ECON5521: Homework 1 Solution

Ting-Chih Hung
r12323020@ntu.edu.tw

March 13, 2024

4 Roy Model

Consider the simple model

$$\begin{aligned} w_0 &= \mu_0 + \varepsilon_0, \\ w_1 &= \mu_1 + \varepsilon_1, \end{aligned}$$

where migrant is 1, non-migrant is 0. Assume that the migration cost is C and the error terms are jointly normally distributed;[1]

$$\begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix} \right).$$

Let I be the event that an individual migrates. One will migrate if the wage gain is greater than the migration cost, $w_1 - w_0 > C$. Let $\nu := \varepsilon_1 - \varepsilon_0$ and $z := \frac{\mu_0 - \mu_1 + C}{\nu}$.

4.1 Review

1. Show that

$$\mathrm{E}\left[w_0 \mid I\right] = \mu_0 + \frac{\sigma_0 \sigma_1}{\sigma_\nu} \left(\rho - \frac{\sigma_0}{\sigma_1} \right) \left(\frac{\phi(z)}{1 - \Phi(z)} \right), \tag{1}$$

$$\mathrm{E}\left[w_1 \mid I\right] = \mu_0 + \frac{\sigma_0 \sigma_1}{\sigma_\nu} \left(\frac{\sigma_1}{\sigma_0} - \rho \right) \left(\frac{\phi(z)}{1 - \Phi(z)} \right). \tag{2}$$

> *Solution.* For those who migrate, their average wage that they would have earned if they had not migrated is
>
> $$\begin{aligned} & \mathrm{E}\left[w_0 \mid I\right] \\ = & \mathrm{E}\left[\mu_0 + \varepsilon_0 \mid I\right] && \text{(Def. of potential wage)} \end{aligned}$$

[1] The joint normality assumption is necessary for the following derivation. Note that even if the error terms are marginally normally distributed, they may not be jointly normally distributed.

$$
\begin{aligned}
&= \mu_0 + \mathrm{E}\left[\varepsilon_0 \mid I\right] \\
&= \mu_0 + \mathrm{E}\left[\varepsilon_0 \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right] && \left(I = \left\{\tfrac{\nu}{\sigma_\nu} > z\right\}\right) \\
&= \mu_0 + \sigma_0 \mathrm{E}\left[\frac{\varepsilon_0}{\sigma_0} \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right] && \text{(Linearity of expectation)} \\
&= \mu_0 + \sigma_0 \mathrm{E}\left\{\mathrm{E}\left[\frac{\varepsilon_0}{\sigma_0} \,\middle|\, \frac{\nu}{\sigma_\nu}\right] \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right\} && \text{(Iterated expectation)} \\
&= \mu_0 + \sigma_0 \mathrm{E}\left\{\frac{\mathrm{Cov}\left(\frac{\varepsilon_0}{\sigma_0}, \frac{\nu}{\sigma_\nu}\right)}{\mathrm{Var}\left(\frac{\nu}{\sigma_\nu}\right)} \frac{\nu}{\sigma_\nu} \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right\} && \text{(Regression formula)} \\
&= \mu_0 + \rho_{0\nu}\sigma_0 \mathrm{E}\left\{\frac{\nu}{\sigma_\nu} \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right\} && \left(\rho_{0\nu} = \mathrm{Cov}\left(\tfrac{\varepsilon_0}{\sigma_0}, \tfrac{\nu}{\sigma_\nu}\right)\right) \\
&= \mu_0 + \rho_{0\nu}\sigma_0 \frac{\phi(z)}{1-\Phi(z)} && \text{(Mean of truncated normal dist.)} \\
&= \mu_0 + \frac{\sigma_{0\nu}}{\sigma_0\sigma_\nu}\sigma_0 \frac{\phi(z)}{1-\Phi(z)} && \text{(Def. of correlation coef.)} \\
&= \mu_0 + \frac{\sigma_{01} - \sigma_0^2}{\sigma_\nu} \frac{\phi(z)}{1-\Phi(z)} && \left(\sigma_{0\nu} = \sigma_0^2 - \sigma_{01}\right) \\
&= \mu_0 + \frac{\sigma_0\sigma_1}{\sigma_\nu}\left(\rho - \frac{\sigma_0}{\sigma_1}\right)\left(\frac{\phi(z)}{1-\Phi(z)}\right).
\end{aligned}
$$

