

Homework 1

ECON 7023: Econometrics II

Maghfira Ramadhani

January 26, 2022

Spring 2023

Part 1

Problem 3.5

Let $\{y_i : i = 1, 2, 3, \dots\}$ be an independent, identically distributed sequence with $E(y_i^2) < \infty$. Let $\mu = E(y_i)$ and $\sigma^2 = \text{Var}(y_i)$.

- a. Let \bar{y}_N denote the sample average based on a sample size of N . Find $\text{Var}[\sqrt{N}(\bar{y}_N - \mu)]$.

Answer:

By definition of sample mean, we have $\text{Var}(\bar{y}_N) = \text{Var}\left(\frac{1}{N} \sum_{i=1}^N y_i\right)$. Recall property of variance that $\text{Var}(ax) = a^2 \text{Var}(x)$ for random variable x and constant a . Since N is constant, thus we have $\text{Var}(\bar{y}_N) = \left(\frac{1}{N}\right)^2 \text{Var}\left(\sum_{i=1}^N y_i\right)$. It is given that $y_i \stackrel{iid}{\sim} (\mu, \sigma^2)$ for $i \in \{1, 2, \dots, n\}$. Therefore $\text{Var}(y_i) = \sigma^2$ and is constant, and since y_i 's are independent we have $\text{Var}\left(\sum_{i=1}^N y_i\right) = \sum_{i=1}^N \text{Var}(y_i)$. Finally, combining these results, we have

$$\text{Var}(\bar{y}_N) = \left(\frac{1}{N}\right)^2 \sum_{i=1}^N \text{Var}(y_i) = \left(\frac{1}{N}\right)^2 N \sigma^2 = \sigma^2/N.$$

Now, we can calculate $\text{Var}[\sqrt{N}(\bar{y}_N - \mu)]$. Since N and μ are constant, we have

$$\text{Var}[\sqrt{N}(\bar{y}_N - \mu)] = \text{Var}[\sqrt{N}(\bar{y}_N)] = (\sqrt{N})^2 \text{Var}(\bar{y}_N) = N(\sigma^2/N) = \sigma^2.$$

□

- b. What is the asymptotic variance of $\sqrt{N}(\bar{y}_N - \mu)$?

Answer:

By Central Limit Theorem, since $y_i \stackrel{iid}{\sim} (\mu, \sigma^2)$, then we have $\sqrt{N}(\bar{y}_N - \mu) \stackrel{a}{\sim} \text{Normal}(0, \sigma^2)$. So, the asymptotic variance of $\sqrt{N}(\bar{y}_N - \mu)$ is $\text{Avar}[\sqrt{N}(\bar{y}_N - \mu)] = \sigma^2$.

□

- c. What is the asymptotic variance of \bar{y}_N ? Compare this with $\text{Var}(\bar{y}_N)$.

Answer:

From the result in (c) we have $\text{Avar}[\sqrt{N}(\bar{y}_N - \mu)] = \sigma^2$. Since N and μ are constant we have $\text{Avar}[\sqrt{N}(\bar{y}_N - \mu)] = \text{Avar}[\sqrt{N}(\bar{y}_N)]$. Recall property of variance that $\text{Var}(ax) = a^2 \text{Var}(x)$ for random variable x and constant a . Applying this property to the asymptotic variance we have $\text{Avar}(\bar{y}_N) = \frac{1}{N} \text{Avar}[\sqrt{N}(\bar{y}_N)] = \frac{1}{N} \text{Avar}[\sqrt{N}(\bar{y}_N - \mu)] = \sigma^2/N$. We get the same result as in (a).

□

- d. What is the asymptotic standard deviation of \bar{y}_N ?

Answer:

By definition, the asymptotic standard deviation is the square root of the asymptotic variance that we have from previous result, formally, $\sqrt{\text{Avar}(\bar{y}_N)} = \sqrt{\sigma^2/N} = \sigma/\sqrt{N}$.

□

e. How would you obtain the asymptotic standard error of \bar{y}_N ?

