Gerber and Green (2012) Chapter 3 Problem 9

Margaret Moor and Alexander Coppock, Yale University

January 30, 2018

This script shows how to conduct the randomization inference procedure in Gerber and Green (2012) Chapter 3 Problem 9 three different ways: using the `ri2` package, using the `ri` package, and by hand with a loop.

Problem 3.9

Camerer reports the results of an experiment in which he tests whether large, early bets placed at horse tracks affect the betting behavior of other bettors. Selecting pairs of long-shot horses running in the same race whose betting odds were approximately the same when betting opened, he placed two $500 bets on one of the two horses approximately 15 minutes before the start of the race. Because odds are determined based on the proportion of total bets placed on each horse, this intervention causes the betting odds for the treatment horse to decline and the betting odds of the control horse to rise. Because Camerer's bets were placed early, when the total betting pool was small, his bets caused marked changes in the odds presented to other bettors. (A few minutes before each race started, Camerer cancelled his bets.) While the experimental bets were still "live," were other bettors attracted to the treatment horse (because other bettors seemed to believe in the horse) or repelled by it (because the diminished odds meant a lower return for each wager)? Seventeen pairs of horses in this study are listed below. The outcome measure is the number of dollars that were placed on each horse (not counting Camerer's own wagers on the treatment horses) during the test period, which begins 16 minutes before each race (roughly 2 minutes before Camerer began placing his bets) and ends 5 minutes before each race (roughly 2 minutes before Camerer withdrew his bets).

(a) One interesting feature of this study is that each pair of horses ran in the same race. Does this design feature violate the non-interference assumption, or can potential outcomes be de ned so that the non-interference assumption is satisfied?

NOT SHOWN

(b) A researcher interested in conducting a randomization check might assess whether, as expected, treatment and control horses attract similarly sized bets prior to the experimental intervention. Use randomization inference to test the sharp null hypothesis that the bets had no effect prior to being placed.

SEE BELOW

(c) Calculate the average increase in bets during the experimental period for treatment horses and control horses. Compare treatment and control means, and interpret the estimated ATE.

NOT SHOWN

(d) Show that the estimated ATE is the same when you subtract the control group outcome from the treatment group outcome for each pair, and calculate the average difference for the 17 pairs.

NOT SHOWN

(e) Use randomization inference to test the sharp null hypothesis of no treatment effect for any subject. When setting up the test, remember to construct the simulation to account for the fact that random assignment takes place within each pair. Interpret the results of your hypothesis test and explain why a two-tailed test is appropriate in this application.

```
# Data from http://isps.yale.edu/FEDAI
library(haven)
library(randomizr)
```

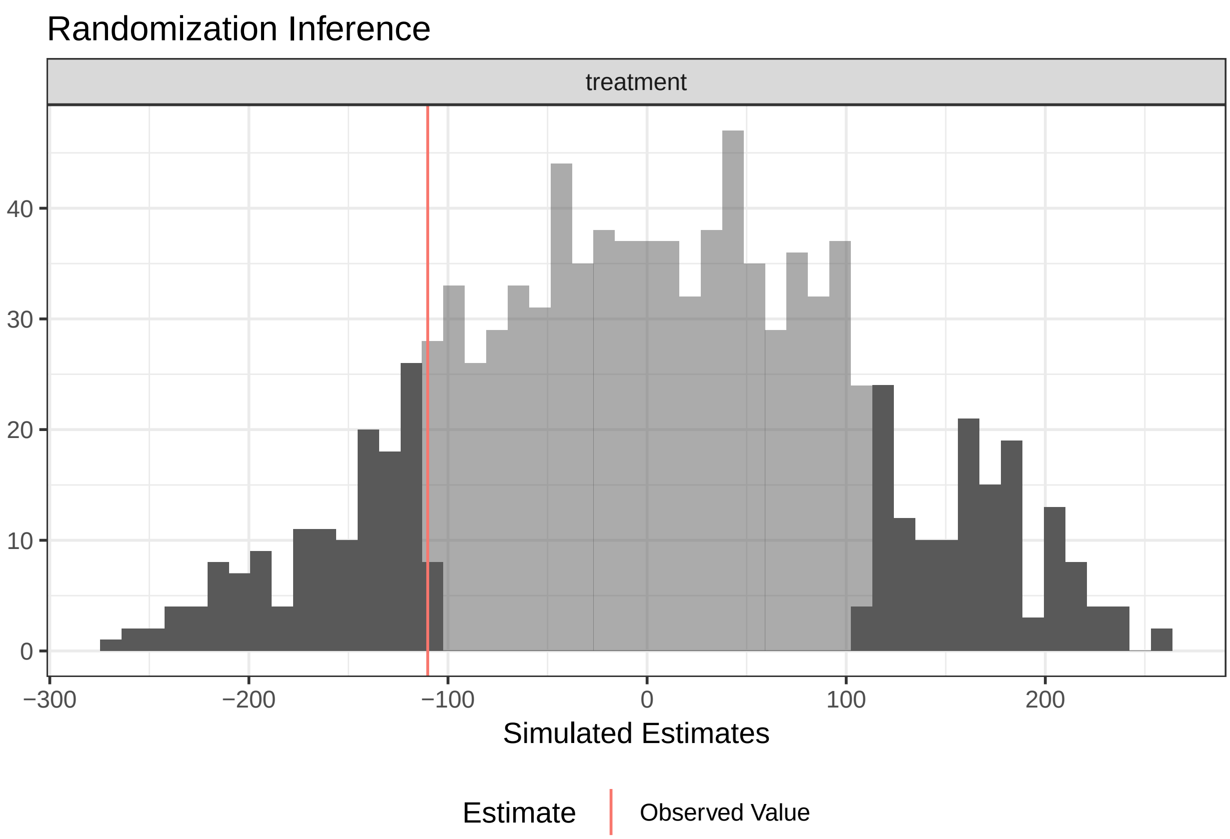
```r
data3.9 <- read_dta("datasets/3.9.dta")
# Number of sims the same for all three methods
sims <- 1000
```

