Problem Set 4

Alex, Micah, Scott, and David

12/01/2021

Contents

1 Consider Designs

1.1 Game night!

Suppose that you're advertising a board-game or online game to try and increase sales. You decide to individually randomly-assign into treatment and control. After you randomize, you learn that some treatment-group members are friends with control-group members IRL.

- What is the causal quantity that you would have **liked** to estimate?
- What is the causal quantity that you have **in fact** estimated?
- Is there any relationship between the two? Do you think that what you have estimated will be higher, lower, or about the same effect as the causal quantity that you would have liked to estimate?

Answer: - The ideal causal quantity would be the ATE of advertising on an individual on their probability of purchasing the game. That is, how likely an individual is to purchase a game holding their friends unexposed to the game/ad. - Given the scenario, the causal quantity that is in fact estimated is a biased ATE. That is, the treatment effect is conditional on whether an individual has friends in the other group or there is spill over across the groups. Therefore, you would want to take into account friends in the control group and estimate the ATE for treated individuals or the ATT. So the causal quantity estimated will be the ATT of ads on the probability that an individual purchases the game.
- There is a relationship between the two given both include individuals from the treatment group. The estimated ATE will likely be smaller given the fact that people in the treatment group will share information with people in the control group, making the control group look more like the treatment group.

1.2 Bonus time!

As we're writing this question, end-of-year bonuses are being given out in people's companies. (This is not a concept we your instructors have in the program – each day with your smiling faces is reward enough – and who needs money anyways?)

Suppose that you're interested in knowing whether this is a good idea from the point of view of worker productivity and so you agree to randomly assign bonuses to some people.

- What is the causal quantity that you would have **liked** to estimate?
- What is the causal quantity that you have **in fact** estimated?
- Is there any relationship between the two? Do you think that what you have estimated will be higher, lower, or about the same effect as the causal quantity that you would have liked to estimate?

Answer: … - The causal quantity that we would like to estimate is the ATE of bonuses on the an individual's productivity. - The causal quantity that we have estimated would be a biased ATE of that. We would estimate the ATT of bonuses on individual's productivity for those that received bonuses. - Yes there is a relationship. The degree of the difference will depend on the level of spill over, the motivation of the individuals, and the culture of the office. If there is spill over and the control group follows suit in terms of motivation as the treatment group (motivated to work harder or not motivated to work harder) then the estimated value would be lower than what we would have liked. If the spill over results in an opposite effect to motivation/productivity in the control group, then the estimated value will be greater than what we would have liked.

2 Noncompliance in Recycling Experiment

2.1 Intent to treat effect

What is the ITT? Do the work to compute it, and store it into the object `recycling_itt`. Provide a short narrative using inline R code, such as r inline_reference.

```
'ITT = E[Y | Z = 1] - E[ Y | Z = 0]'
```

```
## [1] "ITT = E[Y | Z = 1] - E[ Y | Z = 0]"
```

```
group_a = (600/1500)
#group_a
group_b = (800/1500)
#group_b
group_c = (3000/3000)
#group_c
Y_t_a = (500/600)
#Y_t_a
Y_t_c = (600/3000)
#Y_t_c
#I do not believe we can use the control group percentage for those that were assigned to treatment but
#recycling_itt_b <- (Y_t_a*group_a)+(Y_t_c*group_b) - (Y_t_c)
#recycling_itt_b
recycling_itt <- (Y_t_a*group_a)+(0*group_b) - (Y_t_c)
recycling_itt
```

```
## [1] 0.1333333
```

Answer: The ITT is 0.133, which tells us that the effect if receiving information about the benefits of recycling results in a 13.3% increase in recycling as compared to not receiving such information. This tells us the treatment effect for treatment/control assignments, regardless of compliance.

2.2 Compliers average causal effect

What is the CACE? Do the work to compute it, and store it into the object `recycling_cace`. Provide a short narrative using inline R code.

```
ITT_D = (700/1500)
recycling_cace <- recycling_itt/ITT_D
recycling_cace
```

```
## [1] 0.2857143
```

Answer: The CACE is 0.286, which tells us that, for compliers, receiving information about recycling resulted in a 51.4% increase in recycling as compared to not receiving the information.

