

# Gerber and Green (2012) Chapter 3 Problem 6

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

## Chapter 3 Problem 6

The Clingingsmith, Khwaja, and Kremer study discussed in section 3.5 may be used to test the sharp null hypothesis that winning the visa lottery for the pilgrimage to Mecca had no effect on the views of Pakistani Muslims toward people from other countries. Assume that the Pakistani authorities assigned visas using complete random assignment. Conduct 10,000 simulated random assignments under the sharp null hypothesis. How many of the simulated random assignments generate an estimated ATE that is at least as large as the actual estimate of the ATE? What is the implied one-tailed p-value? How many of the simulated random assignments generate an estimated ATE that is at least as large in absolute value as the actual estimate of the ATE? What is the implied two-tailed p-value?

```
# Data from http://isps.yale.edu/FEDAI
data3.6 <- read.csv("datasets/3.6.csv")
# Number of sims the same for all three methods
sims <- 1000
```

### In ri2

```
library(ri2)

# Declare randomization procedure
declaration <- declare_ra(N = 958, m = 479)

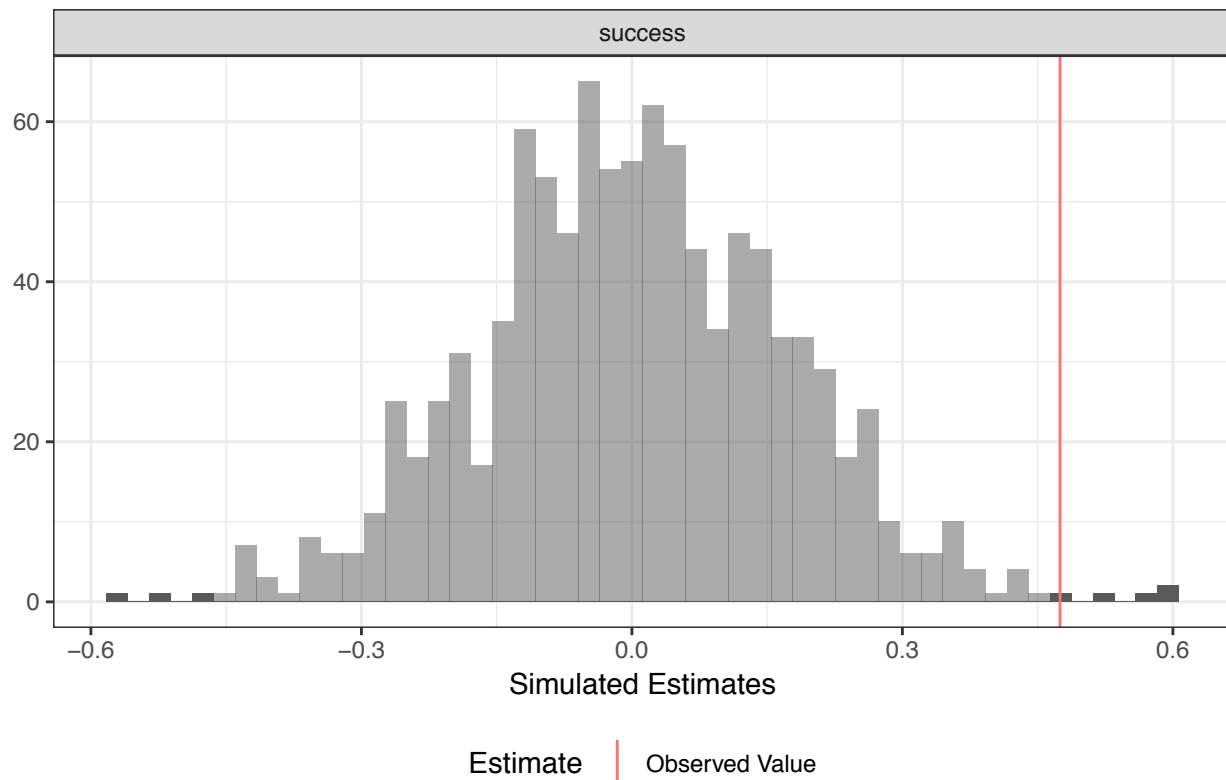
# Conduct Randomization Inference
ri2_out <- conduct_ri(
  formula = views ~ success,
  declaration = declaration,
  assignment = "success",
  sharp_hypothesis = 0,
  data = data3.6,
  sims = sims
)

summary(ri2_out)

##   coefficient estimate two_tailed_p_value null_ci_lower null_ci_upper
## 1      success 0.4748337             0.008      -0.3299582      0.3424843

plot(ri2_out)
```