Answer:

Applying **Definition 3.10** from the textbook, the asymptotic standard error of \bar{y}_N is $se(\bar{y}_N) = \hat{\sigma}/\sqrt{N}$, with $\hat{\sigma}$ a consistent estimator for σ , formally, $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$. Because we deal with y_i as a single random variable, the consistent estimator for σ^2 will be $\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y}_N)^2$. Then take the $|\sqrt{\hat{\sigma}^2}|$ as the $\hat{\sigma}$.

Proof. We need to show that $\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y}_N)^2$ is a consistent and unbiased estimator for σ^2 . We know that $y_i \stackrel{iid}{\sim} (\mu, \sigma^2)$. From the Weak Law of Large Number, we have $E(\bar{y}_N) = \mu$. From previous result, we already have the following properties $\text{Var}(\bar{y}_N) = \sigma^2/N$. Recall that $\text{Var}(\bar{y}_N) = E(\bar{y}_N^2) - E(\bar{y}_N)^2$, thus we have $E(\bar{y}_N^2) = \text{Var}(\bar{y}_N) + E(\bar{y}_N)^2 = \sigma^2/N + \mu^2$. We also know that $\mu = E(y_i)$ and $\sigma^2 = \text{Var}(y_i)$, thus we have $E(y_i^2) = \text{Var}(y_i) + E(y_i)^2 = \sigma^2 + \mu^2$. Now take expectation of our estimator

$$\begin{aligned} E(\hat{\sigma}^2) &= \frac{1}{N-1} E\left(\sum_{i=1}^N (y_i - \bar{y}_N)^2\right) \\ &= \frac{1}{N-1} E\left(\sum_{i=1}^N (y_i^2 - 2y_i\bar{y}_N + \bar{y}_N^2)\right) \\ &= \frac{1}{N-1} E\left(\sum_{i=1}^N y_i^2 - 2\bar{y}_N \sum_{i=1}^N y_i + \sum_{i=1}^N \bar{y}_N^2\right) \\ &= \frac{1}{N-1} E\left(\sum_{i=1}^N y_i^2 - 2N\bar{y}_N^2 + N\bar{y}_N^2\right) \\ &= \frac{1}{N-1} E\left(\sum_{i=1}^N y_i^2 - N\bar{y}_N^2\right) \\ &= \frac{1}{N-1} \left[E\left(\sum_{i=1}^N y_i^2\right) - NE(\bar{y}_N^2) \right] \\ &= \frac{1}{N-1} [N(\sigma^2 + \mu^2) - N(\sigma^2/N + \mu^2)] \\ &= \frac{1}{N-1} (N-1)\sigma^2 \\ &= \sigma^2. \end{aligned}$$

We show that the estimator is unbiased. Now we need to show consistency. By Weak Law of Large Number, we have $\bar{y}_N \xrightarrow{P} \mu$. Again, by Weak Law of Large Number we have $\frac{1}{N} \sum_{i=1}^N y_i^2 \xrightarrow{P} E(y_i^2)$. Recall our estimator

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y}_N)^2 = \frac{1}{N-1} \left(\sum_{i=1}^N y_i^2 - N\bar{y}_N^2 \right) = \frac{N}{N-1} \left(\frac{1}{N} \sum_{i=1}^N y_i^2 - \bar{y}_N^2 \right). \quad (1)$$

Take the *plim* of equation (1) we have

$$\begin{aligned} \text{plim} \hat{\sigma}^2 &= \text{plim} \left(\frac{N}{N-1} \left(\frac{1}{N} \sum_{i=1}^N y_i^2 - \bar{y}_N^2 \right) \right) \\ &= \lim_{N \rightarrow \infty} \left(\frac{N}{N-1} \right) \left(\text{plim} \left(\frac{1}{N} \sum_{i=1}^N y_i^2 \right) - \text{plim}(\bar{y}_N^2) \right) \quad [\text{by Slutsky's theorem}] \\ &= 1 \cdot (E(y_i^2) - E(y_i)^2) \\ &= \text{Var}(y_i) = \sigma^2. \end{aligned}$$

We showed that the estimator is consistent. □

Problem 4.1

Consider a standard $\log(\text{wage})$ equation for men under the assumption that all explanatory variables are exogenous:

$$\begin{aligned}\log(\text{wage}) &= \beta_0 + \beta_1 \text{married} + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma} + u, \\ E(u|\text{married}, \text{educ}, \mathbf{z}) &= 0,\end{aligned}\tag{4.49}$$

where \mathbf{z} contains factors other than marital status and education that can affect wages. When β_1 is small, $100 \cdot \beta_1$ is approximately the ceteris paribus difference in wages between married and unmarried men. When β_1 is large, it might be preferable to use the exact percentage difference in $E(\text{wage}|\text{married}, \text{educ}, \mathbf{z})$. Call this θ_1 .

