Homework 1
Intermediate Econometrics

Thomas DESIRE - Hugo SEGARD

October 5, 2023

Exercise 1

Question 1

First, let's study the consistency for β of:

$$\tilde{\beta} = \frac{\sum_{i=1}^n x_i^2 y_i}{\sum_{i=1}^n x_i^3}.$$

Let's make β appear on the right hand side:

$$\begin{aligned}
\tilde{\beta} &= \frac{\sum_{i=1}^n x_i^2 y_i}{\sum_{i=1}^n x_i^3} \\
&= \frac{\sum_{i=1}^n x_i^2 (x_i \beta + u_i)}{\sum_{i=1}^n x_i^3} \text{ [by plugging in the true model for } y_i] \\
&= \frac{\sum_{i=1}^n (x_i^3 \beta + x_i^2 u_i)}{\sum_{i=1}^n x_i^3} \\
&= \frac{\sum_{i=1}^n x_i^3 \beta + \sum_{i=1}^n x_i^2 u_i}{\sum_{i=1}^n x_i^3} \\
&= \frac{\sum_{i=1}^n x_i^3 \beta}{\sum_{i=1}^n x_i^3} + \frac{\sum_{i=1}^n x_i^2 u_i}{\sum_{i=1}^n x_i^3} \\
&= \beta + \frac{\sum_{i=1}^n x_i^2 u_i}{\sum_{i=1}^n x_i^3}.
\end{aligned}$$

Then we use this expression with isolated β and we multiply the numerator and denominator of the fraction by $\frac{1}{n}$:

$$\tilde{\beta} = \beta + \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2 u_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^3}.$$

Then by using the fact that we have an i.i.d. sample and that the variables have bounded moments of order 4, we can apply the Weak Law of Large Numbers and find the following probability limits:

$$\text{plim}(\frac{1}{n}\sum_{i=1}^{n} x_i^3) = E(x_i^3) \neq 0 \text{ (by assumption)}$$

$$\text{plim}(\frac{1}{n}\sum_{i=1}^{n} x_i^2 u_i) = E(x_i^2 u_i) = E(E(x_i^2 u_i \mid x_i)) = E(x_i^2 E(u_i \mid x_i)) = 0$$

(by Law of Iterated Expectations, and by assumption $E(u_i \mid x_i) = 0$).

Then we use the Continuous Mapping Theorem (knowing that the inverse function is continuous, except in 0 which is different from $E(x_i^3)$), and we get:

$$\begin{aligned}
\text{plim}(\tilde{\beta}) &= \text{plim}(\beta + \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2 u_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^3}) \\
&= \beta + \text{plim}(\frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2 u_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^3}) \\
&= \beta + \frac{\text{plim}(\frac{1}{n}\sum_{i=1}^{n} x_i^2 u_i)}{\text{plim}(\frac{1}{n}\sum_{i=1}^{n} x_i^3)} \text{ [Continuous Mapping Theorem]} \\
&= \beta + \frac{0}{E(x_i^3)} \\
&= \beta.
\end{aligned}$$

So $\tilde{\beta}$ is a consistent estimator for β.

Second, let's study the consistency for β of:

$$\check{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} + \frac{1}{n}.$$

Let's make β appear on the right hand side:

$$\begin{aligned}
\check{\beta} &= \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} + \frac{1}{n} \\
&= \frac{\sum_{i=1}^n x_i(x_i\beta + u_i)}{\sum_{i=1}^n x_i^2} + \frac{1}{n} \text{ [by plugging in the true model for } y_i\text{]} \\
&= \frac{\sum_{i=1}^n (x_i^2\beta + x_i u_i)}{\sum_{i=1}^n x_i^2} + \frac{1}{n} \\
&= \frac{\sum_{i=1}^n x_i^2\beta + \sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2} + \frac{1}{n} \\
&= \beta + \frac{\sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2} + \frac{1}{n}.
\end{aligned}$$

Then we use this expression with isolated β and we multiply the numerator and denominator of the fraction by $\frac{1}{n}$:

$$\check{\beta} = \beta + \frac{\frac{1}{n}\sum_{i=1}^n x_i u_i}{\frac{1}{n}\sum_{i=1}^n x_i^2} + \frac{1}{n}.$$

Then by using the fact that we have an i.i.d. sample and that the variables have bounded moments of order 4, we can apply the Weak Law of Large Numbers and find the following probability limits:

$$\text{plim}(\frac{1}{n}\sum_{i=1}^n x_i^2) = E(x_i^2) \neq 0 \text{ (by assumption)}$$

$$\text{plim}(\frac{1}{n}\sum_{i=1}^n x_i u_i) = E(x_i u_i) = E(x_i E(u_i \mid x_i)) = 0$$

(by Law of Iterated Expectations, and by assumption $E(u_i \mid x_i) = 0$).