2.3 Mike's CACE

What is the CACE if Mike is correct? Provide a short narrative using inline R code.

```
'ITT = E[Y | Z = 1] - E[ Y | Z = 0]'
```

```
## [1] "ITT = E[Y | Z = 1] - E[ Y | Z = 0]"
```

```
group_a = (500/1500)
group_b = (1000/1500)
group_c = (3000/3000)
Y_t_a = (500/500)
```

```
Y_t_c = (600/3000)

recycling_itt <- (Y_t_a*group_a)+(0*group_b) - (Y_t_c)
ITT_D = (500/1500)
cace_mike <- recycling_itt/ITT_D
cace_mike
```

```
## [1] 0.4
```

Answer: The CACE is 0.4, which tells us that as a result of receiving information on recycling people are 40% more likely to recycle as compared to not receiving this information. This is a ~1.2% increase in effect size as compared to the original numbers.

2.4 Andy's CACE

What is the CACE if Andy is correct? Provide a short narrative using inline R code.

```
'ITT = E[Y | Z = 1] - E[ Y | Z = 0]'
```

```
## [1] "ITT = E[Y | Z = 1] - E[ Y | Z = 0]"
```

```
group_a = (600/1500)
group_b = (900/1500)
group_c = (3000/3000)
Y_t_a = (500/600)
Y_t_c = (600/3000)

recycling_itt <- (Y_t_a*group_a)+(0*group_b) - (Y_t_c)
ITT_D = (600/1500)
cace_andy <- recycling_itt/ITT_D
cace_andy
```

```
## [1] 0.3333333
```

Answer: The CACE is 0.333, which tells us that as a result of receiving information on recycling people are 40% more likely to recycle as compared to not receiving this information. This is a ~0.7% reduction in effect size as compared to the Mike's numbers but still ~0.5% increase from the original numbers.

2.5 Effect of false reporting

What was the impact of the undergraduates's false reporting on our estimates of the treatment's effectiveness?

Answer: The impact was that our CACE estimate was below what the true estimate may be. If we are to believe either mike or andy, then we woul have underestimated the true estimate by ~0.5%-1.2%.

2.6 Effect of false reporting... on what quantity?

Does your answer change depending on whether you choose to focus on the ITT or the CACE?

Answer: The ITT remains the same but CACE changes. Given that we do not have any information on non-compliers for treatment, the ITT is not changed.

3 Fun with the placebo

3.1 Make data

Construct a data set that would reproduce the table. (Too frequently we receive data that has been summarized up to a level that is not useful for our analysis. Here, we're asking you to "un-summarize" the data to conduct the rest of the analysis for this question.)

```
generate_data <- function(assignment, treated, n, turnout) {
  d <- data.table(id = 1:n)
  n_turnout <- round(n * turnout)
  d[ , assignment := assignment]
  d[ , treated := treated]
  d[ , turnout := c(rep(1, n_turnout), rep(0, n - n_turnout))]
}
baseline_no     <- generate_data("Baseline", "No", 2463, 0.3008)
treatment_yes   <- generate_data("Treatment", "Yes", 512, 0.3890)
treatment_no    <- generate_data("Treatment", "No", 1898, 0.3160)
placebo_yes     <- generate_data("Placebo", "Yes", 476, 0.3002)
placebo_no      <- generate_data("Placebo", "No", 2108, 0.3145)

d <- rbindlist(list(
  baseline_no,
  treatment_yes,
  treatment_no,
  placebo_yes,
  placebo_no
))
```

3.2 Estimate the compliance rate using the treatment group

Estimate the proportion of compliers by using the data on the treatment group. Provide a short narrative using inline R code, such as r inline_reference.

```
#d[assignment=='Treatment' & treated == 'Yes', .N]
compliance_rate_t <- d[assignment=='Treatment' & treated == 'Yes', .N] / d[assignment=='Treatment', .N]
compliance_rate_t
```

```
## [1] 0.2124481
```

Answer: The complier rate is 0.212 or 21.2%, which tells us that ~21% of the treatment group actually received the treatment.

3.3 Estimate the compliance rate using the control group

C. Estimate the proportion of compliers by using the data on the placebo group. Provide a short narrative using inline R code.

```
compliance_rate_p <- d[assignment=='Placebo' & treated == 'No', .N] / d[assignment=='Placebo', .N]
compliance_rate_p
```

```
## [1] 0.8157895
```

Answer: he complier rate is 0.816 or 81.6%, which tells us that ~82% of the placebo group actually received the placebo. This is appears to be much higher than the treatment group.