- a. Show that if u is independent of all explanatory variables in equation (4.49), then $\theta_1 = 100 \cdot [\exp(\beta_1) - 1]$. (Hint: Find $E(\text{wage}|\text{married}, \text{educ}, \mathbf{z})$ for $\text{married} = 1$ and $\text{married} = 0$, and find the percentage difference). A natural, consistent, estimator of θ_1 is $\hat{\theta}_1 = 100 \cdot [\exp(\hat{\beta}_1) - 1]$, where $\hat{\beta}_1$ is the OLS estimator from equation (4.49).

Answer:

Rewrite equation (4.49),

$$\begin{aligned}\log(\text{wage}) &= \beta_0 + \beta_1 \text{married} + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma} + u \\ \Leftrightarrow \text{wage} &= \exp(\beta_0 + \beta_1 \text{married} + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma} + u) \\ \Leftrightarrow \text{wage} &= \exp(u) \exp[\beta_0 + \beta_1 \text{married} + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma}].\end{aligned}\tag{2}$$

Now take the expectation of (2) with respect to all explanatory variables, we have

$$E(\text{wage}|\text{married}, \text{educ}, \mathbf{z}) = E(\exp(u)|\text{married}, \text{educ}, \mathbf{z}) \exp(\beta_0 + \beta_1 \text{married} + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma}).\tag{3}$$

If u is independent of all explanatory variables, then $\exp(u)$ will also be independent with all those variables, i.e. $E(\exp(u)|\text{married}, \text{educ}, \mathbf{z}) = E(\exp(u)) = c$, where c is a constant. Now from (3) we have

$$E(\text{wage}|\text{married}, \text{educ}, \mathbf{z}) = c \exp(\beta_0 + \beta_1 \text{married} + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma}).\tag{4}$$

Then, we can compute the exact percentage different $\theta_1 = 100 \cdot \frac{E(\text{wage}|\text{married}=1) - E(\text{wage}|\text{married}=0)}{E(\text{wage}|\text{married}=0)}$ from (4), we have

$$\begin{aligned}\theta_1 &= 100 \cdot \left[\frac{c \exp(\beta_0 + \beta_1(1) + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma}) - c \exp(\beta_0 + \beta_1(0) + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma})}{c \exp(\beta_0 + \beta_1(0) + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma})} \right] \\ &= 100 \cdot \left[\frac{\exp(\beta_0 + \beta_1 + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma})}{\exp(\beta_0 + \beta_2 \text{educ} + \mathbf{z}\boldsymbol{\gamma})} - 1 \right] \\ &= 100 \cdot [\exp(\beta_1) - 1].\end{aligned}$$

□

- b. Use the delta method (see Section 3.5.2) to show that asymptotic standard error of $\hat{\theta}_1$ is $100 \cdot [\exp(\hat{\beta}_1)] \cdot \text{se}(\hat{\beta}_1)$.

Answer:

From previous result we have $\theta_1 : \beta_1 \rightarrow \mathbb{R}$ defined as $\theta_1 = 100 \cdot [\exp(\beta_1) - 1]$. We can use the OLS estimator $\hat{\beta}_1$ from (4.49) to estimate $\hat{\theta}_1 = 100 \cdot [\exp(\hat{\beta}_1) - 1]$. We are interested in finding $\text{se}(\hat{\theta}_1)$. By Delta Method we have

$$\text{Avar}[\hat{\theta}_1(\hat{\beta}_1)] = \left(\frac{\partial \hat{\theta}_1}{\partial \hat{\beta}_1} \right)^2 \text{Avar}(\hat{\beta}_1),\tag{5}$$

with the Jacobian in this case is only 1×1 , i.e., a scalar. From (5) we take the square root to find $se(\hat{\theta}_1)$, we have

$$\begin{aligned} se[\hat{\theta}_1(\hat{\beta}_1)] &= \left(\frac{\partial \hat{\theta}_1}{\partial \hat{\beta}_1} \right) se(\hat{\beta}_1) \\ &= 100 \cdot \exp(\hat{\beta}_1) \cdot se(\hat{\beta}_1). \end{aligned}$$

□

- c. Repeat parts a and b by finding the exact percentage change in $E(wage|married, educ, \mathbf{z})$ for any given change in $educ$, $\Delta educ$. Call this θ_2 . Explain how to estimate θ_2 and obtain its asymptotic standard error.