For those who migrate, their average wage is

$$
\begin{aligned}
&\mathrm{E}\left[w_1 \mid I\right] \\
&= \mathrm{E}\left[\mu_1 + \varepsilon_1 \mid I\right] && \text{(Def. of potential wage)} \\
&= \mu_1 + \mathrm{E}\left[\varepsilon_1 \mid I\right] \\
&= \mu_1 + \mathrm{E}\left[\varepsilon_1 \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right] && \left(I = \left\{\tfrac{\nu}{\sigma_\nu} > z\right\}\right) \\
&= \mu_1 + \sigma_1 \mathrm{E}\left[\frac{\varepsilon_1}{\sigma_1} \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right] && \text{(Linearity of expectation)} \\
&= \mu_1 + \sigma_1 \mathrm{E}\left\{\mathrm{E}\left[\frac{\varepsilon_1}{\sigma_1} \,\middle|\, \frac{\nu}{\sigma_\nu}\right] \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right\} && \text{(Iterated expectation)} \\
&= \mu_1 + \sigma_1 \mathrm{E}\left\{\frac{\mathrm{Cov}\left(\frac{\varepsilon_1}{\sigma_1}, \frac{\nu}{\sigma_\nu}\right)}{\mathrm{Var}\left(\frac{\nu}{\sigma_\nu}\right)} \frac{\nu}{\sigma_\nu} \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right\} && \text{(Regression formula)} \\
&= \mu_1 + \sigma_1 \mathrm{E}\left\{\mathrm{Cov}\left(\frac{\varepsilon_1}{\sigma_1}, \frac{\nu}{\sigma_\nu}\right) \frac{\nu}{\sigma_\nu} \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right\} && \left(\mathrm{Var}\left(\tfrac{\nu}{\sigma_\nu}\right) = 1\right) \\
&= \mu_1 + \rho_{1\nu}\sigma_1 \mathrm{E}\left\{\frac{\nu}{\sigma_\nu} \,\middle|\, \frac{\nu}{\sigma_\nu} > z\right\} && \left(\rho_{1\nu} = \mathrm{Cov}\left(\tfrac{\varepsilon_1}{\sigma_1}, \tfrac{\nu}{\sigma_\nu}\right)\right) \\
&= \mu_1 + \rho_{1\nu}\sigma_1 \frac{\phi(z)}{1-\Phi(z)} && \text{(Mean of truncated normal dist.)}
\end{aligned}
$$

$$
\begin{aligned}
&= \mu_1 + \frac{\sigma_{1\nu}}{\sigma_1 \sigma_\nu} \sigma_1 \frac{\phi(z)}{1-\Phi(z)} && \text{(Def. of correlation coef.)} \\
&= \mu_1 + \frac{\sigma_1^2 - \sigma_{01}}{\sigma_\nu} \frac{\phi(z)}{1-\Phi(z)} && (\sigma_{1\nu} = \sigma_1^2 - \sigma_{01}) \\
&= \mu_1 + \frac{\sigma_0 \sigma_1}{\sigma_\nu} \left(\frac{\sigma_1}{\sigma_0} - \rho \right) \frac{\phi(z)}{1-\Phi(z)}.
\end{aligned}
$$

2. Let $Q_0 = \mathrm{E}[\varepsilon_0 \mid I]$ and $Q_1 = \mathrm{E}[\varepsilon_1 \mid I]$. Is it possible that $Q_0 > 0$, $Q_1 < 0$?

