

# Gerber and Green (2012) Chapter 3 Problem 7

*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 7 three different ways: using the `ri2` package, using the `ri` package, and by hand with a loop.

## Chapter 3 Problem 7

A diet and exercise program advertises that it causes everyone who is currently starting a diet to lose at least seven pounds more than they otherwise would have during the first two weeks. Use randomization inference (the procedure described in section 3.4) to test the hypothesis that  $\tau_i = 7$  for all  $i$ . The treatment group's weight losses after two weeks are  $\{2, 11, 14, 0, 3\}$  and the control group's weight losses are  $\{1, 0, 0, 4, 3\}$ . In order to test the hypothesis  $\tau_i = 7$  for all  $i$  using the randomization inference methods discussed in this chapter, subtract 7 from each outcome in the treatment group so that the exercise turns into the more familiar test of the sharp null hypothesis that  $\tau_i = 0$  for all  $i$ . When describing your results, remember to state the null hypothesis clearly, and explain why you chose to use a one-sided or two-sided test.

```
# Create dataset
d <- rep(c(0, 1), c(5, 5))
y <- rep(c(1, 0, 0, 4, 3, -5, 4, 7, -7, -4))
data3.7 <- data.frame(d, y)
```

## In ri2

```
library(ri2)

# Declare randomization procedure
declaration <- declare_ra(N = 10, m = 5)

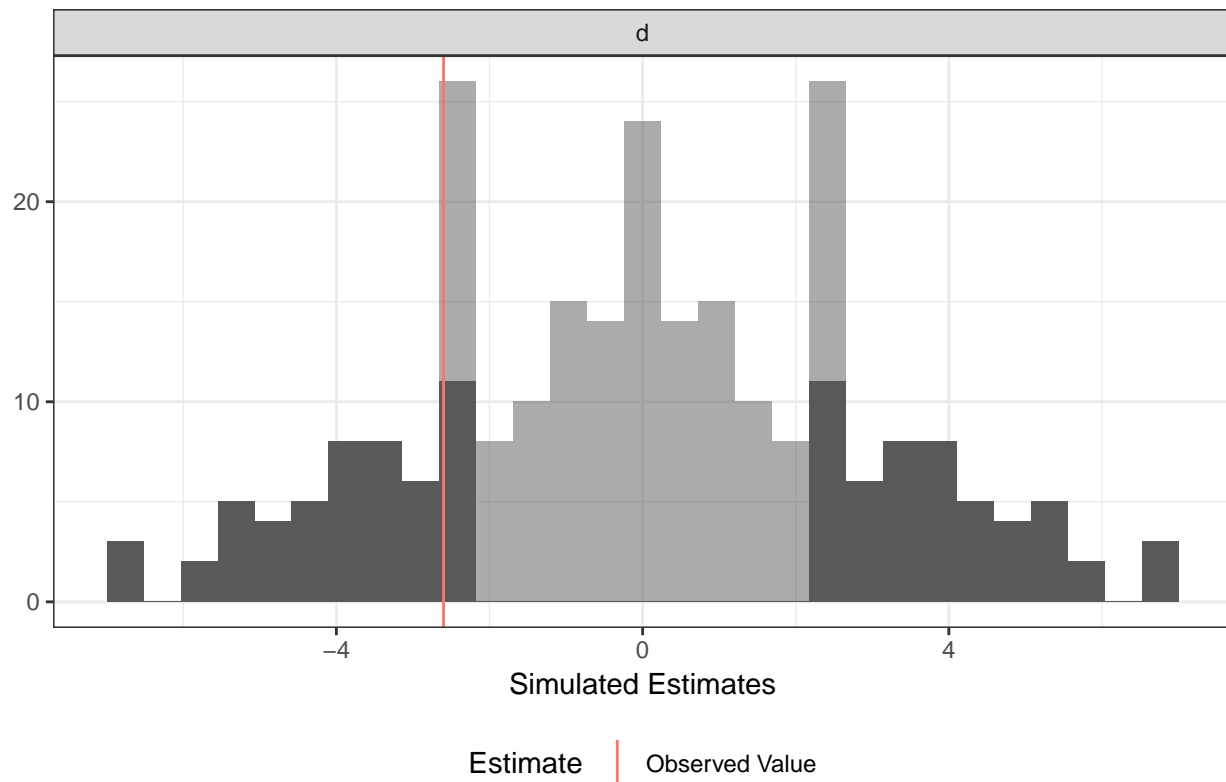
# Conduct Randomization Inference
ri2_out <- conduct_ri(y ~ d,
                      declaration = declaration,
                      assignment = "d",
                      sharp_hypothesis = 0,
                      data = data3.7)

summary(ri2_out)

##      coefficient estimate two_tailed_p_value null_ci_lower null_ci_upper
## 1             d      -2.6             0.4126984          -5.4           5.4

plot(ri2_out)
```