Answer:

The step involved is similar with the previous result. Recall (4), we have

$$E(wage|married, educ, \mathbf{z}) = c \exp(\beta_0 + \beta_1 married + \beta_2 educ + \mathbf{z}\gamma).$$

Then, we can compute the exact percentage different for an additional years of education $\Delta educ$, that is $\theta_2 = 100 \cdot \frac{E(wage|educ=\epsilon+\Delta educ) - E(wage|educ=\epsilon)}{E(wage|educ=\epsilon)}$ from (4), we have

$$\begin{aligned} \theta_2 &= 100 \cdot \left[\frac{c \exp(\beta_0 + \beta_1 married + \beta_2(\epsilon + \Delta educ) + \mathbf{z}\gamma) - c \exp(\beta_0 + \beta_1 married + \beta_2(\epsilon) + \mathbf{z}\gamma)}{c \exp(\beta_0 + \beta_1 married + \beta_2(\epsilon) + \mathbf{z}\gamma)} \right] \\ &= 100 \cdot \left[\frac{c \exp(\beta_0 + \beta_1 married + \beta_2(\epsilon + \Delta educ) + \mathbf{z}\gamma)}{c \exp(\beta_0 + \beta_1 married + \beta_2(\epsilon) + \mathbf{z}\gamma)} - 1 \right] \\ &= 100 \cdot [\exp(\beta_2 \Delta educ) - 1]. \end{aligned}$$

We can use the OLS estimator $\hat{\beta}_2$ from (4.49) to estimate $\hat{\theta}_2 = 100 \cdot [\exp(\beta_2) \Delta educ - 1]$ for a given change in $educ$, $\Delta educ$.

□

To estimate the asymptotic standard error we can use the Delta Method. We have that $\theta_2 : \beta_2 \rightarrow \mathbb{R}$ defined as $\theta_2 = 100 \cdot [\exp(\beta_2 \Delta educ) - 1]$. Here we treat $\Delta educ$ as a constant because it is given. By Delta Method we have

$$Avar[\hat{\theta}_2(\hat{\beta}_2)] = \left(\frac{\partial \hat{\theta}_2}{\partial \hat{\beta}_2} \right)^2 Avar(\hat{\beta}_2), \quad (6)$$

with the Jacobian in this case is only 1×1 , i.e., a scalar. From (6) we take the square root to find $se(\hat{\theta}_2)$, we have

$$\begin{aligned} se[\hat{\theta}_2(\hat{\beta}_2)] &= \left(\frac{\partial \hat{\theta}_2}{\partial \hat{\beta}_2} \right) se(\hat{\beta}_2) \\ &= 100 \cdot |\Delta educ| \cdot \exp(\hat{\beta}_2 \Delta educ) \cdot se(\hat{\beta}_2). \end{aligned}$$

□

- d. Use the data in NLS80.RAW to estimate equation (4.49), where \mathbf{z} contains the remaining variables in equation (4.29) (except ability, of course). Find $\hat{\theta}_1$ and its standard error; find $\hat{\theta}_2$ and its standard error when $\Delta educ = 4$.

Answer:

From the data we get these regression result in Table 1.

