standard
##           Estimate Std. Error t value  Pr(>|t|)  CI Lower
## Treatment       70   48.35057 1.44776 0.1814026 -39.27398
##           CI Upper      DF
## Treatment  179.274 9.05578
```

# t-test in R using `difference_in_means()`

Identical, but

- no one-sided tests
- allows for easy design-based estimation when you use blocking and clustering (see homework)
- *I encourage this package since it's built with experiments in mind*

```
options(width = 60)
estimatr::difference_in_means(Donation ~ Treatment,
                              data = df)
```

```
## Design:  Standard
##           Estimate Std. Error t value  Pr(>|t|)  CI Lower
## Treatment       70   48.35057 1.44776 0.1814026 -39.27398
##           CI Upper      DF
## Treatment  179.274 9.05578
```

# t-test in R using `difference_in_means()`

Identical, but

- no one-sided tests
- allows for easy design-based estimation when you use blocking and clustering (see homework)
- *I encourage this package since it's built with experiments in mind*

```
options(width = 60)
estimatr::difference_in_means(Donation ~ Treatment,
                              data = df)
```

```
## Design:  Standard
##           Estimate Std. Error t value  Pr(>|t|)  CI Lower
## Treatment       70   48.35057 1.44776 0.1814026 -39.27398
##           CI Upper      DF
## Treatment  179.274 9.05578
```

# t-test in R using `difference_in_means()`

Identical, but

- no one-sided tests
- allows for easy design-based estimation when you use blocking and clustering (see homework)
- *I encourage this package since it's built with experiments in mind*

```r
options(width = 60)
estimatr::difference_in_means(Donation ~ Treatment,
                              data = df)
```