In ri2

```r
library(ri2)

# Declare randomization procedure
declaration <- declare_ra(N = 34, blocks = data3.9$pair)

# Conduct Randomization Inference
ri2_out <- conduct_ri(experimentbets ~ treatment,
                      declaration = declaration,
                      assignment = "treatment",
                      sharp_hypothesis = 0,
                      sims = sims,
                      data = data3.9)

summary(ri2_out)
```

```
##   coefficient  estimate two_tailed_p_value null_ci_lower null_ci_upper
## 1   treatment -110.1765              0.294     -202.3118      205.3618
```

```r
plot(ri2_out)
```

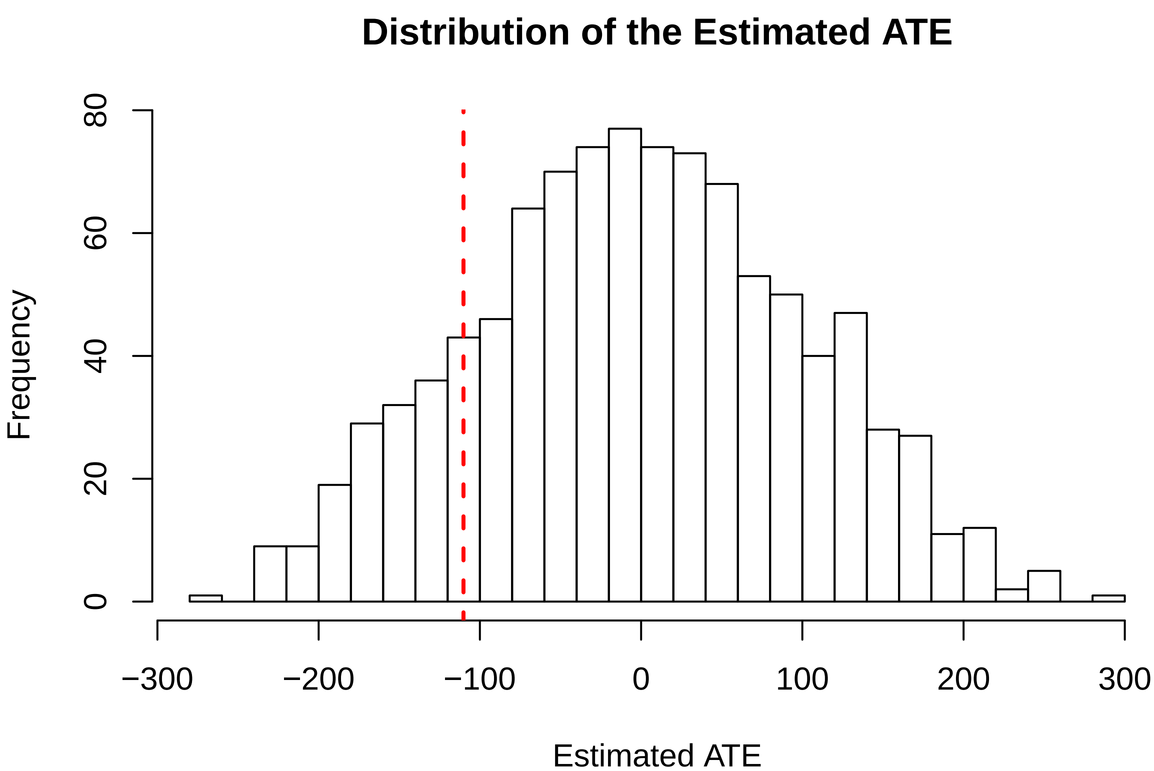


In ri

```r
library(ri)

# all possible permutations
perms <- genperms(data3.9$treatment, blockvar = data3.9$pair, maxiter = sims)
```

```
## Too many permutations to use exact method.
## Defaulting to approximate method.
## Increase maxiter to at least 131072 to perform exact estimation.
```

```r
# probability of treatment
probs <- genprobexact(data3.9$treatment, blockvar = data3.9$pair)
# estimate the ATE
ate <- estate(data3.9$experimentbets, data3.9$treatment, prob = probs)

## Conduct Sharp Null Hypothesis Test of Zero Effect for Each Unit

# generate potential outcomes under sharp null of no effect
Ys <- genouts(data3.9$experimentbets, data3.9$treatment, ate = 0)
# generate sampling dist. under sharp null
distout <- gendist(Ys, perms, prob = probs)
# display characteristics of sampling dist. for inference
ri_out <- dispdist(distout, ate)
```




```r
ri_out
```

```
## $two.tailed.p.value
## [1] 0.308
##
```