Table 1: Regression result for (4.49)

	$\log(wage)$
years of work experience	0.014*** (0.003)
years with current employer	0.012*** (0.003)
=1 if married	0.199*** (0.040)
=1 if live in south	-0.091*** (0.027)
=1 if live in SMSA	0.184*** (0.027)
=1 if black	-0.188*** (0.037)
years of education	0.065*** (0.006)
Constant	5.395*** (0.113)
Observations	935

Standard errors in parentheses

Data: NLS80.DTA

Wooldrige (2011)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Recall the previous result, we can calculate

$$\begin{aligned}\hat{\theta}_1 &= 100 \cdot [\exp(\hat{\beta}_1) - 1] = 100 \cdot [\exp(0.199) - 1] = 22.018\% \\ \text{se}[\hat{\theta}_1] &= 100 \cdot \exp(\hat{\beta}_1) \cdot \text{se}(\hat{\beta}_1) = 100 \cdot \exp(0.199) \cdot (0.040) = 4.881\% \\ \hat{\theta}_2 &= 100 \cdot [\exp(\hat{\beta}_2 \Delta \text{educ}) - 1] = 100 \cdot [\exp(0.065 \cdot 4) - 1] = 29.693\% \\ \text{se}[\hat{\theta}_2] &= 100 \cdot |\Delta \text{educ}| \cdot \exp(\hat{\beta}_2 \Delta \text{educ}) \cdot \text{se}(\hat{\beta}_2) = 100 \cdot 4 \cdot \exp(0.065 \cdot 4) \cdot (0.006) = 3.113\%.\end{aligned}$$

□

Problem 4.2

- a. Show that, under random sampling and the zero conditional mean assumption $E(u|\mathbf{x}) = 0$, $E(\hat{\beta}|\mathbf{X}) = \beta$ if $\mathbf{X}'\mathbf{X}$ is nonsingular.

Answer:

Recall our OLS estimator

$$\hat{\beta} = \left(N^{-1} \sum_{i=1}^N \mathbf{x}_i' \mathbf{x}_i \right)^{-1} \left(N^{-1} \sum_{i=1}^N \mathbf{x}_i' y_i \right) = \beta + \left(N^{-1} \sum_{i=1}^N \mathbf{x}_i' \mathbf{x}_i \right)^{-1} \left(N^{-1} \sum_{i=1}^N \mathbf{x}_i' u_i \right),$$

or in matrix notation

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y}) = \beta + (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{u}), \quad (7)$$

with \mathbf{X} a $N \times K$ data matrix of regressors with i th row \mathbf{x}_i , \mathbf{y} a $N \times 1$ data vector with i th element y_i , \mathbf{u} a $N \times 1$ matrix of errors. Now take the expectation of (7) with respect to \mathbf{X} , we have

$$E(\hat{\beta}|\mathbf{X}) = \beta + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' E(\mathbf{u}|\mathbf{X}) = \beta + (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{0} = \beta. \quad (8)$$

In order for equation (8) to hold, we need the zero conditional mean assumption, that is given, and we also need $(\mathbf{X}'\mathbf{X})$ to be nonsingular so that the inverse, $(\mathbf{X}'\mathbf{X})^{-1}$, exists.

□

- b. In addition to the assumption from part a, assume that $\text{Var}(u|\mathbf{x}) = \sigma^2$. Show that $\text{Var}(\hat{\beta}|\mathbf{X}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.

Answer:

Recall previous result from (a) in equation (7) we have

$$\hat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{u}).$$

In this β a vector of constants because it is a population property. Thus, if we take the conditional variance, we have

$$\begin{aligned}\text{Var}(\hat{\beta}|\mathbf{X}) &= \text{Var}((\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \mathbf{u}|\mathbf{X}) \\ &= [(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'] \text{Var}(u|\mathbf{X}) [(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}']' \\ &= [(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'] \text{Var}(u|\mathbf{X}) [\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] \end{aligned} \quad (9)$$

In equation (9) we need to get rid of $\text{Var}(\mathbf{u}|\mathbf{X})$, so recall that

$$\begin{aligned}\text{Var}(\mathbf{u}|\mathbf{X}) &= E[\mathbf{u}'\mathbf{u}|\mathbf{X}] - E[\mathbf{u}|\mathbf{X}]^2 \\ &= E[\mathbf{u}'\mathbf{u}|\mathbf{X}],\end{aligned}$$

which is a variance covariance matrix, since it is given in (a) that $E[\mathbf{u}|\mathbf{X}] = 0$. We are also given $\text{Var}(u_i|\mathbf{x}_i) = E(u_i^2|\mathbf{x}_i) = \sigma^2$. Note that since \mathbf{x}_i is iid across i and that the matrix is symmetric, we have

