

Homework 4 Answers

February 2, 2022

Instructions

- This homework is due Wednesday, February 9 at 3pm EST.
- Submit via GitHub. Remember to commit and push online so I can see it.
- Please format your homework solutions using R Markdown. You are welcome to simply add your answers below each question.
 - If the question requires a figure, make sure you have informative title, axis labels, and legend if needed.
 - Note: When I've given the framework of an answer's code, I've included the option `eval=FALSE` in the R chunk. When you start filling in your answer, you'll need to switch this to `eval=TRUE`.
- Turn in both the `.rmd` file and the knitted `.pdf` or `.html` file.
 - Knitting the `.rmd` file to a `.pdf` or `.html` file should help ensure your code runs without errors, but double check the output is what you expected.

Question 1

Conduct randomization inference hypothesis test of the sharp null hypothesis for the following data *by hand*. It's an experiment that used complete randomization to assign 2 of 4 units to treatment. The data from the experiment are stored in `df`.

I've shown what the results are using the `conduct_ri` function. Your task is to replicate these results. Specifically, you should calculate the sampling distribution, plot it as a histogram, indicate where the observed estimate (our test-statistic) falls on the histogram, and report a two-tailed pvalue.

Note: The number of possible random assignments isn't that big, so don't feel like you have to write a for loop for this question. You can if you want, of course. I'm more concerned with you gaining an intuition for what's going on with randomization inference.

```
df <- data.frame("Z" = c(0,0,1,1),
                 "Y" = c(4,6,2,2))
ra <- randomizr::declare_ra(N = 4, m = 2)
ra

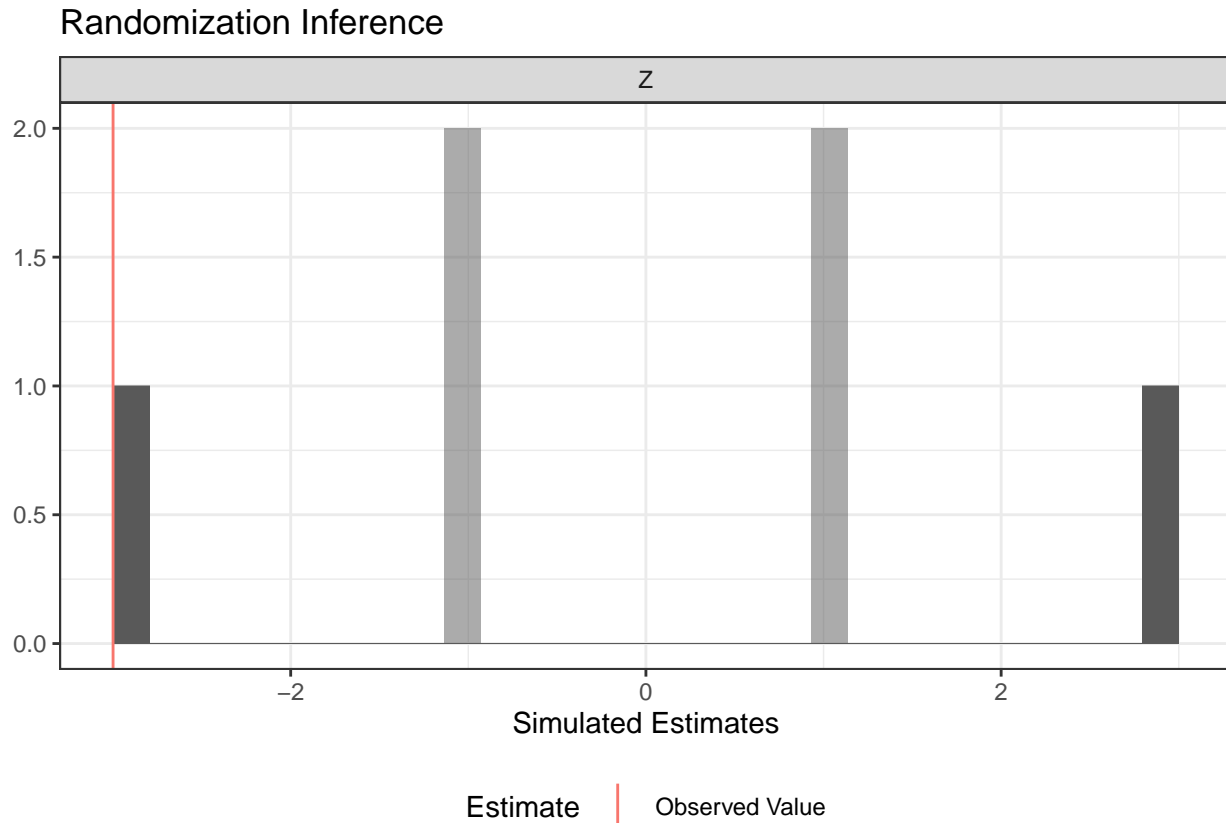
## Random assignment procedure: Complete random assignment
## Number of units: 4
## Number of treatment arms: 2
## The possible treatment categories are 0 and 1.
## The number of possible random assignments is 6.
## The probabilities of assignment are constant across units:
## prob_0 prob_1
##      0.5    0.5

out <- ri2::conduct_ri(formula = Y ~ Z,
                       sharp_hypothesis = 0,
                       data = df,
                       declaration = ra,
                       outcome = "Y",
```

```
assignment = "Z")
out
```

```
## term estimate two_tailed_p_value
## 1 Z -3 0.3333333
```

```
plot(out)
```