Then we use the Continuous Mapping Theorem (knowing that the inverse function is continuous, except in 0 which is different from $E(x_i^2)$), and we get:

$$\begin{aligned}
\text{plim}(\check{\beta}) &= \text{plim}(\beta + \frac{\frac{1}{n}\sum_{i=1}^n x_i u_i}{\frac{1}{n}\sum_{i=1}^n x_i^2} + \frac{1}{n}) \\
&= \beta + \text{plim}(\frac{\frac{1}{n}\sum_{i=1}^n x_i u_i}{\frac{1}{n}\sum_{i=1}^n x_i^2}) + \text{plim}(\frac{1}{n}) \\
&= \beta + \frac{\text{plim}(\frac{1}{n}\sum_{i=1}^n x_i u_i)}{\text{plim}(\frac{1}{n}\sum_{i=1}^n x_i^2)} + 0 \\
&= \beta + \frac{0}{E(x_i^2)} \\
&= \beta.
\end{aligned}$$

So $\check{\beta}$ is a consistent estimator for β.

Third, let's study the consistency for β of:

$$\tilde{\tilde{\beta}} = \frac{1}{2}(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} + \tilde{\beta}).$$

From our study of $\check{\beta}$, we already know that:

$$\text{plim}(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}) = \beta.$$

Thus we have:

$$\begin{aligned}
\text{plim}(\tilde{\tilde{\beta}}) &= \text{plim}(\frac{1}{2}(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} + \tilde{\beta})) \\
&= \frac{1}{2}(\text{plim}(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}) + \text{plim}(\tilde{\beta})) \\
&= \frac{1}{2}(\beta + \beta) \\
&= \beta.
\end{aligned}$$

So $\tilde{\tilde{\beta}}$ is a consistent estimator for β.

Question 2

First, let's find the asymptotic distribution of $\tilde{\beta}$. From the first question, we have:

$$\tilde{\beta} = \beta + \frac{\sum_{i=1}^n x_i^2 u_i}{\sum_{i=1}^n x_i^3} \implies \sqrt{n}(\tilde{\beta} - \beta) = \sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^n x_i^2 u_i}{\frac{1}{n}\sum_{i=1}^n x_i^3} = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^n x_i^2 u_i}{\frac{1}{n}\sum_{i=1}^n x_i^3}.$$

For the denominator, we know that:

$$\text{plim}(\frac{1}{n}\sum_{i=1}^n x_i^3) = E(x_i^3) \neq 0.$$

For the numerator, we know that:

$$E(x_i^2 u_i) = 0 \text{ [as in question 1, with Law of Iterated Expectations]}$$

$$\begin{aligned} Var(x_i^2 u_i) &= E((x_i^2 u_i)^2) - E(x_i^2 u_i)^2 \\ &= E(x_i^4 u_i^2) \\ &= E(E(x_i^4 u_i^2 \mid x_i)) \text{ [L.I.E.]} \\ &= E(x_i^4 E(u_i^2 \mid x_i)) \\ &= E(x_i^4 \sigma^2) \\ &= \sigma^2 E(x_i^4). \end{aligned}$$

Now, we can apply the Central Limit Theorem (our data is i.i.d., and the variables have finite 4th moment):

$$\sqrt{n}(\frac{1}{n}\sum_{i=1}^n x_i^2 u_i - E(x_i^2 u_i)) \xrightarrow{d} \mathcal{N}(0, Var(x_i^2 u_i))$$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n x_i^2 u_i \xrightarrow{d} \mathcal{N}(0, \sigma^2 E(x_i^4)).$$

We have a ratio of two random variables, one of which converges in probability to a constant $\neq 0$ and one of which converges in distribution to a random element:

$$\text{plim}(\frac{1}{n}\sum_{i=1}^n x_i^3)^{-1} = E(x_i^3)^{-1}$$

[CMT, inverse function continuous, except in $0 \neq E(x_i^3)$]

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i^2 u_i \xrightarrow{d} \mathcal{N}(0, \sigma^2 E(x_i^4)).$$

Thus, we can use the Slutsky Theorem:

$$\sqrt{n}(\tilde{\beta} - \beta) = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i^2 u_i}{\frac{1}{n} \sum_{i=1}^{n} x_i^3} \xrightarrow{d} \frac{1}{E(x_i^3)} \mathcal{N}(0, \sigma^2 E(x_i^4)) \sim \mathcal{N}(0, \frac{\sigma^2 E(x_i^4)}{E(x_i^3)^2}).$$

Second, let's find the asymptotic distribution of $\check{\beta}$. From the first question, we have:

$$\check{\beta} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} + \frac{1}{n} \implies \sqrt{n}(\check{\beta} - \beta) = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i u_i}{\frac{1}{n} \sum_{i=1}^{n} x_i^2} + \frac{1}{\sqrt{n}}.$$

