ECON312 Problem Set 1: question 2e

Futing Chen, Will Parker

04/12/2020

Contents


```
library(tidyverse)
library(knitr)
```

Roy model simulator function

```
roy_model <- function(sigma, rho, N = 10000, means = c(0,0), seed = 12346){
  set.seed(seed)
  cov_matrix <- matrix(c(sigma^2, rho*sigma, rho*sigma, 1), nrow = 2)

  sample <- mvtnorm::rmvnorm(n = N, means, cov_matrix)

  colnames(sample) <- c("U1", "U0")

  sample <- as_tibble(sample) %>%
    mutate(D = ifelse(U1 > U0, 1, 0),
           beta = U1-U0,
           Y = D*U1 + (1-D)*U0)

  ATE <- sample$beta %>% mean()

  ATT <- filter(sample, D ==1)$beta %>% mean()

  ATUT <- filter(sample, D == 0)$beta %>% mean()

  # beta OLS
  model <- lm(Y ~ D, data = sample)

  beta_OLS <- model$coefficients[2]

  E_Y_D1 = filter(sample, D ==1)$Y %>% mean()

  E_Y_D0 = filter(sample, D ==0)$Y %>% mean()

  E_diff = E_Y_D1 - E_Y_D0


  results <- tibble(
    `quantity of interest` = c("ATE",
                               "ATT",
                               "ATUT",
                               "Beta_OLS",
                               "E[Y|D =1] - E[Y|D =0]"),
    estimate = c(ATE, ATT, ATUT, beta_OLS, E_diff),
    `theoretical result` = c(0,
                           2*sqrt(sigma^2 + 1 -2*rho*sigma)*dnorm(0),
                           -2*sqrt(sigma^2 + 1 -2*rho*sigma)*dnorm(0),
                           2*dnorm(0)*(sigma^2 - 1)/sqrt(sigma^2 + 1 -2*rho*sigma),
                           2*dnorm(0)*(sigma^2 - 1)/sqrt(sigma^2 + 1 -2*rho*sigma))
  ) %>%
    mutate(estimate = round(estimate, 2),
           `theoretical result` = round(`theoretical result`,2))

  kable(results)
}
```

```r
plot_treatment_effects <- function(sigma, rho, N = 10000, means = c(0,0), seed = 12346){
  set.seed(seed)
  cov_matrix <- matrix(c(sigma^2, rho*sigma, rho*sigma, 1), nrow = 2)

  sample <- mvtnorm::rmvnorm(n = N, means, cov_matrix)

  colnames(sample) <- c("U1", "U0")

  sample <- as_tibble(sample) %>%
    mutate(D = ifelse(U1 > U0, 1, 0),
           beta = U1-U0,
           Y = D*U1 + (1-D)*U0)

  sample %>%
    ggplot(aes(x = beta, fill = factor(D))) +
    geom_histogram(binwidth = 1, center = 0) +
    ggtitle(paste0("sigma =", sigma, ", rho=", rho)) +
    labs(x = "U1 - U0", fill = "treatment selection")
}
```

Simulation results

```r
roy_model(2, 0.5)
```

quantity of interest	estimate	theoretical result
ATE	0.02	0.00
ATT	1.36	1.38
ATUT	-1.40	-1.38
Beta_OLS	1.36	1.38
E[Y|D =1] - E[Y|D =0]	1.36	1.38

The simulation shows that

$$E[Y|D=1] - E[Y|D=0] = \beta_{OLS}$$

```r
roy_model(2, 0)
```

quantity of interest	estimate	theoretical result
ATE	0.02	0.00
ATT	1.75	1.78
ATUT	-1.81	-1.78
Beta_OLS	1.03	1.07
E[Y|D =1] - E[Y|D =0]	1.03	1.07

```r
roy_model(2, -0.5)
```

quantity of interest	estimate	theoretical result
ATE	0.02	0.00
ATT	2.07	2.11
ATUT	-2.14	-2.11
Beta_OLS	0.86	0.90
E[Y|D =1] - E[Y|D =0]	0.86	0.90