$$E(u_i^2|\mathbf{x}_i) = E(u_i^2|\mathbf{x}_i, \mathbf{x}_j) = E(u_i^2|\mathbf{X}) = \sigma^2,$$

where u_i^2 is a diagonal of variance covariance matrix. To transform the matrix quadratic form in (9) to scalar multiplication, we need the covariance part of the matrix to be zero, i.e., $E(u_i u_j | \mathbf{X}) = E(u_i u_j | \mathbf{x}_i, \mathbf{x}_j) = 0$. Note that since u_i and \mathbf{x}_i are iid across i , by Law of Iterated Expectation we have

$$\begin{aligned} E(u_i u_j | \mathbf{x}_i, \mathbf{x}_j) &= E(u_i u_j | \mathbf{x}_i, \mathbf{x}_j, u_j) && [\text{since } u_i, u_j \text{ are independent}] \\ &= u_j E(u_i | \mathbf{x}_i, \mathbf{x}_j, u_j) \\ &= u_j E(u_i | \mathbf{x}_i, \mathbf{x}_j) \\ &= u_j \cdot 0 && [\text{from given assumption in point (a)}] \\ &= 0. \end{aligned}$$

Now we have that $\text{Var}(\mathbf{u} | \mathbf{X}) = E[\mathbf{u}' \mathbf{u} | \mathbf{X}]$ is a diagonal matrix with the diagonal elements σ^2 . Thus from (9) the matrix inside the sandwich form can be transformed into scalar multiplication. Finally, we have

$$\begin{aligned} \text{Var}(\hat{\beta} | \mathbf{X}) &= [(\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \text{Var}(\mathbf{u} | \mathbf{X}) \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1}] \\ &= \sigma^2 [(\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1}] \\ &= \sigma^2 [(\mathbf{X}' \mathbf{X})^{-1} \mathbf{I}] \\ &= \sigma^2 (\mathbf{X}' \mathbf{X})^{-1}. \end{aligned}$$

□

Problem 4.4

Show that estimator $\hat{\mathbf{B}} \equiv N^{-1} \sum_{i=1}^N \hat{u}_i^2 \mathbf{x}_i' \mathbf{x}_i$ is consistent for $\mathbf{B} = E(u^2 \mathbf{x}' \mathbf{x})$ by showing that

$$N^{-1} \sum_{i=1}^N \hat{u}_i^2 \mathbf{x}_i' \mathbf{x}_i = N^{-1} \sum_{i=1}^N u_i^2 \mathbf{x}_i' \mathbf{x}_i + o_p(1).$$

(Hint: Write $\hat{u}_i^2 = u_i^2 - 2\mathbf{x}_i u_i (\hat{\beta} - \beta) + [\mathbf{x}_i (\hat{\beta} - \beta)]^2$, and use the facts that sample averages are $O_p(1)$ when expectations exist and that $\hat{\beta} - \beta = o_p(1)$. Assume that all necessary expectations exist and are finite.)

Answer:

Rewrite $\hat{\mathbf{B}}$ as follow

$$\begin{aligned} \hat{\mathbf{B}} &= N^{-1} \sum_{i=1}^N \hat{u}_i^2 \mathbf{x}_i' \mathbf{x}_i \\ &= N^{-1} \sum_{i=1}^N \left[u_i^2 - 2\mathbf{x}_i u_i (\hat{\beta} - \beta) + [\mathbf{x}_i (\hat{\beta} - \beta)]^2 \right] \mathbf{x}_i' \mathbf{x}_i \\ &= N^{-1} \sum_{i=1}^N u_i^2 \mathbf{x}_i' \mathbf{x}_i - 2N^{-1} \sum_{i=1}^N \mathbf{x}_i u_i (\hat{\beta} - \beta) \mathbf{x}_i' \mathbf{x}_i + N^{-1} \sum_{i=1}^N [\mathbf{x}_i (\hat{\beta} - \beta)]^2 \mathbf{x}_i' \mathbf{x}_i. \end{aligned}$$