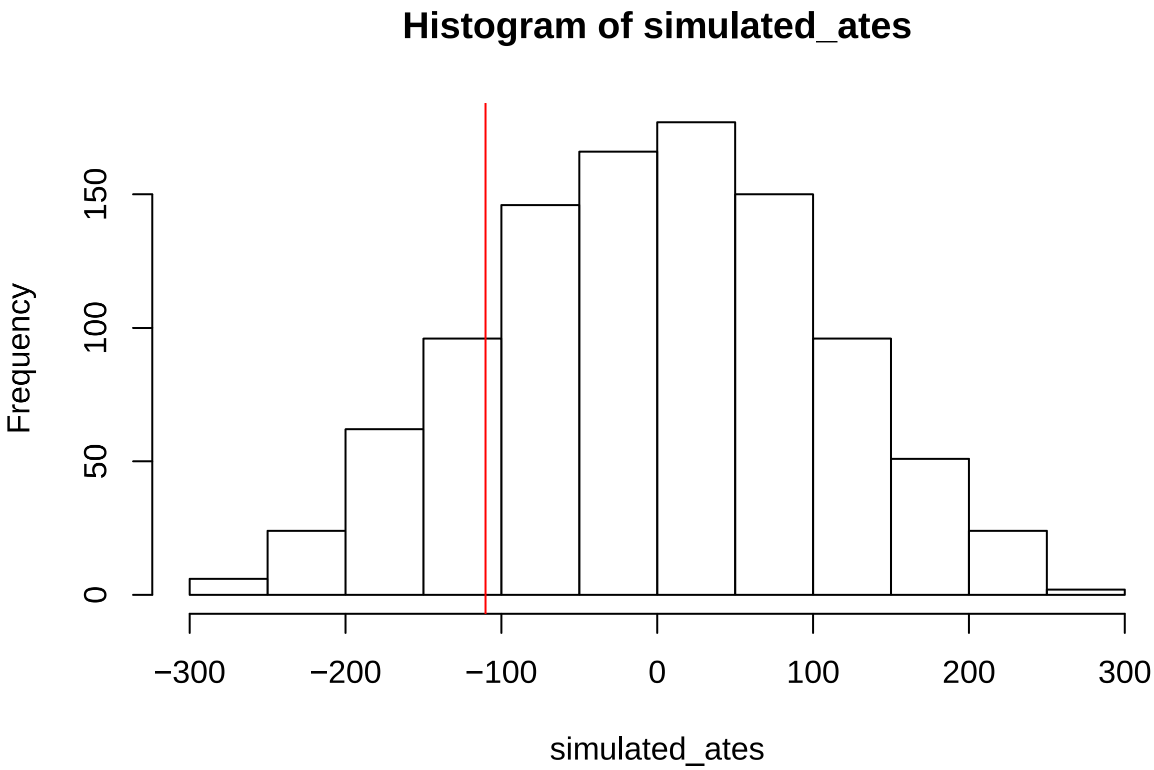
```
## $two.tailed.p.value.abs
## [1] 0.302
##
## $greater.p.value
## [1] 0.846
##
## $lesser.p.value
## [1] 0.154
##
## $quantile
##      2.5%     97.5%
## -192.7676  186.9029
##
## $sd
## [1] 101.206
##
## $exp.val
## [1] -1.976824
```

By hand

```
library(randomizr)
observed_ate <-
  with(data3.9, mean(experimentbets[treatment == 1]) - mean(experimentbets[treatment == 0]))

simulated_ates <- rep(NA, sims)

for (i in 1:sims) {
  data3.9$Z_sim <- with(data3.9, block_ra(blocks = pair))
  simulated_ates[i] <-
    with(data3.9, mean(experimentbets[Z_sim == 1]) - mean(experimentbets[Z_sim == 0]))
}

p_two_tailed <- mean(abs(simulated_ates) >= abs(observed_ate))
p_upper <- mean(simulated_ates >= observed_ate)
p_lower <- mean(simulated_ates <= observed_ate)

c(observed_ate, p_two_tailed, p_upper, p_lower)
```

```
## [1] -110.1765    0.3190    0.8330    0.1670
```

```
hist(simulated_ates, breaks = 10)
abline(v = observed_ate, col = "red")
```


Histogram of simulated_ates
Frequency
0
50
100
150
−300
−200
−100
0
100
200
300
simulated_ates