simulation

2023-03-13

```
rm(list = ls())

library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyve
rse 2.0.0 ──
## ✔ dplyr     1.1.0     ✔ readr     2.1.4
## ✔ forcats   1.0.0     ✔ stringr   1.5.0
## ✔ ggplot2   3.4.1     ✔ tibble    3.1.8
## ✔ lubridate 1.9.2     ✔ tidyr     1.3.0
## ✔ purrr     1.0.1
## ── Conflicts ────────────────────────────────────────── tidyverse_co
nflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the ]8;;http://conflicted.r-lib.org/conflicted package]8;; to
force all conflicts to become errors

library(MASS)

##
## 載入套件：'MASS'
##
## 下列物件被遮斷自 'package:dplyr':
##
##     select

set.seed(0)
#1 set of parameters
m0 <- 1
m1 <- 2
sd0 <- 1
sd1 <- 2
sd01 <- 1.5
C <- 3
#2 Simulate the (ϵ0, ϵ1) for N equals to 10 million individuals
N <- 10 ^ 7
sig <-rbind (c( sd0 ^2 , sd01 ) ,c( sd01 , sd1 ^2) )
mu <- c(0, 0)
data <- as.data.frame(mvrnorm(n = N, mu = mu, Sigma = sig))
colnames(data) <- c("e0", "e1")
#3, 4 Create the columns for w0 and w1,and enerate the column I that ta
ke binary value.
```

```r
data <- data %>%
  mutate(w0 = m0 + e0,
         w1 = m1 + e1,
         I = ifelse(w1 - w0 - C > 0, 1, 0))
#5 Calculate E[w0 | I], E[w1 | I], Q0, Q1
data_I <- data %>% filter(., I ==1)
sapply(data_I, FUN = mean)
```

```
##       e0       e1       w0       w1        I
## 0.660948 3.299858 1.660948 5.299858 1.000000
```

```r
#6 Calculate RHS of equation
sd_v <- sqrt(sd0 ^ 2 + sd1 ^ 2 - 2 * sd01)
rho <- sd01 / (sd0 * sd1)
z <- (m0 - m1 + C)/ sd_v
E_w0_I <- m0 + sd0 * sd1 / sd_v * (rho - sd0 / sd1) * (dnorm(z) / (1 -
pnorm(z)))
E_w1_I <- m0 + sd0 * sd1 / sd_v * (sd1 / sd0 - rho) * (dnorm(z) / (1 -
pnorm(z)))
Q_0 <- E_w0_I - m0
Q_1 <- E_w1_I - m1
E_w0_I
```

```
## [1] 1.659742
```

```r
E_w1_I
```

```
## [1] 4.298709
```

```r
Q_0
```

```
## [1] 0.6597419
```

```r
Q_1
```

```
## [1] 2.298709
```