3.4 Compare these compliance rates

Are the two compliance rates statistically significantly different from each other? Provide *a test* – this means that you cannot simply "look at" or "eyeball" the coefficients and infer some conclusion – and a description about why you chose that particular test, and why you chose that particular set of data.

```
proportions_difference_test <- prop.test(c(compliance_rate_t, compliance_rate_p), c(d[assignment=='Trea
proportions_difference_test
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(compliance_rate_t, compliance_rate_p) out of c(d[assignment == "Treatment", .N], d[assignmer
## X-squared = 7.7883e-31, df = 1, p-value = 1
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.0012359394  0.0007808289
## sample estimates:
##       prop 1       prop 2
## 8.815275e-05 3.157080e-04
```

Answer: I chose the prop.test because it can be used for testing the null that the proportions in several groups are the same. I chose the treatment data based on being placed in treatment and receiving treatment, since this shows the proportion of individuals that complied with their assignment. I chose the placebo data base on being placed in placebo and not receiving treatment, since this shows the proportion of individuals that complied with their assignment. With a p-value of 1, we fail to reject the null that the proportions are the same. In other words, the two compliance rates are not statistically significantly different from each other. This suggests that the compliance behavior does not differ meaningfully between the two groups.

3.5 Evaluate assumptions

What critical assumption does this comparison of the two groups' compliance rates test? Given what you learn from the test, how do you suggest moving forward with the analysis for this problem?

Answer: It tests the assumption of monotonicity and random assignment, or that the only difference between the groups is the assignment itself and not other factors that influence compliance. More specifically, it tests whether assignment to treatment affects compliance behavior, and whether compliance in the placebo group reflects what compliance would have looked like in the absence of treatment. Based on the results of the test, which show no evidence of statistically different compliance rates, I would continue forward with the belief that assumptions are met.

3.6 Compliers average treatement effect... of the placebo?

Estimate the CACE of receiving the placebo. Is the estimate consistent with the assumption that the placebo has no effect on turnout?

```
#'ITT = E[Y | Z = 1] - E[ Y | Z = 0]'
placebo_comply = d[assignment=='Placebo' & treated == 'No', .N] / d[assignment=='Placebo', .N]
#placebo_comply
placebo_noncomply = d[assignment=='Placebo' & treated == 'Yes', .N] / d[assignment=='Placebo', .N]
#placebo_noncomply
baseline_turnout = d[assignment=='Baseline', mean(turnout)]
#baseline_turnout
placebo_comply_turnout = d[assignment=='Placebo' & treated == 'No', mean(turnout)]
#placebo_comply_turnout
placebo_noncomply_turnout = d[assignment=='Placebo' & treated == 'Yes', mean(turnout)]
#placebo_noncomply_turnout
```

```
placebo_itt <- (placebo_comply_turnout*placebo_comply)+(placebo_noncomply_turnout*placebo_noncomply) -
cace_estimate <- placebo_itt/placebo_comply
cace_estimate
```

```
## [1] 0.01356586
```

Answer: Yes, given the very small (close to zero) CACE I would say that it meets the assumption that placebo has no effect on turnout.

3.7 Diference in means estimator

Using a difference in means (i.e. not a linear model), compute the ITT using the appropriate groups' data. Then, divide this ITT by the appropriate compliance rate to produce an estiamte the CACE. Provide a short narrative using inline R code.

```
treat_comply = d[assignment=='Treatment' & treated == 'Yes', .N] / d[assignment=='Treatment', .N]
treat_noncomply = d[assignment=='Treatment' & treated == 'No', .N] / d[assignment=='Treatment', .N]
baseline_turnout = d[assignment=='Baseline', mean(turnout)]
treat_comply_turnout = d[assignment=='Treatment' & treated == 'Yes', mean(turnout)]
treat_noncomply_turnout = d[assignment=='Treatment' & treated == 'No', mean(turnout)]
itt        <- (treat_comply_turnout*treat_comply)+(treat_noncomply_turnout*treat_noncomply) - (baseline_
cace_means <- itt/treat_comply
cace_means
```

```
## [1] 0.1444242
```

Answer: The CACE is 0.144, which tells us that, of the compliers, receiving encouragement to vote results in ~14% increase in the likelihood of voting as compared to not receiving encouragement.