```
## Design:  Standard
##           Estimate Std. Error t value  Pr(>|t|)  CI Lower
## Treatment       70   48.35057 1.44776 0.1814026 -39.27398
##           CI Upper      DF
## Treatment  179.274 9.05578
```

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when…
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard…

Next, an alternative approach that starts from a different null hypothesis…

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# t-test big picture

- We're making assumptions about sampling distribution
  - sampling dist for DIM estimate follows $t$ distribution
- Good assumption when...
  - Outcomes distributed normally, or
  - Outcome distributed non-normally, but big sample (invoke CLT for sampling distribution to be approximately normal)
- In general, we should be wary when we're using small and/or skewed samples
  - "Rules of thumb" are hard...

Next, an alternative approach that starts from a different null hypothesis...

# Randomization inference

# Sharp null hypothesis of no effect

$$H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$$

- Think, "no effect means no effect"
- Different than no **average** treatment effect
  - No average treatment effect does not imply sharp null.
- Example on board

# Sharp null hypothesis of no effect

$$H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$$

- Think, "no effect means no effect"
- Different than no **average** treatment effect
  - No average treatment effect does not imply sharp null.
- Example on board

# Sharp null hypothesis of no effect

$$H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$$

- Think, "no effect means no effect"
- Different than no **average** treatment effect
  - No average treatment effect does not imply sharp null.
- Example on board

# Sharp null hypothesis of no effect

$$H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$$

- Think, "no effect means no effect"
- Different than no **average** treatment effect
  - No average treatment effect does not imply sharp null.
- Example on board

# Sharp null hypothesis of no effect

$$H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$$

- Think, "no effect means no effect"
- Different than no **average** treatment effect
  - No average treatment effect does not imply sharp null.
- Example on board

# Recall our hypothesis testing steps

1. We state a null hypothesis
   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$
2. We choose a test statistic
   - Keep in mind, many to choose from
   - Let's stick with DIM
3. Determine the distribution of the test statistic under the null
   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.
4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis
   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$
2. We choose a test statistic
   - Keep in mind, many to choose from
   - Let's stick with DIM
3. Determine the distribution of the test statistic under the null
   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.
4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis
   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$
2. We choose a test statistic
   - Keep in mind, many to choose from
   - Let's stick with DIM
3. Determine the distribution of the test statistic under the null
   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.
4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis
   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$
2. We choose a test statistic
   - Keep in mind, many to choose from
   - Let's stick with DIM
3. Determine the distribution of the test statistic under the null
   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.
4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis

   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$

2. We choose a test statistic

   - Keep in mind, many to choose from
   - Let's stick with DIM

3. Determine the distribution of the test statistic under the null

   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.

4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis
   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$
2. We choose a test statistic
   - Keep in mind, many to choose from
   - Let's stick with DIM
3. Determine the distribution of the test statistic under the null
   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.
4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis

– $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
– Or simply, $H_0 : Y_i(1) = Y_i(0)$

2. We choose a test statistic

– Keep in mind, many to choose from
– Let's stick with DIM

3. Determine the distribution of the test statistic under the null

– **This is where we shake things up!** Example next.
– Null hypothesis is about *individual* rather than *average* effect
– We have a link between observed data and the potential outcomes.
– **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.

4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis

- $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
- Or simply, $H_0 : Y_i(1) = Y_i(0)$

2. We choose a test statistic

- Keep in mind, many to choose from
- Let's stick with DIM

3. Determine the distribution of the test statistic under the null

- **This is where we shake things up!** Example next.
- Null hypothesis is about *individual* rather than *average* effect
- We have a link between observed data and the potential outcomes.
- **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.

4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis
   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$
2. We choose a test statistic
   - Keep in mind, many to choose from
   - Let's stick with DIM
3. Determine the distribution of the test statistic under the null
   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.
4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis

- $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
- Or simply, $H_0 : Y_i(1) = Y_i(0)$

2. We choose a test statistic

- Keep in mind, many to choose from
- Let's stick with DIM

3. Determine the distribution of the test statistic under the null

- **This is where we shake things up!** Example next.
- Null hypothesis is about *individual* rather than *average* effect
- We have a link between observed data and the potential outcomes.
- **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.

4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis

- $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
- Or simply, $H_0 : Y_i(1) = Y_i(0)$

2. We choose a test statistic

- Keep in mind, many to choose from
- Let's stick with DIM

3. Determine the distribution of the test statistic under the null

- **This is where we shake things up!** Example next.
- Null hypothesis is about *individual* rather than *average* effect
- We have a link between observed data and the potential outcomes.
- **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.

4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis

   - $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
   - Or simply, $H_0 : Y_i(1) = Y_i(0)$

2. We choose a test statistic

   - Keep in mind, many to choose from
   - Let's stick with DIM

3. Determine the distribution of the test statistic under the null

   - **This is where we shake things up!** Example next.
   - Null hypothesis is about *individual* rather than *average* effect
   - We have a link between observed data and the potential outcomes.
   - **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.

4. Calculate the probability ($p$-value) of the test statistic under the null.

# Recall our hypothesis testing steps

1. We state a null hypothesis

- $H_0 : \tau_i = Y_i(1) - Y_i(0) = 0 \quad \forall i$
- Or simply, $H_0 : Y_i(1) = Y_i(0)$

2. We choose a test statistic

- Keep in mind, many to choose from
- Let's stick with DIM

3. Determine the distribution of the test statistic under the null

- **This is where we shake things up!** Example next.
- Null hypothesis is about *individual* rather than *average* effect
- We have a link between observed data and the potential outcomes.
- **No assumptions about distributions needed;** we can calculate the sampling distribution under the null.

4. Calculate the probability ($p$-value) of the test statistic under the null.

# How do we determine the distribution of the test statistic under the null?

1. What do we know about potential outcomes from what we *observe*?

| Unit | $d_i$ | $Y_i$ | $Y_i(0)$ | $Y_i(1)$ |
|---|---|---|---|---|
| 1 | 0 | 1 | ? | ? |
| 2 | 0 | 4 | ? | ? |
| 3 | 0 | 2 | ? | ? |
| 4 | 0 | 1 | ? | ? |
| 5 | 1 | 5 | ? | ? |
| 6 | 1 | 7 | ? | ? |
| 7 | 1 | 3 | ? | ? |
| 8 | 1 | 3 | ? | ? |

# How do we determine the distribution of the test statistic under the null?

1. What do we know about potential outcomes from what we *observe*?

- $Y_i(0)|d_i = 0$
- $Y_i(1)|d_i = 1$

| Unit | $d_i$ | $Y_i$ | $Y_i(0)$ | $Y_i(1)$ |
|---|---|---|---|---|
| 1 | 0 | 1 | 1 | ? |
| 2 | 0 | 4 | 4 | ? |
| 3 | 0 | 2 | 2 | ? |
| 4 | 0 | 1 | 1 | ? |
| 5 | 1 | 5 | ? | 5 |
| 6 | 1 | 7 | ? | 7 |
| 7 | 1 | 3 | ? | 3 |
| 8 | 1 | 3 | ? | 3 |

# How do we determine the distribution of the test statistic under the null?

2. What do we know about potential outcomes from what we *assume*?

| Unit | $d_i$ | $Y_i$ | $Y_i(0)$ | $Y_i(1)$ |
|---|---|---|---|---|
| 1 | 0 | 1 | 1 | ? |
| 2 | 0 | 4 | 4 | ? |
| 3 | 0 | 2 | 2 | ? |
| 4 | 0 | 1 | 1 | ? |
| 5 | 1 | 5 | ? | 5 |
| 6 | 1 | 7 | ? | 7 |
| 7 | 1 | 3 | ? | 3 |
| 8 | 1 | 3 | ? | 3 |

# How do we determine the distribution of the test statistic under the null?

2. What do we know about potential outcomes from what we *assume*?

- We can use the sharp null to fill in remaining potential outcomes
- $Y_i(1) - Y_i(0) = 0$

| Unit | $d_i$ | $Y_i$ | $Y_i(0)$ | $Y_i(1)$ |
|---|---|---|---|---|
| 1 | 0 | 1 | 1 | (1) |
| 2 | 0 | 4 | 4 | (4) |
| 3 | 0 | 2 | 2 | (2) |
| 4 | 0 | 1 | 1 | (1) |
| 5 | 1 | 5 | (5) | 5 |
| 6 | 1 | 7 | (7) | 7 |
| 7 | 1 | 3 | (3) | 3 |
| 8 | 1 | 3 | (3) | 3 |

# How do we determine the distribution of the test statistic under the null?

3. Since we have *all* potential outcomes, we can calculate test-statistic for all hypothetical randomizations!!!

– What is this called?

# How do we determine the distribution of the test statistic under the null?

3. Since we have *all* potential outcomes, we can calculate test-statistic for all hypothetical randomizations!!!

- What is this called? **Sampling distribution**
- Also called "randomization distribution"

4. Compare our test statistic to all possible test statistics under the null to get p-value.

# Example

| Unit | $d_i$ | $Y_i$ | $Y_i(0)$ | $Y_i(1)$ |
|---|---|---|---|---|
| 1 | 0 | 1 | 1 | (1) |
| 2 | 0 | 4 | 4 | (4) |
| 3 | 0 | 2 | 2 | (2) |
| 4 | 0 | 1 | 1 | (1) |
| 5 | 1 | 5 | (5) | 5 |
| 6 | 1 | 7 | (7) | 7 |
| 7 | 1 | 3 | (3) | 3 |
| 8 | 1 | 3 | (3) | 3 |