Solution. By the previous result, we have

$$
\begin{aligned}
Q_0 &:= \mathrm{E}[\varepsilon_0 \mid I] = \frac{\sigma_0 \sigma_1}{\sigma_\nu} \left(\rho - \frac{\sigma_0}{\sigma_1} \right) \left(\frac{\phi(z)}{1-\Phi(z)} \right), \\
Q_1 &:= \mathrm{E}[\varepsilon_1 \mid I] = \frac{\sigma_0 \sigma_1}{\sigma_\nu} \left(\frac{\sigma_1}{\sigma_0} - \rho \right) \left(\frac{\phi(z)}{1-\Phi(z)} \right).
\end{aligned}
$$

This is impossible because it requires that $\rho > 1$.

4.2 Simulation

The R code (using `tidyverse`) is provided in the Appendix B.

1. Pick your favorite value for this set of parameters $(\mu_0, \mu_1, \sigma_0, \sigma_1, \sigma_{01}, C)$.

Solution. Feel free to pick any value you like. I set $\mu_0 = 0.8$, $\mu_1 = 1$, $\sigma_0 = 0.8$, $\sigma_1 = 0.4$, $\sigma_{01} = 0.25$, and $C = 0$.

2. Simulate the $(\varepsilon_0, \varepsilon_1)$ for N equals to 10 million individuals.

Solution. See the R code in the Appendix B.

3. Create the columns for w_0 and w_1.

Solution. See the R code in the Appendix B.

4. Generate the column I that take binary value.

Solution. See the R code in the Appendix B.

5. Calculate $\mathrm{E}[w_0 \mid I]$, $\mathrm{E}[w_1 \mid I]$, Q_0, Q_1 from data without invoking Equation (1) and (2).

> *Solution.* See the R code in the Appendix B. The simulation results are approximately 0.3857, 0.9049, −0.4143, and 0.0951, respectively.

6. Calculate RHS of Equation (1) and (2) to compare with the previous question.

> *Solution.* The RHS of Equation (1) and (2) are calculated in the R code in the Appendix B. The theoretical values are approximately 0.3863 and 0.9046, respectively. They are close to the simulation results.

7. Which columns are observed in the real world? Which of them are not?

> *Solution.* We can observe the choice of migration I. Also, for those who migrate, we can observe their wage w_1, and for those who do not migrate, we can observe their wage w_0. The error terms ε_0 and ε_1 are not observed. For those who do not migrate, we cannot observe their potential wage w_1, and for those who migrate, we cannot observe their potential wage w_0. The average potential wage μ_0 and μ_1 are also not observed.

A Useful Results

> Theorem A.1 (Density of Truncated Normal Distribution). Suppose that X is standard normally distributed. The probability of X falling in the interval (a, b) is
> $$P(a < X < b) = \Phi(b) - \Phi(a).$$
> The conditional density of X given that $X \in (a, b)$ is
> $$f(x \mid x \in (a, b)) = \frac{\phi(x)}{\Phi(b) - \Phi(a)},$$
> where $\phi(\cdot)$ is the standard normal density function and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

> Theorem A.2 (Mean of Truncated Normal Distribution). The mean of $X \sim N(0, 1)$ given that $X \in (a, b)$ is
> $$\mathrm{E}[X \mid a < X < b] = \frac{\phi(a) - \phi(b)}{\Phi(b) - \Phi(a)}.$$

THEOREM A.3 (Regession Formula). Suppose that X and Y are jointly normally distributed with zero mean. Then, the conditional expectation of Y given that $X = x$ is

$$\mathrm{E}\left[Y \mid X = x\right] = \frac{\mathrm{Cov}\left(X, Y\right)}{\mathrm{Var}\left(X\right)} x.$$

Proof. Here is a proof of more general case. Suppose that

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} \right).$$

Then, their joint density function is

$$\begin{aligned} f(x, y) = {} & \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \\ & \exp\left[\frac{-1}{2(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)\right]. \end{aligned}$$