For the denominator, we know that:

$$\text{plim}(\frac{1}{n} \sum_{i=1}^{n} x_i^2) = E(x_i^2) \neq 0.$$

For the numerator, we know that:

$$E(x_i u_i) = E(x_i E(u_i \mid x_i)) = 0 \text{ [L.I.E.]}$$

$$\begin{aligned} Var(x_i u_i) &= E((x_i u_i)^2) - E(x_i u_i)^2 \\ &= E(x_i^2 u_i^2) \\ &= E(E(x_i^2 u_i^2 \mid x_i)) \\ &= E(x_i^2 E(u_i^2 \mid x_i)) \\ &= E(x_i^2 \sigma^2) \\ &= \sigma^2 E(x_i^2). \end{aligned}$$

Now, we can apply the Central Limit Theorem (our data is i.i.d., and the variables have finite 4th moment):

$$\sqrt{n}(\frac{1}{n} \sum_{i=1}^{n} x_i u_i - E(x_i u_i)) \xrightarrow{d} \mathcal{N}(0, Var(x_i u_i))$$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i u_i \xrightarrow{d} \mathcal{N}(0, \sigma^2 E(x_i^2)).$$

We have a ratio of two random variables, one of which converges in probability to a constant $\neq 0$ and one of which converges in distribution to a random element:

$$\text{plim}(\frac{1}{n}\sum_{i=1}^{n} x_i^2)^{-1} = E(x_i^2)^{-1}$$

[CMT, inverse function continuous, except in $0 \neq E(x_i^2)$]

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i u_i \xrightarrow{d} \mathcal{N}(0, \sigma^2 E(x_i^2)).$$

Thus, we can use the Slutsky Theorem:

$$\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i u_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \xrightarrow{d} \frac{1}{E(x_i^2)}\mathcal{N}(0, \sigma^2 E(x_i^2)) \sim \mathcal{N}(0, \frac{\sigma^2}{E(x_i^2)}).$$

Finally, since:

$$\text{plim}(\frac{1}{\sqrt{n}}) = 0$$

then we can use the Slutsky Theorem:

$$\sqrt{n}(\breve{\beta} - \beta) = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i u_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^2} + \frac{1}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \frac{\sigma^2}{E(x_i^2)}).$$

Exercise 2

Question 1

- The hypotheses are:

$$H_0 : \beta_E \leq 0.20 \text{ against } H_1 : \beta_E > 0.20.$$

- We use the following test statistic:

$$\hat{t} = \frac{\hat{\beta}_E - 0.20}{s.e.(\hat{\beta}_E)}.$$

  Under H_0, $\hat{t} \xrightarrow{d} \mathcal{N}(0, 1)$.

- The rejection rule for a 5% level test is to reject H_0 if:

$$\hat{t} > z_{1-0.05} = z_{0.95} = 1.6449.$$

- The value of the test statistic is:

$$\hat{t} = \frac{0.258 - 0.20}{0.013} = 4.46.$$

- Since 4.46 is larger than 1.6449, we reject the null in favor of the alternative at 5%.

From an economic point of view, we can conclude that working men aged 25 to 55 with a general baccalaureate or a higher education diploma level of education receive a hourly wage rate **at least** 20% superior than the hourly wage rate of the workers who do not have this level of education.

Question 2

- The hypotheses are:

$$H_0 : \beta_E = \beta_A \text{ against } H_1 : \beta_E \neq \beta_A$$

$$\text{i.e. } H_0 : \beta_E - \beta_A = 0 \text{ against } H_1 : \beta_E - \beta_A \neq 0.$$

- We can use the following t-test statistic:

$$\hat{t} = \frac{\hat{\beta_E} - \hat{\beta_A} - 0}{\sqrt{\hat{Var}(\hat{\beta_E} - \hat{\beta_A})}}.$$

  Under H_0, $\hat{t} \xrightarrow{d} \mathcal{N}(0,1)$.

- The rejection region is when $\left|\hat{t}\right| > z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$.

- We have $\hat{\beta_E} - \hat{\beta_A} = 0.258 - 0.132 = 0.126$. However, $\hat{Var}(\hat{\beta_E} - \hat{\beta_A}) = \hat{Var}(\hat{\beta_E}) + \hat{Var}(\hat{\beta_A}) - 2\hat{Cov}(\hat{\beta_E}, \hat{\beta_A})$. We do not have the estimated covariance between $\hat{\beta_E}$ and $\hat{\beta_A}$, so we cannot calculate this test statistic.

Question 3

The asymptotic confidence interval for β_A at confidence level 95% is:

$$[\hat{\beta_A} - z_{0.975} s.e.(\hat{\beta_A}); \hat{\beta_A} + z_{0.975} s.e.(\hat{\beta_A})]$$

$$[0.132 - 1.96 \times 0.012); 0.132 + 1.96 \times 0.012]$$

$$[0.10848; 0.15552].$$

This means that working men over 40 years old are estimated to earn between 10.848% and 15.552% more in hourly wage rate compared to those under 40 years old. If we repeated the study on several samples, the proportion of calculated confidence intervals encompassing the true value of β_A would tend toward 95%. Finally, the parameter estimate is statistically significant because 0 does not belong to the interval. So we are confident to say that a higher age positively influences the hourly wage rate of men aged 25 to 55.