Answer:

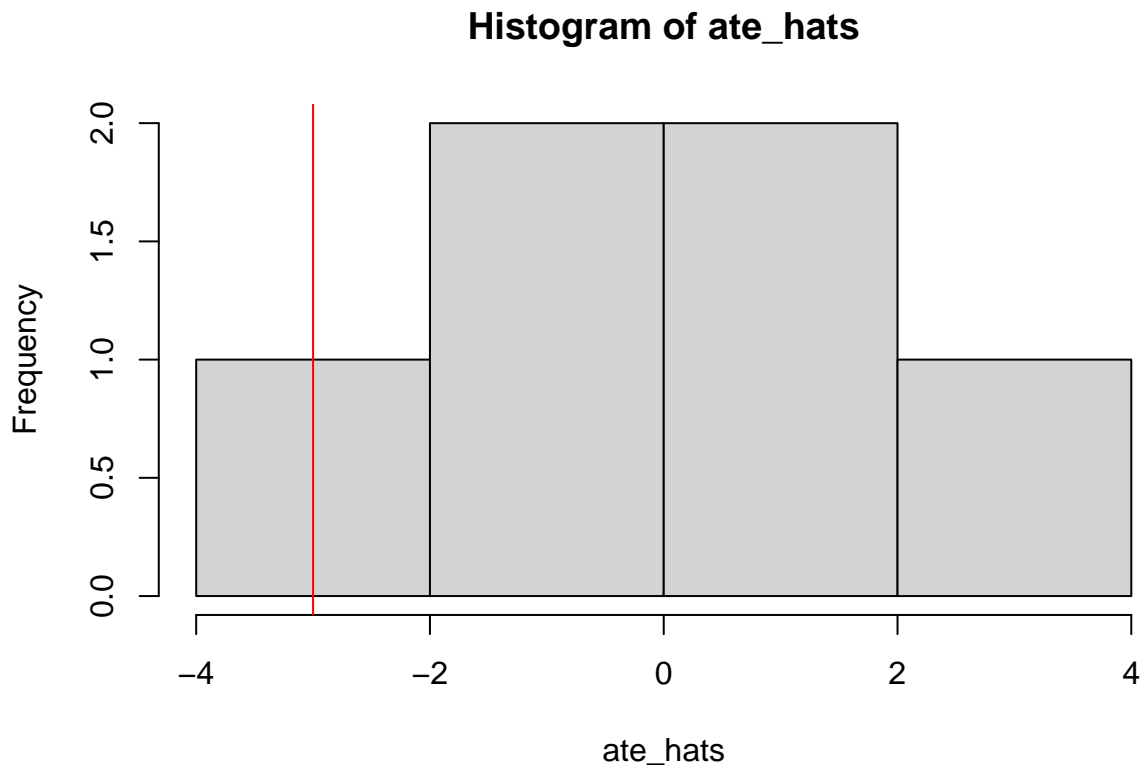
First, I manually list all possible random assignments consistent with the randomization procedure (complete RA) and store them in a list called D. Then, for each possible randomization, I calculate \hat{ATE} and store it.

```
D <- list(d1 = c(0,0,1,1),
          d2 = c(0,1,0,1),
          d3 = c(0,1,1,0),
          d4 = c(1,1,0,0),
          d5 = c(1,0,0,1),
          d6 = c(1,0,1,0))
ate_hats <- c()
for(d in D){
  ate_hats <- c(ate_hats, mean(df$Y[d==1]) - mean(df$Y[d==0]))
}
```

Next, I make a histogram of ATE estimates across all possible randomizations consistent with the design under the null, on other words, I plot the sampling distribution.