```
df <- data.frame(Z = c(0,0,0,0,1,1,1,1),
                 Y = c(1,4,2,1,5,7,3,3))
```

# Example

- For randomization inference (often called RI), you need exact randomization procedure
- Why is this so important?

```
random_assignment <- randomizr::declare_ra(N = 8, m = 4)
random_assignment
```

```
## Random assignment procedure: Complete random assignment
## Number of units: 8
## Number of treatment arms: 2
## The possible treatment categories are 0 and 1.
## The number of possible random assignments is 70.
## The probabilities of assignment are constant across units:
## prob_0 prob_1
##    0.5    0.5
```

# Example

- For randomization inference (often called RI), you need exact randomization procedure
- Why is this so important?

```
random_assignment <- randomizr::declare_ra(N = 8, m = 4)
random_assignment
```

```
## Random assignment procedure: Complete random assignment
## Number of units: 8
## Number of treatment arms: 2
## The possible treatment categories are 0 and 1.
## The number of possible random assignments is 70.
## The probabilities of assignment are constant across units:
## prob_0 prob_1
##    0.5    0.5
```

# Example

- For randomization inference (often called RI), you need exact randomization procedure
- Why is this so important?

```
random_assignment <- randomizr::declare_ra(N = 8, m = 4)
random_assignment
```

```
## Random assignment procedure: Complete random assignment
## Number of units: 8
## Number of treatment arms: 2
## The possible treatment categories are 0 and 1.
## The number of possible random assignments is 70.
## The probabilities of assignment are constant across units:
## prob_0 prob_1
##    0.5    0.5
```

# Example

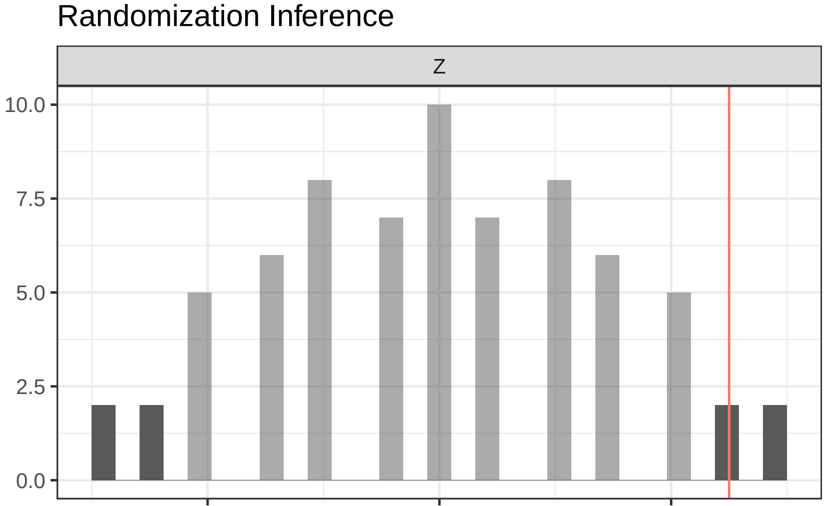
```r
out <- ri2::conduct_ri(formula = Y ~ Z,
                       data = df,
                       declaration = random_assignment)
out
```