Denote the first, second, and last term respectively as P, Q, R . Rewrite the second term as follow

$$\begin{aligned} Q &= -2N^{-1} \sum_{i=1}^N \mathbf{x}_i u_i (\hat{\beta} - \beta) \mathbf{x}_i' \mathbf{x}_i \\ &= -2 \left[\sum_{k=1}^N N^{-1} \left(\sum_{i=1}^N x_{ik} u_i (\hat{\beta}_k - \beta_k) \mathbf{x}_i' \mathbf{x}_i \right) \right] \\ &= -2 \left[\sum_{k=1}^N (\hat{\beta}_k - \beta_k) N^{-1} \left(\sum_{i=1}^N x_{ik} u_i \mathbf{x}_i' \mathbf{x}_i \right) \right]. \end{aligned}$$

Note that $\hat{\beta}_j - \beta_j = o_p(1)$, and $N^{-1} \left(\sum_{i=1}^N x_{ij} u_i \mathbf{x}'_i \right) = O_p(1)$. Therefore, we have the second term $Q = o_p(1) \cdot O_p(1) = o_p(1)$. Then, rewrite the third term as follow

$$\begin{aligned} R &= N^{-1} \sum_{i=1}^N [\mathbf{x}_i(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]^2 \mathbf{x}'_i \mathbf{x}_i \\ &= \sum_{k=1}^N \sum_{j=1}^N \left[(\hat{\beta}_k - \beta_k)(\hat{\beta}_j - \beta_j) N^{-1} \left(\sum_{i=1}^N [x_{ij} x_{ik}] \mathbf{x}'_i \mathbf{x}_i \right) \right]. \end{aligned}$$

Similarly, note that $\hat{\beta}_j - \beta_j = o_p(1)$, and $N^{-1} \left(\sum_{i=1}^N [x_{ij} x_{ik}] \mathbf{x}'_i \mathbf{x}_i \right) = O_p(1)$. Therefore, we have the third term $R = o_p(1) \cdot o_p(1) \cdot O_p(1) = o_p(1)$. Finally, we have

$$\begin{aligned} \hat{\mathbf{B}} &= P + Q + R \\ &= P + o_p(1) + o_p(1) \\ &= N^{-1} \sum_{i=1}^N u_i^2 \mathbf{x}'_i \mathbf{x}_i + o_p(1). \end{aligned}$$

□

Problem 4.9

Consider a linear model where the dependent variable is in logarithmic form, and the lag of $\log(y)$ is also an explanatory variable:

$$\log(y) = \beta_0 + \mathbf{x}\boldsymbol{\beta} + \alpha_1 \log(y_{-1}) + u, \quad E(u|\mathbf{x}, y_{-1}) = 0,$$

where the inclusion of y_{-1} might be to control for correlation between policy variables in \mathbf{x} and a previous value of y ; see Example 4.4.

- a. For estimating $\boldsymbol{\beta}$, why do we obtain the same estimator if the *growth* in y , $\log(y) - \log(y_{-1})$, is used instead as the dependent variable?

Answer:

Rearrange the equation by defining $\Delta \log(y) = \log(y) - \log(y_{-1})$, we have

$$\Delta \log(y) = \beta_0 + \mathbf{x}\boldsymbol{\beta} + (\alpha_1 - 1) \log(y_{-1}) + u.$$

So if we regress using the dependent variable $\Delta \log(y)$ we will have the same estimator for $\boldsymbol{\beta}$. The only change in the estimator will happen with the regression on $\log(y_{-1})$.

□

- b. Suppose that there are no covariates \mathbf{x} in the equation. Show that, if the distributions of y and y_{-1} are identical, then $|\alpha_1| < 1$. This is the *regression-to-the-mean* phenomenon in a dynamic setting. (Hint: Show that $\alpha_1 = \text{Corr}[\log(y), \log(y_{-1})]$.)