3.8 Linear model estimator

Use two separate linear models to estimate the CACE of receiving the treatment by first estimating the ITT and then dividing by ITT_D. Use the `coef()` extractor and in line code evaluation to write a descriptive statement about what you learn after your code.

```
d[, Assignment_numeric := ifelse(assignment == "Treatment", 1, 0)]
d[, Treated_numeric := ifelse(treated == "Yes", 1, 0)]
itt_model   <- lm(turnout ~ assignment, data=d)
itt_d_model <- lm(Treated_numeric ~ assignment, data=d)
itt <- coef(itt_model)['assignmentTreatment']
itt_d <- coef(itt_d_model)['assignmentTreatment']
itt
```

```
## assignmentTreatment
##          0.03068265
```

```
itt_d
```

```
## assignmentTreatment
##           0.2124481
```

```
itt/itt_d
```

```
## assignmentTreatment
##           0.1444242
```

Answer: The CACE is 0.144, which tells us that, of the compliers, receiving encouragement to vote results in ~14% increase in the likelihood of voting as compared to not receiving encouragement.

3.9 Data subset estimator

When a design uses a placebo group, one additional way to estiamte the CACE is possible – subset to include only compliers in the treatment and placebo groups, and then estimate a linear model. Produce that estimate here. Provide a short narrative using inline R code.

```
treat_comply = d[assignment=='Treatment' & treated == 'Yes', ]
placebo_comply = d[assignment=='Placebo' & treated == 'No', ]
d_subset = rbind(treat_comply,placebo_comply)
itt_model   <- lm(turnout ~ assignment, data=d_subset)
itt_d_model <- lm(Treated_numeric ~ assignment, data=d_subset)
itt <- coef(itt_model)['assignmentTreatment']
itt_d <- coef(itt_d_model)['assignmentTreatment']
cace_subset_model <- itt/itt_d
cace_subset_model
```

```
## assignmentTreatment
##          0.07415575
```

Answer: The CACE is 0.0742, which tells us that, of the subset of compliers in treatment and placebo, receiving encouragement to vote results in ~7% increase in the likelihood of voting as compared to not receiving encouragement. This is less than the prior CACEs due to the sub-setting of the data.

3.10 Evaluate estimators

In large samples (i.e. "in expectation") when the design is carried out correctly, we have the expectation that the results from 7, 8, and 9 should be the same. Are they? If so, does this give you confidence that these methods are working well. If not, what explains why these estimators are producing different estimates?

Answer: They are not. 7 & 8 are the same but 9 is different. I believe that they would be providing similar answers if they had similar compliance rates. Given that there is large difference between the compliance rates, by sub-setting on the complier data we are not doing an apples to apples comparison since it would indicate that the compliers in treatment are different from compliers in placebo.

4 Another Turnout Question

4.1 Simple treatment effect

Load the data and estimate a `lm` model that compares the rates of turnout in the control group to the rate of turnout among anybody who received *any* letter. This model combines all the letters into a single condition – "treatment" compared to a single condition "control". Report robust standard errors, and include a narrative sentence or two after your code using inline R code, such as r inline_reference.

Answer: The results of this model tells us that the estimated effect of receiving any letter is 0.001, or 0.1% increase to turnout. However, this coefficient is not stat sig and has a SE at a scale that is greater than the coefficient itself.

4.2 Letter-specific treatment effects

Suppose that you want to know whether different letters have different effects. To begin, what are the effects of each of the letters, as compared to control? Estimate an appropriate linear model and use robust standard errors. Provide a short narrative using inline R code.

Answer: The results show varying signs and scales for the coefficients but they all remain stat insignificant with large SE.

4.3 Test for letter-specific effects

Does the increased flexibilitiy of a different treatment effect for each of the letters improve the performance of the model? Test, using an F-test. What does the evidence suggest, and what does this mean about whether there **are** or **are not** different treatment effects for the different letters?

Answer: The increased flexibility does not result in any improvement to our model. This suggests that the simpler model provides just as much insight as the other. This further's the evidence that there are no observable treatment effects from the different types of letters received.