```
##   term estimate two_tailed_p_value
## 1    Z      2.5          0.1142857
```

# Example

```
plot(out)
```

```
## Warning: It is deprecated to specify `guide = FALSE` to
## remove a guide. Please use `guide = "none"` instead.
```



# Example 2

Let's re-do the GG t-test example

```
df <- data.frame("Donation" = c(500, 100, 100, 50,
                                25, 25, 0, 0, 0, 0,
                                25, 20, 15, 15, 10,
                                5, 5, 5, 0, 0),
                 "Treatment" = c(rep(1,10),
                                 rep(0, 10)))
head(df)
```

```
##   Donation Treatment
## 1      500         1
## 2      100         1
## 3      100         1
## 4       50         1
## 5       25         1
## 6       25         1
```

# Example 2

```
random_assignment <- randomizr::declare_ra(N = 20)
random_assignment
```

```
## Random assignment procedure: Complete random assignment
## Number of units: 20
## Number of treatment arms: 2
## The possible treatment categories are 0 and 1.
## The number of possible random assignments is 184756.
## The probabilities of assignment are constant across units:
## prob_0 prob_1
##    0.5    0.5
```

# Example 2

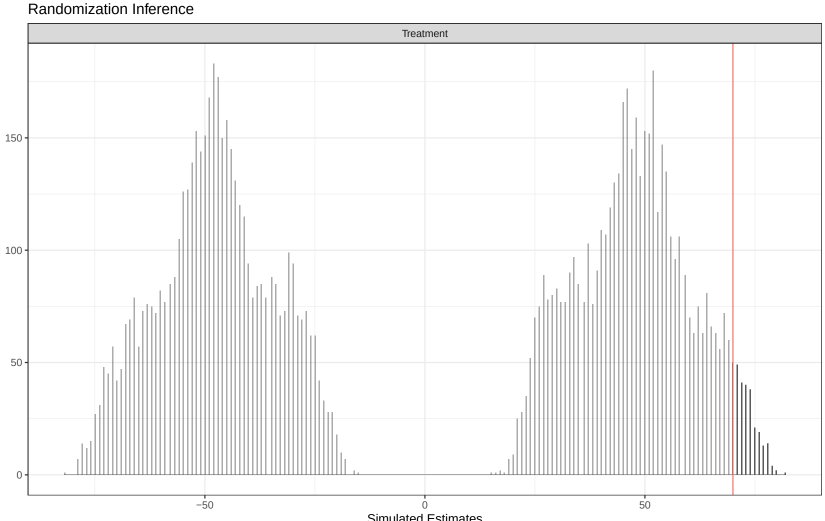
```r
out <- ri2::conduct_ri(formula = Donation ~ Treatment,
                       outcome = "Donation",
                       assignment = "Treatment",
                       p = "upper",
                       sims = 10000,
                       data = df,
                       declaration = random_assignment)
out
```

```
##        term estimate upper_p_value
## 1 Treatment       70        0.0292
```

# Example 2