Answer:

Based on the prompt we will have the following linear model,

$$\log(y) = \beta_0 + \alpha_1 \log(y_{-1}) + u.$$

It is also given that the distribution of y and y_{-1} are identical, formally, $y, y_{-1} \stackrel{iid}{\sim} (\mu, \sigma^2)$ for some finite mean μ , and variance σ^2 . Recall the coefficient for simple regression, for this case we have

$$\alpha_1 = \frac{\text{Cov}[\log(y), \log(y_{-1})]}{\text{Var}[\log(y_{-1})]} \quad (10)$$

Recall the correlation relation with covariance and variance, we have

$$\begin{aligned}\text{Corr}[\log(y), \log(y_{-1})] &= \frac{\text{Cov}[\log(y), \log(y_{-1})]}{\sqrt{\text{Var}[\log(y_{-1})]\text{Var}[\log(y)]}} \\ \Leftrightarrow \text{Cov}[\log(y), \log(y_{-1})] &= \text{Corr}[\log(y), \log(y_{-1})] \sqrt{\text{Var}[\log(y_{-1})]\text{Var}[\log(y)]}\end{aligned}\quad (11)$$

Combining equations (10) and (11), and our assumption, $\text{Var}[\log(y_{-1})] = \text{Var}[\log(y)] = \sigma^2$, we have

$$\begin{aligned}\alpha_1 &= \frac{\text{Corr}[\log(y), \log(y_{-1})] \sqrt{\text{Var}[\log(y_{-1})]\text{Var}[\log(y)]}}{\text{Var}[\log(y_{-1})]} \\ &= \frac{\text{Corr}[\log(y), \log(y_{-1})] \sqrt{\sigma^2 \cdot \sigma^2}}{\sigma^2} \\ &= \text{Corr}[\log(y), \log(y_{-1})]\end{aligned}$$

Since we know $-1 < \alpha_1 = \text{Corr}[\log(y), \log(y_{-1})] < 1$, thus we can conclude that $|\alpha_1| < 1$. □

Part 2

Show that for a regression model, if a regressor x_j is measured with error, then it will be endogenous.

Answer:

Consider a regression with j variables:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_j x_j^* + v, \quad (12)$$

where y, x_1, \dots, x_{j-1} are observable, but x_j^* is not, instead we observe x_j that is measured with error. We assume that v has zero mean, i.e $E(v) = 0$, and is uncorrelated with $x_1, \dots, x_{j-1}, x_j^*$. We also assume v is uncorrelated with x_j , i.e $\text{Cov}(x_j, v) = 0$. Now define the measurement error of the population, u_j , as follows

$$x_j = x_j^* + u_j \Leftrightarrow x_j^* = x_j - u_j. \quad (13)$$

We assume that u_j has zero mean, i.e $E(u_j) = 0$, and is uncorrelated with $x_1, \dots, x_{j-1}, x_j^*$. Now substituting equation (13) into (12) we have

$$\begin{aligned}y &= \beta_0 + \beta_1 x_1 + \dots + \beta_j (x_j - u_j) + v \\ \Leftrightarrow y &= \beta_0 + \beta_1 x_1 + \dots + \beta_j x_j + (v - \beta_j u_j) \\ \Leftrightarrow y &= \beta_0 + \beta_1 x_1 + \dots + \beta_j x_j + e\end{aligned}\quad (14)$$

Let's denote $e = v - \beta_j u_j$ in equation (14), and we assume u_j and v are independent. With this measurement error, equation is the actual regression that we will be estimating rather than the original one. Previously, we have assume that $E(v) = 0$, and $E(u_j) = 0$, and are uncorrelated with each x_i including x_j , thus, we also have $E(e) = 0$. We want to assume that x_j is exogenous, i.e $\text{Cov}(x_j, e) = E(x_j e) = 0$. But we will have a problem of endogeneity in this case, that is

$$\begin{aligned}\text{Cov}(x_j, e) &= E(x_j e) \\ &= E((x_j^* + u_j)(v - \beta_j u_j)) \\ &= E(x_j^* v) + E(u_j v) - \beta_j E(x_j^* u_j) - \beta_j E(u_j^2) \\ &= 0 + 0 + 0 - \beta_j E(u_j^2) && [\text{because } x_j^* \perp v, x_j^* \perp u_j, u_j \perp v] \\ &= -\beta_j \text{Var}(u_j) \\ &\neq 0. && [\text{if } \beta_j \neq 0, \text{Var}(u_j) \neq 0]\end{aligned}$$

Not that in the presence of measurement error then $\text{Var}(u_j) \neq 0$. Thus in our case of explanatory variable measured with error, we have endogeneity. □