The conditional density of Y given $X = x$ is

$$\begin{aligned} & f(y \mid x) \\ = {} & \frac{f(x, y)}{f(x)} \\ = {} & \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)\right]}{\frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left[\frac{-1}{2}\frac{(x-\mu_X)^2}{\sigma_X^2}\right]} \\ = {} & \frac{\exp\left[\frac{-1}{2(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right) + \frac{1}{2}\frac{(x-\mu_X^2)}{\sigma_X^2}\right]}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \\ = {} & \frac{\exp\left[\frac{-1}{2(1-\rho^2)\sigma_Y^2}\left(\frac{(x-\mu_X)^2\sigma_Y^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)\sigma_Y}{\sigma_X} + (y-\mu_Y)^2 - \frac{(1-\rho^2)(x-\mu_X)^2\sigma_Y^2}{\sigma_X^2}\right)\right]}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \\ = {} & \frac{\exp\left[\frac{-1}{2(1-\rho^2)\sigma_Y^2}\left(-2\rho\frac{(x-\mu_X)(y-\mu_Y)\sigma_Y}{\sigma_X} + (y-\mu_Y)^2 + \frac{\rho^2(x-\mu_X)^2\sigma_Y^2}{\sigma_X^2}\right)\right]}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \\ = {} & \frac{\exp\left[\frac{-1}{2(1-\rho^2)\sigma_Y^2}\left((y-\mu_Y) - \left(\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)^2\right)\right)^2\right]}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}. \end{aligned}$$

Hence,

$$Y \mid X = x \sim N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X), (1-\rho^2)\sigma_Y^2\right).$$

Therefore,

$$\mathrm{E}\left[Y \mid X = x\right] = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

$\square$

B R Code

```r
library(data.table)
library(MASS)

# Set a random seed to make the results reproducible
set.seed(1234)

# Set the parameters
N <- 1e6
mu_0 <- 0.8
mu_1 <- 1
sigma_00 <- 0.64
sigma_11 <- 0.16
sigma_01 <- 0.25

# Create a data.table
dt <- data.table(
  ID = 1:N,
  mu_0 = mu_0,
  mu_1 = mu_1
)

# Specify the error structure
mean_vector <- c(0, 0)
cov_matrix <- matrix(
  c(
    sigma_00, sigma_01,
    sigma_01, sigma_11
  ),
  nrow = 2, byrow = TRUE
)

# Generate the errors with MASS::mvrnorm()
epsilons <- as.data.table(mvrnorm(n = N, mu = mean_vector, Sigma = cov_matrix))

# Create the columns for the errors and the observed outcomes
dt[, c("epsilon_0", "epsilon_1") := epsilons]
dt[, c("w_0", "w_1") := .(mu_0 + epsilon_0, mu_1 + epsilon_1)]
dt[, I := fifelse(w_0 >= w_1, 0, 1)]
dt[, W := fifelse(I == 0, w_0, w_1)]

# Calculate E[w_0 | I = 1] and E[w_1 | I = 1]
mean(dt[I == 1, w_0])
mean(dt[I == 1, w_1])
```

```r
# Q_0 and Q_1
mean(dt[I == 1, w_0]) - mu_0
mean(dt[I == 1, w_1]) - mu_1

# Use the formula to calculate the E[w_0 | I = 1] and E[w_1 | I = 1]
sd_nu <- sqrt(sigma_00 + sigma_11 - 2 * sigma_01)
sd_0 <- sqrt(sigma_00)
sd_1 <- sqrt(sigma_11)
z <- (mu_0 - mu_1) / sd_nu
rho <- sigma_01 / sqrt(sigma_00 * sigma_11)
imr <- (dnorm(z) / (1 - pnorm(z)))
Q_0 <- (sd_0 * sd_1 / sd_nu) * (rho - sd_0 / sd_1) * imr
Q_1 <- (sd_0 * sd_1 / sd_nu) * (sd_1 / sd_0 - rho) * imr
mu_0 + Q_0
mu_1 + Q_1
```