```
plot(out)
```

```
## Warning: It is deprecated to specify `guide = FALSE` to
## remove a guide. Please use `guide = "none"` instead.
```



# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - » Then we can impute all potential outcomes
  - » Then we can find sampling distribution of our test stat by looking at all $D_i$
  - » Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - » Tells us nothing about individual causal effects
  - » Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Null hypotheses summary

**Sharp null hypothesis of no effect**

- Null of $Y_i(1) - Y_i(0) = 0$ allows us to say $Y_i(1) = Y_i(0)$
  - Then we can impute all potential outcomes
  - Then we can find sampling distribution of our test stat by looking at all $D_i$
  - Then we can get exact p-values!

**Null hypothesis of no average effect**

- Only allows us to say that $E[Y_i(1)] = E[Y_i(0)]$
  - Tells us nothing about individual causal effects
  - Requires assumption about sampling distribution, assumptions might not be met

Remember:

- We don't need to believe either hypothesis; we're looking for evidence *against* them!

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Pros and cons

Randomization inference:

- good when assumptions about sampling distribution are worrisome
  - small $N$
  - skewed outcomes
- good when we have complicated randomization procedures
  - blocking, clustering, etc.
- it's simple and exact!
- but null hypothesis might be less interesting
- regardless, it forces you to take a moment to think carefully about what the null hypothesis is and how it should be tested

Probably not worth it:

- "when an experiment involves random assignment of individual subjects, outcomes are distributed more or less symmetrically around the mean, and the number of subjects is greater than 100, the difference between conventional p-values and RI p-values may be negligible."

# Lab 1

# Covariates in experimental design

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - » However, covariates can have important roles in design and analysis (or can be mis-used!)
  - » Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - » Rescaling the outcome
  - » Regression adjustment
  - » Balance tests
  - » Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# Big picture

Recall,

- Randomization addresses the problem of omitted variables that plagues observational data
- Therefore, covariates are not a primary concern for causal inference
  - However, covariates can have important roles in design and analysis (or can be mis-used!)
  - Incorporation of covariates is best though-out during the design of experiments
- Roles:
  - Rescaling the outcome
  - Regression adjustment
  - Balance tests
  - Block randomization (last time)

# What are covariates?

- "Supplementary variables that predict outcomes" GG pg 95
- We assume that they are fixed pre-treatment
  - i.e., assignment to treatment and control doesn't change them
  - Usually measured pre-treatment
- Given random assignment, explain for covariate $X$ how

$$E[X_i] = E[X_i | D_i = 1] = E[X_i | D_i = 0]$$

# What are covariates?

- "Supplementary variables that predict outcomes" GG pg 95
- We assume that they are fixed pre-treatment
  - i.e., assignment to treatment and control doesn't change them
  - Usually measured pre-treatment
- Given random assignment, explain for covariate $X$ how

$$E[X_i] = E[X_i|D_i = 1] = E[X_i|D_i = 0]$$

# What are covariates?

- "Supplementary variables that predict outcomes" GG pg 95
- We assume that they are fixed pre-treatment
  - i.e., assignment to treatment and control doesn't change them
  - Usually measured pre-treatment
- Given random assignment, explain for covariate $X$ how

$$E[X_i] = E[X_i|D_i = 1] = E[X_i|D_i = 0]$$

# What are covariates?

- "Supplementary variables that predict outcomes" GG pg 95
- We assume that they are fixed pre-treatment
    - i.e., assignment to treatment and control doesn't change them
    - Usually measured pre-treatment
- Given random assignment, explain for covariate $X$ how

$$E[X_i] = E[X_i|D_i = 1] = E[X_i|D_i = 0]$$

# What are covariates?

- "Supplementary variables that predict outcomes" GG pg 95
- We assume that they are fixed pre-treatment
  - i.e., assignment to treatment and control doesn't change them
  - Usually measured pre-treatment
- Given random assignment, explain for covariate $X$ how

$$E[X_i] = E[X_i|D_i = 1] = E[X_i|D_i = 0]$$

# What are covariates?

- "Supplementary variables that predict outcomes" GG pg 95
- We assume that they are fixed pre-treatment
  - i.e., assignment to treatment and control doesn't change them
  - Usually measured pre-treatment
- Given random assignment, explain for covariate $X$ how

$$E[X_i] = E[X_i | D_i = 1] = E[X_i | D_i = 0]$$

# Use #1. Rescaling the outcome

Redefine the DV:

$$Y_i^* = Y_i - X_i$$

Difference-in-differences estimator:

$$E[\hat{ATE}] = E[(Y_i - X_i)|D_i = 1] - E[(Y_i - X_i)|D_i = 0]$$

- Prove DID is also an **unbiased estimator**.
- Choice between DID and DIM is a matter of **precision**

# Use #1. Rescaling the outcome

Redefine the DV:

$$Y_i^* = Y_i - X_i$$

Difference-in-differences estimator:

$$E[A\hat{T}E] = E[(Y_i - X_i)|D_i = 1] - E[(Y_i - X_i)|D_i = 0]$$

- Prove DID is also an **unbiased estimator**.
- Choice between DID and DIM is a matter of **precision**

# Use #1. Rescaling the outcome

Redefine the DV:

$$Y_i^* = Y_i - X_i$$

Difference-in-differences estimator:

$$E[A\hat{T}E] = E[(Y_i - X_i)|D_i = 1] - E[(Y_i - X_i)|D_i = 0]$$