## Randomization Inference



## In ri

```
library(ri)

# all possible permutations
perms <- genperms(data3.6$success, maxiter = sims)

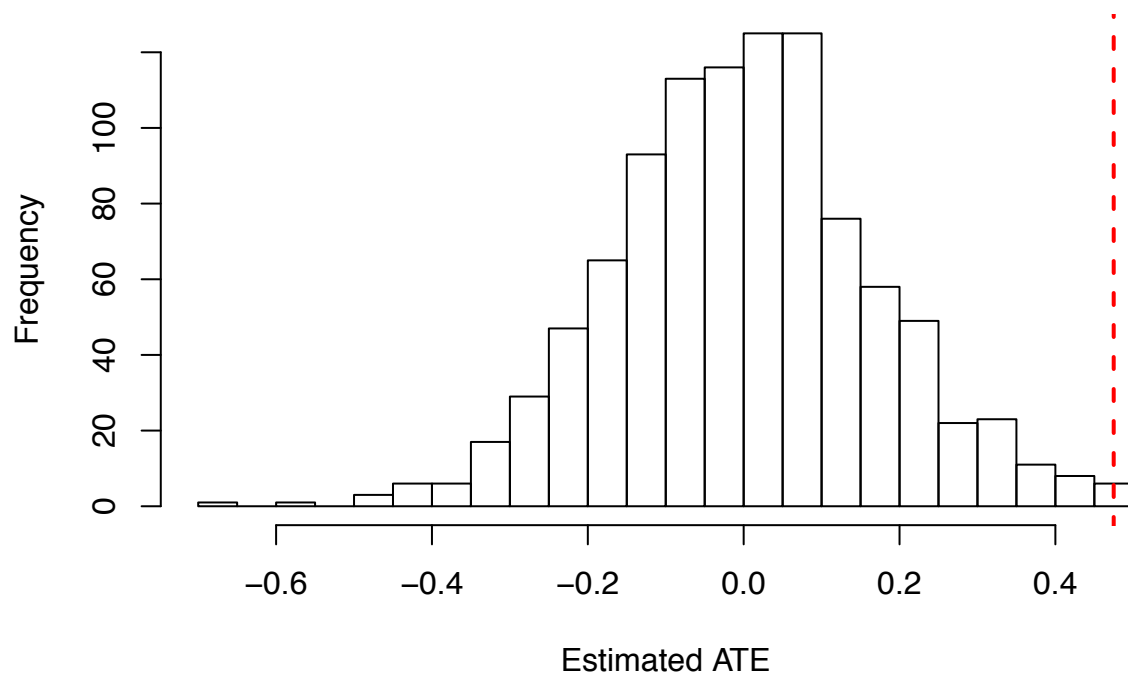
## Too many permutations to use exact method.
## Defaulting to approximate method.
## Increase maxiter to at least 8.45030476380969e+285 to perform exact estimation.

# probability of treatment
probs <- genprobexact(data3.6$success)
# estimate the ATE
ate <- estate(data3.6$views, data3.6$success, prob = probs)

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

# generate potential outcomes under sharp null of no effect
Ys <- genouts(data3.6$views, data3.6$success, 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] 0.006
##
## $two.tailed.p.value.abs
## [1] 0.006
##
## $greater.p.value
## [1] 0.003
##
## $lesser.p.value
## [1] 0.997
##
## $quantile
##      2.5%      97.5%
## -0.3303132  0.3491507
##
## $sd
## [1] 0.1684416
##
## $exp.val
## [1] 0.0004039741
```

## By hand

```
observed_ate <- with(data3.6, mean(views[success == 1]) - mean(views[success == 0]))
simulated_ates <- rep(NA, sims)
```

```

for (i in 1:sims) {
  Z_sim <- complete_ra(N = 958)
  simulated_ates[i] <-
    with(data3.6, mean(views[Z_sim == 1]) - mean(views[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] 0.4748337 0.0050000 0.0020000 0.9980000
hist(simulated_ates, breaks = 10)
abline(v = observed_ate, col = "red")

```

**Histogram of simulated\_ates**