Then, I calculate the *observed* estimate and add it to the histogram.

```
hist(ate_hats)
observed_est <- mean(df$Y[df$Z==1]) - mean(df$Y[df$Z==0])
abline(v = observed_est, col = "red")
```



Finally, I report a pvalue. The pvalue is the proportion of \hat{ATE} 's that are at least as big as our observed estimate. The sampling distribution plots data we would observe if the sharp null of no effect is true. When we compare our observed estimate to the sampling distribution, we can see how suprising our result is *if the null is true*. If we wouldn't expect our result under the null (e.g., $p < .05$) we can reject the null!

```
sum(abs(ate_hats) >= abs(observed_est)) / length(ate_hats)
```

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

Question 2

Below I've read in a dataset called `alo_star_men` included in the `estimatr` packate. The data is from the following paper:

Lin, Winston. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics*. Stat. 7(1): 295-318. doi:10.1214/12-AOAS583.

You can see the definitions of each variable in the help file associated with the dataset.

Here is the abstract from the paper:

This paper reports on an experimental evaluation of strategies designed to improve academic performance among college freshmen. One treatment group was offered academic support services. Another was offered financial incentives for good grades. A third group combined both interventions. Service use was highest for women and for subjects in the combined group. The combined treatment also raised the grades and improved the academic standing of women. These differentials persisted through the end of second year, though incentives were given in the first

year only. This suggests study skills among some treated women increased. In contrast, the program had no effect on men.

According to the help file, “This data was originally taken from the following paper, subset to men who showed up to college, were in one of the arms with the support condition, and had GPA data for their first year in college.”

Reading the abstract and the dataset help file (`?alo_star_men`) should provide enough information to answer the following questions. I’ll add one note that treatment *was* randomized, and since I couldn’t find a lot of detail in the paper, let’s just assume they used complete randomization.

Your task is to assess whether one important covariate—high school GPA—is balanced across the two experimental conditions.

```
library(estimatr)
data(alo_star_men)
head(alo_star_men)
```

```
##      gpa0 sfsp ssp GPA_year1 GPA_year2
## 17  82.0   0   1      2.25      2.70
## 19  84.8   1   0      3.01      1.95
## 51  78.0   0   1      2.48      2.36
## 95  72.8   1   0      0.99      1.31
## 114 84.7   0   1      1.10      1.77
## 118 87.3   1   0      1.38      2.28
```

2a

Conduct a t-test to assess balance on this covariate and interpret the results. Is this variable balanced? Are you surprised by the results? Why or why not?

Answer:

```
estimatr::difference_in_means(gpa0 ~ sfsp, data = alo_star_men)
```

```
## Design: Standard
##      Estimate Std. Error   t value Pr(>|t|)    CI Lower CI Upper      DF
## sfsp 0.5454715    0.714166 0.7637881 0.4464439 -0.8680589 1.959002 124.0369
```

The `gpa0` variable has a range of 71.3 to 87.3 in the dataset. We see from the t-test, that the difference in the average of this variable across conditions is 0.55, which seems like a small difference. Indeed, it is **not** a significant difference ($p=.45$).

We know that if treatment is randomized, then $E[GPA] = E[GPA|d_i = 1] = E[GPA|d_i = 0]$. We also know our estimates of $E[GPA|d_i = 1]$ and $E[GPA|d_i = 0]$ using the sample mean are unbiased. So with an N of 157, we should expect randomization to do its job, and mean GPA to be similar across conditions. Although, since we only get one experiment, it’s important to check that we didn’t get an unlucky randomization, there wasn’t an administrative error, etc.

2b

Comment on why balance on high school GPA is critical to the causal inferences we draw from this study.

Answer:

If GPA was correlated with treatment administration and the variable was imbalanced, then we might have data where people with higher high school GPA’s got the treatment with both financial incentives and the support program, confounding our results. We wouldn’t be able to say whether the students did better in college because they were already better students or because of the financial incentives and support program treatment.

2c

Consider an experiment you've recently read in a class, and experiment you've conducted before, or an experiment you plan to conduct.

Question 3

Please write up your initial thoughts on your final paper. You can write in whatever text processor you prefer. Email me the document if you'd prefer writing in something other than the .rmd file.

Include, at a minimum, thoughts on the topic, the research question, and one or more hypotheses. You might structure this one-pager by including one paragraph motivating the topic, one paragraph explaining your research question, and one paragraph motivating your hypotheses. Please also formally write your hypotheses, for example, " H_1 : Partisan media increases affective polarization." I should be able to know the important concepts in your study and identify what is X and what is Y .

Keep in mind two things:

1. The goal of the one-pager is to check in with where you're at. Nothing is set in stone!
2. This is your first opportunity for feedback on your idea, so if you want to include *more* than the topic, question, and hypotheses, please do. For instance, if you have in mind the sample, the experimental conditions, etc. feel free to write it up.