- Prove DID is also an **unbiased estimator**.
- Choice between DID and DIM is a matter of **precision**

# Use #1. Rescaling the outcome

Redefine the DV:

$$Y_i^* = Y_i - X_i$$

Difference-in-differences estimator:

$$E[\hat{ATE}] = E[(Y_i - X_i)|D_i = 1] - E[(Y_i - X_i)|D_i = 0]$$

- Prove DID is also an **unbiased estimator**.
- Choice between DID and DIM is a matter of **precision**

# Use #1. Rescaling the outcome

Redefine the DV:

$$Y_i^* = Y_i - X_i$$

Difference-in-differences estimator:

$$E[\hat{ATE}] = E[(Y_i - X_i)|D_i = 1] - E[(Y_i - X_i)|D_i = 0]$$

- Prove DID is also an **unbiased estimator**.
- Choice between DID and DIM is a matter of **precision**

# Use #1. Rescaling the outcome

Redefine the DV:

$$Y_i^* = Y_i - X_i$$

Difference-in-differences estimator:

$$E[\hat{ATE}] = E[(Y_i - X_i)|D_i = 1] - E[(Y_i - X_i)|D_i = 0]$$

- Prove DID is also an **unbiased estimator**.
- Choice between DID and DIM is a matter of **precision**

# Considerations

1. Pre-treatment measure of outcome will strongly predict potential outcomes, so rescaling outcome will likely lead to large gains in precision! (lab)

   – could require smaller $N$ to be powered for same effect size

2. But, pre-treatment measurement might threaten study's other assumptions, like excludability

   – priming effects
   – social desirability bias
   – etc.

3. Budget concerns

# Considerations

1. Pre-treatment measure of outcome will strongly predict potential outcomes, so rescaling outcome will likely lead to large gains in precision! (lab)

- could require smaller $N$ to be powered for same effect size

2. But, pre-treatment measurement might threaten study's other assumptions, like excludability

- priming effects
- social desirability bias
- etc.

3. Budget concerns

## Considerations

1. Pre-treatment measure of outcome will strongly predict potential outcomes, so rescaling outcome will likely lead to large gains in precision! (lab)
   - could require smaller $N$ to be powered for same effect size
2. But, pre-treatment measurement might threaten study's other assumptions, like excludability
   - priming effects
   - social desirability bias
   - etc.
3. Budget concerns

# Considerations

1. Pre-treatment measure of outcome will strongly predict potential outcomes, so rescaling outcome will likely lead to large gains in precision! (lab)

 – could require smaller $N$ to be powered for same effect size

2. But, pre-treatment measurement might threaten study's other assumptions, like excludability

 – priming effects
 – social desirability bias
 – etc.

3. Budget concerns

# Considerations

1. Pre-treatment measure of outcome will strongly predict potential outcomes, so rescaling outcome will likely lead to large gains in precision! (lab)

- could require smaller $N$ to be powered for same effect size

2. But, pre-treatment measurement might threaten study's other assumptions, like excludability

- priming effects
- social desirability bias
- etc.

3. Budget concerns

# Considerations

1. Pre-treatment measure of outcome will strongly predict potential outcomes, so rescaling outcome will likely lead to large gains in precision! (lab)

– could require smaller $N$ to be powered for same effect size

2. But, pre-treatment measurement might threaten study's other assumptions, like excludability

– priming effects
– social desirability bias
– etc.

3. Budget concerns

# Considerations

1. Pre-treatment measure of outcome will strongly predict potential outcomes, so rescaling outcome will likely lead to large gains in precision! (lab)

- could require smaller $N$ to be powered for same effect size

2. But, pre-treatment measurement might threaten study's other assumptions, like excludability

- priming effects
- social desirability bias
- etc.

3. Budget concerns

# Use # 2. Balance tests

Experimenters bear the burden to demonstrate the soundness of their experimental procedures:

- Give details of the random assignment procedure and any instances in which subjects are lost or excluded after allocation
- Concerned about **imbalances**
  - When a covariate is correlated with treatment/control assignment
  - Present a table decribing the degree of similarity between treatment and control groups in terms of covariates
- Adjustment using covariates (next slide) re-established balance in case of chance (or error) imbalance

# Use # 2. Balance tests

Experimenters bear the burden to demonstrate the soundness of their experimental procedures:

- Give details of the random assignment procedure and any instances in which subjects are lost or excluded after allocation
- Concerned about **imbalances**
  - When a covariate is correlated with treatment/control assignment
  - Present a table decribing the degree of similarity between treatment and control groups in terms of covariates
- Adjustment using covariates (next slide) re-established balance in case of chance (or error) imbalance

# Use # 2. Balance tests

Experimenters bear the burden to demonstrate the soundness of their experimental procedures:

- Give details of the random assignment procedure and any instances in which subjects are lost or excluded after allocation
- Concerned about **imbalances**
  - When a covariate is correlated with treatment/control assignment
  - Present a table decribing the degree of similarity between treatment and control groups in terms of covariates
- Adjustment using covariates (next slide) re-established balance in case of chance (or error) imbalance