## Randomization Inference



## In ri

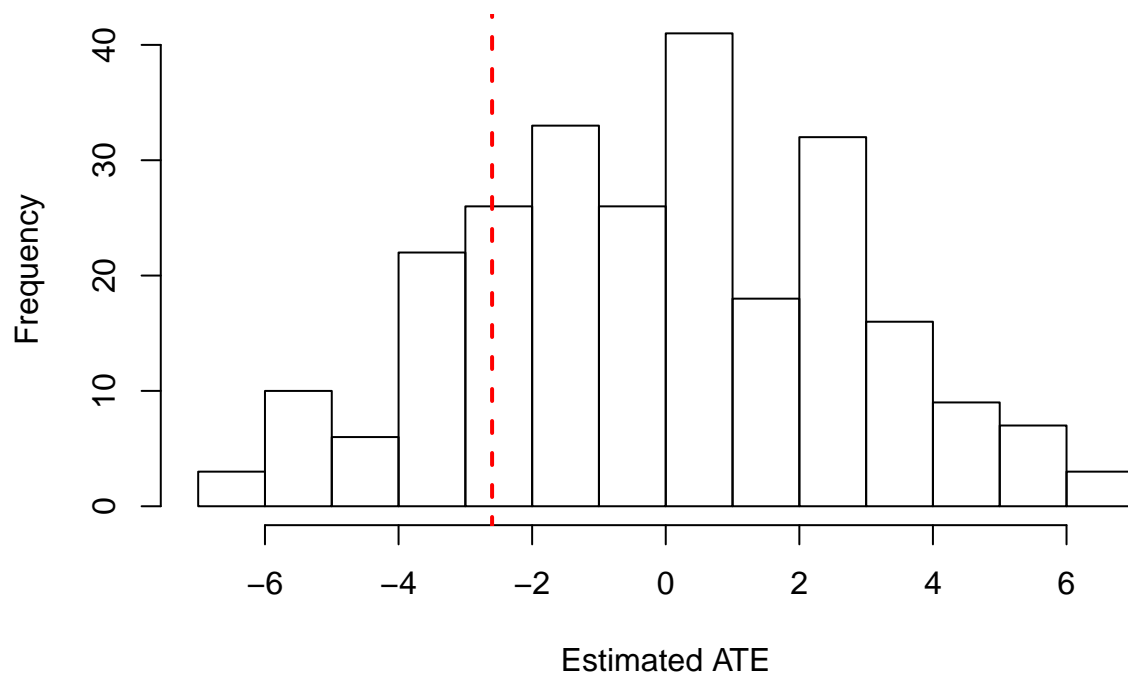
```
library(ri)

# all possible permutations
perms <- genperms(data3.7$d)
# probability of treatment
probs <- genprobexact(data3.7$d)
# estimate the ATE
ate <- estate(data3.7$y, data3.7$d, prob = probs)

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

# generate potential outcomes under sharp null of no effect
Ys <- genouts(data3.7$y, data3.7$d, 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)
```

## Distribution of the Estimated ATE



```
ri_out
```

```
## $two.tailed.p.value
## [1] 0.4126984
##
## $two.tailed.p.value.abs
## [1] 0.4126984
##
## $greater.p.value
## [1] 0.8373016
##
## $lesser.p.value
## [1] 0.2063492
##
## $quantile
## 2.5% 97.5%
## -5.4 5.4
##
## $sd
## [1] 2.829213
##
## $exp.val
## [1] 0
```

## By hand

```
perms <- combn(10, 5, tabulate, nbins = 10)
sims <- ncol(perms)
```

```

observed_ate <- with(data3.7, mean(y[d == 1]) - mean(y[d == 0]))
simulated_ates <- rep(NA, sims)

for (i in 1:sims){
  data3.7$Z_sim <- perms[,i]
  simulated_ates[i] <- with(data3.7, mean(y[Z_sim == 1]) - mean(y[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] -2.600000  0.4126984  0.8373016  0.2063492
hist(simulated_ates, breaks = 10)
abline(v = observed_ate, col = "red")

```

**Histogram of simulated\_ates**