Fixing $\rho = 0.5$ and varying σ

```r
roy_model(1, 0.5)
```

quantity of interest	estimate	theoretical result
ATE	0.01	0.0
ATT	0.79	0.8
ATUT	-0.80	-0.8
Beta_OLS	-0.01	0.0
E[Y|D =1] - E[Y|D =0]	-0.01	0.0

```r
roy_model(2, 0.5)
```

quantity of interest	estimate	theoretical result
ATE	0.02	0.00
ATT	1.36	1.38
ATUT	-1.40	-1.38
Beta_OLS	1.36	1.38
E[Y|D =1] - E[Y|D =0]	1.36	1.38

```r
roy_model(4, 0.5)
```

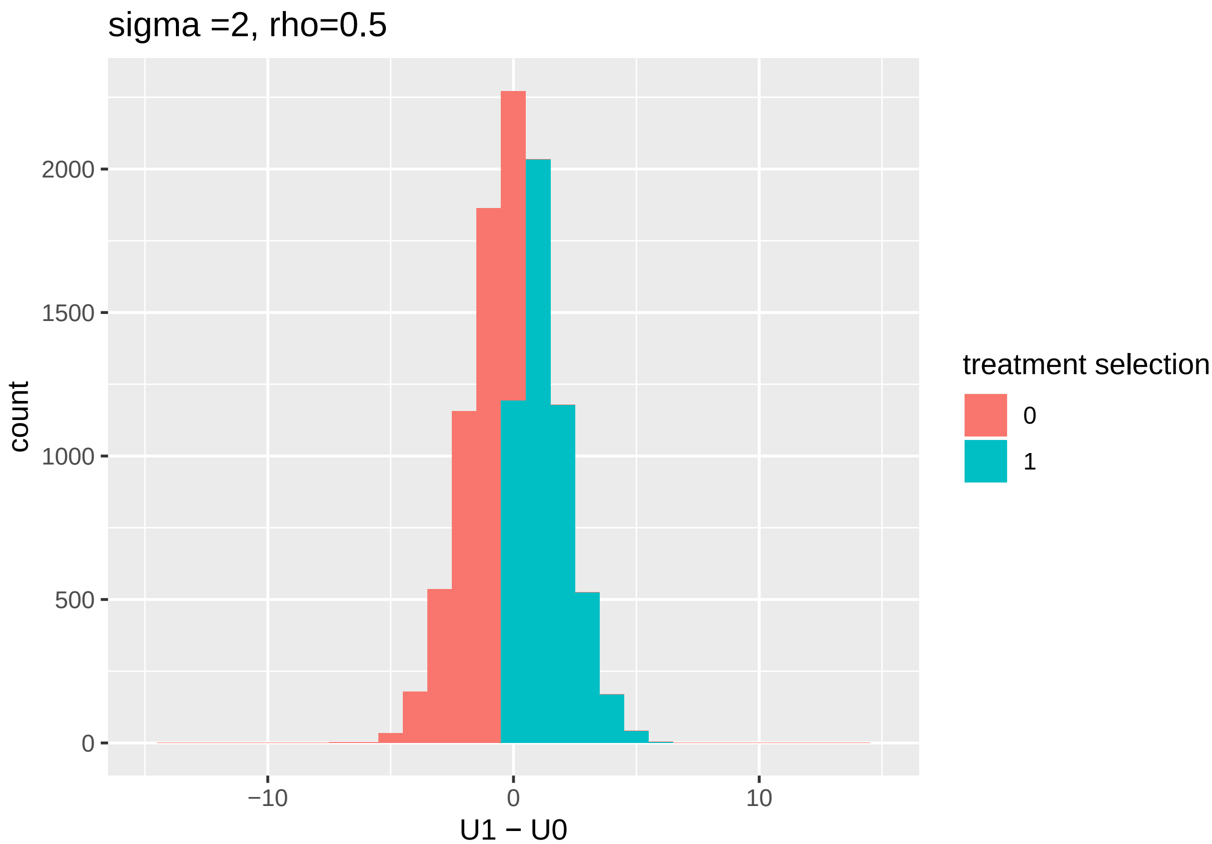
quantity of interest	estimate	theoretical result
ATE	0.05	0.00
ATT	2.84	2.88
ATUT	-2.89	-2.88
Beta_OLS	3.26	3.32
E[Y|D =1] - E[Y|D =0]	3.26	3.32

```r
roy_model(10, 0.5)
```