# Use # 2. Balance tests

Experimenters bear the burden to demonstrate the soundness of their experimental procedures:

- Give details of the random assignment procedure and any instances in which subjects are lost or excluded after allocation
- Concerned about **imbalances**
  - When a covariate is correlated with treatment/control assignment
  - Present a table decribing the degree of similarity between treatment and control groups in terms of covariates
- Adjustment using covariates (next slide) re-established balance in case of chance (or error) imbalance

# Use # 2. Balance tests

Experimenters bear the burden to demonstrate the soundness of their experimental procedures:

- Give details of the random assignment procedure and any instances in which subjects are lost or excluded after allocation
- Concerned about **imbalances**
  - » When a covariate is correlated with treatment/control assignment
  - » Present a table decribing the degree of similarity between treatment and control groups in terms of covariates
- Adjustment using covariates (next slide) re-established balance in case of chance (or error) imbalance

# Use # 2. Balance tests

Experimenters bear the burden to demonstrate the soundness of their experimental procedures:

- Give details of the random assignment procedure and any instances in which subjects are lost or excluded after allocation
- Concerned about **imbalances**
  - When a covariate is correlated with treatment/control assignment
  - Present a table decribing the degree of similarity between treatment and control groups in terms of covariates
- Adjustment using covariates (next slide) re-established balance in case of chance (or error) imbalance

# Use # 2. Balance tests

Experimenters bear the burden to demonstrate the soundness of their experimental procedures:

- Give details of the random assignment procedure and any instances in which subjects are lost or excluded after allocation
- Concerned about **imbalances**
  - When a covariate is correlated with treatment/control assignment
  - Present a table decribing the degree of similarity between treatment and control groups in terms of covariates
- Adjustment using covariates (next slide) re-established balance in case of chance (or error) imbalance

# Use #3. Regression adjustment

Switching equation $Y_i = Y_i(0)(1 - d_i) + Y_i(1)d_i$

Rearrange $Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i$

Rename $Y_i = \mu_{Y(0)} + [\mu_{Y(1)} - \mu Y(0)]d_i + u_i,$

- see definition of $u_i$ in text and see it matches how we think of disturbance in DeclareDesign

Rename $Y_i = a + bd_i + u_i$

- $a$ is average value of untreated potential outcomes for all units
- $b$ is average shift in average potential outcomes for i.e., ATE
- OLS estimator produces equivalent $A\hat{T}E$ as DIM

# Use #3. Regression adjustment

Switching equation $Y_i = Y_i(0)(1 - d_i) + Y_i(1)d_i$

Rearrange $Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i$

Rename $Y_i = \mu_{Y(0)} + [\mu_{Y(1)} - \mu Y(0)]d_i + u_i$,

- see definition of $u_i$ in text and see it matches how we think of disturbance in DeclareDesign

Rename $Y_i = a + bd_i + u_i$

- $a$ is average value of untreated potential outcomes for all units
- $b$ is average shift in average potential outcomes for i.e., ATE
- OLS estimator produces equivalent $A\hat{T}E$ as DIM

# Use #3. Regression adjustment

Switching equation $Y_i = Y_i(0)(1 - d_i) + Y_i(1)d_i$

Rearrange $Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i$

Rename $Y_i = \mu_{Y(0)} + [\mu_{Y(1)} - \mu Y(0)]d_i + u_i$,

- see definition of $u_i$ in text and see it matches how we think of disturbance in DeclareDesign

Rename $Y_i = a + bd_i + u_i$

- $a$ is average value of untreated potential outcomes for all units
- $b$ is average shift in average potential outcomes for i.e., ATE
- OLS estimator produces equivalent $\hat{ATE}$ as DIM

# Use #3. Regression adjustment

Switching equation $Y_i = Y_i(0)(1 - d_i) + Y_i(1)d_i$

Rearrange $Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i$

Rename $Y_i = \mu_{Y(0)} + [\mu_{Y(1)} - \mu Y(0)]d_i + u_i$,

- see definition of $u_i$ in text and see it matches how we think of disturbance in DeclareDesign

Rename $Y_i = a + bd_i + u_i$

- $a$ is average value of untreated potential outcomes for all units
- $b$ is average shift in average potential outcomes for i.e., ATE
- OLS estimator produces equivalent $\hat{ATE}$ as DIM

# Use #3. Regression adjustment

Switching equation $Y_i = Y_i(0)(1 - d_i) + Y_i(1)d_i$

Rearrange $Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i$

Rename $Y_i = \mu_{Y(0)} + [\mu_{Y(1)} - \mu Y(0)]d_i + u_i$,

- see definition of $u_i$ in text and see it matches how we think of disturbance in DeclareDesign

Rename $Y_i = a + bd_i + u_i$

- $a$ is average value of untreated potential outcomes for all units
- $b$ is average shift in average potential outcomes for i.e., ATE
- OLS estimator produces equivalent $A\hat{T}E$ as DIM

# Use #3. Regression adjustment

Switching equation $Y_i = Y_i(0)(1 - d_i) + Y_i(1)d_i$

Rearrange $Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i$

Rename $Y_i = \mu_{Y(0)} + [\mu_{Y(1)} - \mu Y(0)]d_i + u_i$,

- see definition of $u_i$ in text and see it matches how we think of disturbance in DeclareDesign

Rename $Y_i = a + bd_i + u_i$

- $a$ is average value of untreated potential outcomes for all units
- $b$ is average shift in average potential outcomes for i.e., ATE
- OLS estimator produces equivalent $\hat{ATE}$ as DIM