4.4 Compare letter-specific effects

Is one message more effective than the others? The authors have drawn up this design as a full-factorial design. Write a *specific* test for the difference between the *Partisan* message and the *Election Info* message. Write a *specific* test for the difference between *Top-Two Info* and the *Election Info* message. Report robust standard errors on both tests and include a short narrative statement after your estimates.

Answer: Based on the results, each of the coefficients tested remain stat insignificant. This tells us that both "Top-two info" and "Partisan" are not stat sig different or more effective than "election info".

4.5 Count the number of blocks

Blocks? We don't need no stinking blocks? The blocks in this data are defined in the `block.num` variable (which you may have renamed). There are a *many* of blocks in this data, none of them are numerical – they're all category indicators. How many blocks are there?

```
length(d[, unique(block)])
```

```
## [1] 283
```

```
#length(unique(d[, block]))
```

Answer: There are 283 blocks.

4.6 Add block fixed effects

SAVE YOUR CODE FIRST but then try to estimate a `lm` that evaluates the effect of receiving *any letter*, and includes this block-level information. What happens? Why do you think this happens? If this estimate *would have worked* (that's a hint that we don't think it will), what would the block fixed effects have accomplished?

Answer: When controling for blocks, the results become too long to review. This happens because the model is controlling for the nearly 300 different blocks, creating estimates for each. This is not good practice, since it can lead to p hacking. If it had worked, what it would tell us is whether the block has any impact on voter turnout.

4.7 A clever work-around?

Even though we can't estimate this fixed effects model directly, we can get the same information and model improvement if we're *just a little bit clever*. Create a new variable that is the *average turnout within a block* and attach this back to the data.table. Use this new variable in a regression that regresses voting on `any_letter` and this new `block_average`. Then, using an F-test, does the increased information from all these blocks improve the performance of the *causal* model? Use an F-test to check.

Answer: The clever maneuver does us to more easily estimate what we are looking for. The F-test results in stat sig. which tells us that blocks do have an effect on vote turnout and in fact provide us more power in our model.

4.8 Does cleverness create a bad-control?

Doesn't this feel like using a bad-control in your regression? Has the treatment coefficient changed from when you didn't include the `block_average` measure to when you did? Have the standard errors on the treatment coefficient changed from when you didn't include the `block_average` measure to when you did? Why is this OK to do?

Answer: It does kind of feel odd including it as a control, but I'm not sure if I would consider it a bad control just yet. The treatment coefficient has not changed after including the block average and neither has the SE. This is ok to do because this is not a post-treatment variable. Block is also not directly impacted by treatment. Therefore, including this type of variable is ok.

5 Optional Turnout in Dorms

5.1 Use Linear Regressions

1. Estimate the ITT using a linear regression on the appropriate subset of data. Notice that there are two `NA` in the data. Just na.omit to remove these rows so that we are all working with the same data. Given the ways that randomization was conducted, what is the appropriate way to construct the standard errors?

```
dorm_model <- 'fill this in'
```

5.2 Use Randomization Inference

1. How many people are in treatment and control? Does this give you insight into how the scientists might have randomized? As usual, include a narrative sentence after your code.

```
n_treatment <- 'fill this in'
n_control   <- 'fill this in'
```

Narrative: ...

2. Write an algorithm to conduct the Randomization Inference. Be sure to take into account the fact that random assignment was clustered by dorm room.

```
# Use this block for your work
```

3. What is the value that you estimate for the treatment effect?

```
dorm_room_cace <- 'fill this in'
```

Narrative: ...

4. What are the 2.5% and 97.5% quantiles of this distribution?

```
dorm_room_ci <- 'fill this in with a length-two vector; first number 2.5%, second number 97.5%'
```

Narrative: ...

5. What is the p-value that you generate for the test: How likely is this treatment effect to have been generated if the sharp null hypothesis were true.

```
p_value <- 'fill this in'
```

Narrative: ...

6. Assume that the leaflet (which was left in case nobody answered the door) had no effect on turnout. Estimate the CACE either using ITT and ITT_d or using a set of linear models. What is the CACE, the estimated standard error of the CACE, and the p-value of the test you conduct?

```
dorm_room_cace <- 'fill this in'
```

Narrative: ...

7. What if the leaflet that was left actually *did* have an effect? Is it possible to estimate a CACE in this case? Why or why not?

Narrative: ...