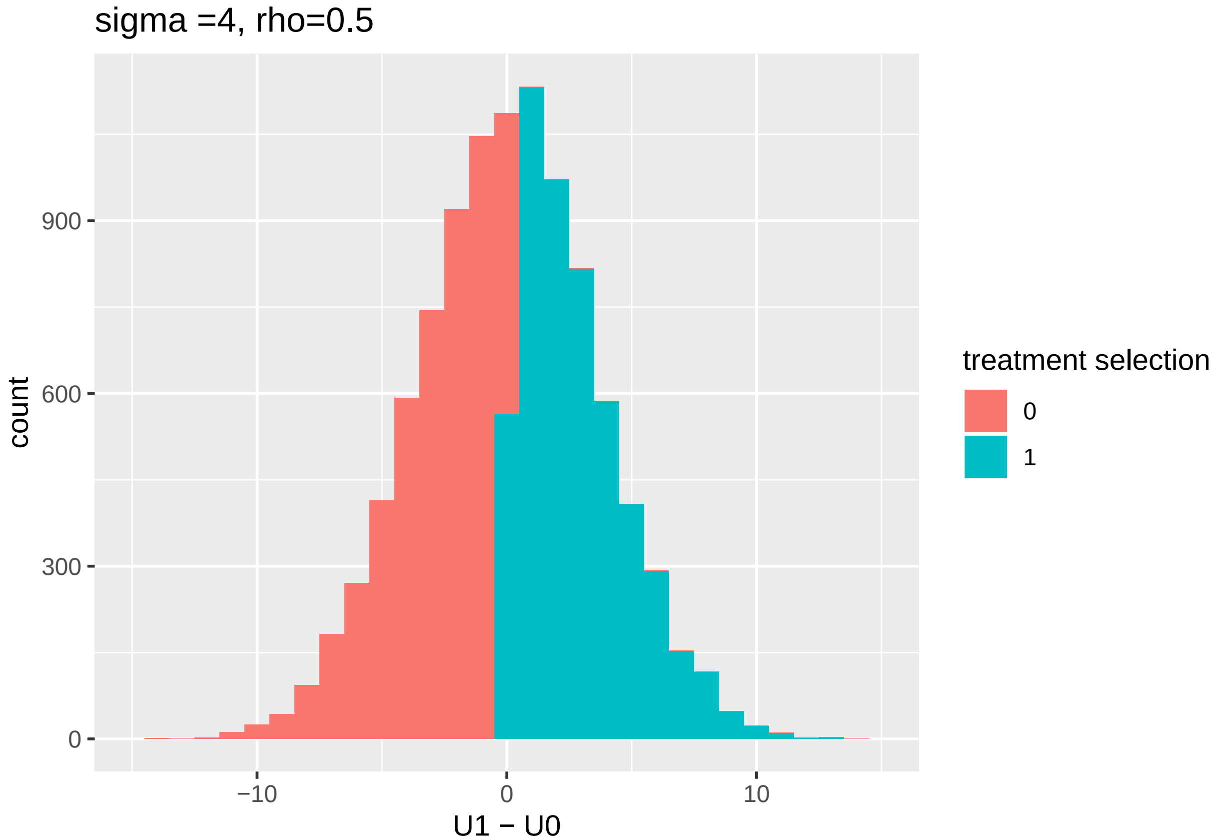
quantity of interest	estimate	theoretical result
ATE	0.13	0.00
ATT	7.53	7.61
ATUT	-7.61	-7.61
Beta_OLS	8.18	8.28
E[Y|D =1] - E[Y|D =0]	8.18	8.28

The increase in magnitude of ATT and ATUT as sigma increases can be explained by a larger variance in $U_1 - U_0$ and the $D = \mathbf{1}[U_1 > U_0]$ selection mechanism, as illustrated by these plots

```r
plot_treatment_effects(2, 0.5) +
      lims(x = c(-15, 15))
```




```r
plot_treatment_effects(4, 0.5) +
      lims(x = c(-15, 15))
```


sigma =4, rho=0.5
count
900
600
300
0
−10
0
10
U1 − U0
treatment selection
0
1