# Use #3. Regression adjustment

Switching equation $Y_i = Y_i(0)(1 - d_i) + Y_i(1)d_i$

Rearrange $Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i$

Rename $Y_i = \mu_{Y(0)} + [\mu_{Y(1)} - \mu Y(0)]d_i + u_i$,

- see definition of $u_i$ in text and see it matches how we think of disturbance in DeclareDesign

Rename $Y_i = a + bd_i + u_i$

- $a$ is average value of untreated potential outcomes for all units
- $b$ is average shift in average potential outcomes for i.e., ATE
- OLS estimator produces equivalent $\hat{ATE}$ as DIM

# Readjustment

Including covariate(s) on RHS

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i = a + bd_i + cX_i + (u_i - cX_i)$$

- When $X_i$ predicts the outcome, we reduce amount of unexplained variation in $Y_i$!
  - Doing so reduces the standard error of $\hat{b}$ like we talked about last time :)
- Restablishes balance if an unlucky randomization, administrative error, etc.

# Readjustment

Including covariate(s) on RHS

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i = a + bd_i + cX_i + (u_i - cX_i)$$

- When $X_i$ predicts the outcome, we reduce amount of unexplained variation in $Y_i$!
  - Doing so reduces the standard error of $\hat{b}$ like we talked about last time :)
- Restablishes balance if an unlucky randomization, administrative error, etc.

# Readjustment

Including covariate(s) on RHS

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i = a + bd_i + cX_i + (u_i - cX_i)$$

- When $X_i$ predicts the outcome, we reduce amount of unexplained variation in $Y_i$!
  - Doing so reduces the standard error of $\hat{b}$ like we talked about last time :)
- Restablishes balance if an unlucky randomization, administrative error, etc.

# Readjustment

Including covariate(s) on RHS

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i = a + bd_i + cX_i + (u_i - cX_i)$$

- When $X_i$ predicts the outcome, we reduce amount of unexplained variation in $Y_i$!
  - Doing so reduces the standard error of $\hat{b}$ like we talked about last time :)
- Restablishes balance if an unlucky randomization, administrative error, etc.

# Readjustment

Including covariate(s) on RHS

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i = a + bd_i + cX_i + (u_i - cX_i)$$

- When $X_i$ predicts the outcome, we reduce amount of unexplained variation in $Y_i$!
  - Doing so reduces the standard error of $\hat{b}$ like we talked about last time :)
- Restablishes balance if an unlucky randomization, administrative error, etc.

# Readjustment

Including covariate(s) on RHS

$$Y_i = Y_i(0) + (Y_i(1) - Y_i(0))d_i = a + bd_i + cX_i + (u_i - cX_i)$$

- When $X_i$ predicts the outcome, we reduce amount of unexplained variation in $Y_i$!
  - Doing so reduces the standard error of $\hat{b}$ like we talked about last time :)
- Restablishes balance if an unlucky randomization, administrative error, etc.

# Summary on using covariates

- Use pre-treatment covariates identified by other researchers, pilot testing, or theory to influence outcome of interest (blocking, adjustment, re-scaling outcome)
- Helps to reduce variability in our estimate
- Including covariates that does not predict outcomes does nothing to the sampling variability of our estimates
- Avoid waiting to determine which covariates to include (tie your hands using a PAP)
- Report both unadjusted estimates of the causal estimates along with adjusted estimates

# Summary on using covariates

- Use pre-treatment covariates identified by other researchers, pilot testing, or theory to influence outcome of interest (blocking, adjustment, re-scaling outcome)
- Helps to reduce variability in our estimate
- Including covariates that does not predict outcomes does nothing to the sampling variability of our estimates
- Avoid waiting to determine which covariates to include (tie your hands using a PAP)
- Report both unadjusted estimates of the causal estimates along with adjusted estimates

# Summary on using covariates

- Use pre-treatment covariates identified by other researchers, pilot testing, or theory to influence outcome of interest (blocking, adjustment, re-scaling outcome)
- Helps to reduce variability in our estimate
- Including covariates that does not predict outcomes does nothing to the sampling variability of our estimates
- Avoid waiting to determine which covariates to include (tie your hands using a PAP)
- Report both unadjusted estimates of the causal estimates along with adjusted estimates

# Summary on using covariates

- Use pre-treatment covariates identified by other researchers, pilot testing, or theory to influence outcome of interest (blocking, adjustment, re-scaling outcome)
- Helps to reduce variability in our estimate
- Including covariates that does not predict outcomes does nothing to the sampling variability of our estimates
- Avoid waiting to determine which covariates to include (tie your hands using a PAP)
- Report both unadjusted estimates of the causal estimates along with adjusted estimates

# Summary on using covariates

- Use pre-treatment covariates identified by other researchers, pilot testing, or theory to influence outcome of interest (blocking, adjustment, re-scaling outcome)
- Helps to reduce variability in our estimate
- Including covariates that does not predict outcomes does nothing to the sampling variability of our estimates
- Avoid waiting to determine which covariates to include (tie your hands using a PAP)
- Report both unadjusted estimates of the causal estimates along with adjusted estimates

# Lab 2