Econometrics I
TA Session 5

Jukina HATAKEYAMA *

May 14, 2024

Contents

*E-mail: u868710a@ecs.osaka-u.ac.jp

1 Asymptotic Properties of OLSE

In this section, we review the asymptotic properties of the OLSE.

1.1 Single regression model

Suppose the single regression

$$y_i = \alpha + \beta x_i + u_i, \tag{1}$$

where $u_i \overset{i.i.d}{\sim} N_{\mathbb{R}}(0, \sigma^2)$.
The OLSE is can be written as

$$\hat{\beta} = \beta + \sum_{i=1}^{n} w_i u_i. \tag{2}$$

Recall that $w_i = (x_i - \bar{x})/\sum_{i=1}^n (x_i - \bar{x})^2$. Then, from the central limit theorem, we obtain

$$\frac{\sum_{i=1}^n w_i u_i - \mathbb{E}[\sum_{i=1}^n w_i u_i]}{\sqrt{\mathrm{Var}(\sum_{i=1}^n w_i u_i)}} = \frac{\hat{\beta} - \beta}{\sigma/\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \xrightarrow[n\to\infty]{d} N_{\mathbb{R}}(0, 1), \tag{3}$$

where $\mathbb{E}[\sum_{i=1}^n w_i u_i] = 0$, $\mathrm{Var}(\sum_{i=1}^n w_i u_i) = \sigma^2 \sum_{i=1}^n w_i^2 = \sigma^2/\sum_{i=1}^n (x_i - \bar{x})^2$ and $\sum_{i=1}^n w_i u_i = \hat{\beta} - \beta$. Additionally, the LLN implies that:

$$\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 \xrightarrow[n\to\infty]{p} \mathbb{E}[(x_i - \mu)^2]. \tag{4}$$

By using (4) and (3), we can apply the CLT as follows:

$$\frac{\sqrt{n}(\hat{\beta} - \beta)}{(\sigma^2 \sum_{i=1}^n (x_i - \bar{x})^2)^{-1/2}} \xrightarrow[n\to\infty]{d} N(0, 1). \tag{5}$$

Therefore, we can derive following relationship:

$$\hat{\beta} - \beta \xrightarrow[n\to\infty]{d} N(0, \sigma^2 \mathbb{E}[(x_i - \bar{x})^2]^{-1}).$$

2 Test Statistics

Ex.) Constraint OLS:

$$y = xb + u \quad where \quad Rb = q, \tag{6}$$

where $b_1 + b_2 = 1$, $R = (1, 1)$ and $q = 1$.
We can check whether $Rb - q = 0$ or not by the Wald test, . Now, we are going to explain a similar statistics, so called χ^2 statistics, which leads to the Wald statistics.

2.1 Chi-Square Distribution

We first review the chi-square distribution since the Wald statistics follows this distribution.

Theorem 2.1. Suppose that a random variable $X \in \mathbb{R}^{n\times 1}$ follows a normal distribution $X \sim N(\mu, V)$, where $V \in \mathbb{R}^{n\times n}$ is positive definite. Then, a random variable $W_0 = (X-\mu)'V^{-1}(X-\mu)$ follows a $\chi^2(n)$ distribution. Its probability density function is given as follows:

$$f_n(w_0) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} w_0^{\frac{n}{2}-1} e^{-\frac{w_0}{2}},$$

where $\Gamma(\frac{n}{2})$ is a gamma function.

Keep in mind that χ^2 distribution with n freedom is defined by the summation of the square of n random variables which follows the standard normal distribution, $N_{\mathbb{R}}(0,1)$. By using this definition, the proof of the above theorem is not so difficult.

Proof. Suppose that V can be decomposed as:

$$V = C'\Lambda C.$$

Then, we can calculate $V = V^{1/2}V^{1/2}$ where $V^{1/2} = C'\Lambda^{1/2}C$. In addition, we can say $Z \equiv V^{-1/2}(X-\mu) \sim N_{\mathbb{R}^{dim(X)}}(0, I)$ by the properties of the multivariate normal distribution. Let $W_0 \equiv Z'Z = \sum_{i=1}^{n} Z_i^2$. Because each Z_i follows the standard normal distribution, Z_i^2 follows $\chi^2(1)$ distribution. Therefore, $W_0 \sim \chi^2(n)$ is proven. $\square$

2.2 Delta Method

Consider the case of a parameter θ_0 to be estimated and sequence of its estimator $\hat{\theta_n}$. If $\hat{\theta_n}$ has an asymptotic normality, $\sqrt{n}(\hat{\theta_n} - \theta_0) \xrightarrow{d} N(0, \Sigma)$, we can derive the following theorem.

Theorem 2.2. Suppose any continuous, differentiable function $g : \mathbb{R}^d \to \mathbb{R}^s$. Let $X_1, X_2, \ldots, X_n$ a sequence of d-dimensional random variables. If $\hat{\theta_n}$ has an asymptotic normality, we can state that:

$$\sqrt{n}[g(\hat{\theta_n}) - g(\theta_0)] \xrightarrow[n\to\infty]{d} N_{\mathbb{R}^{dim(g)}}(0, D_g(\theta_0)\Sigma D'_g(\theta_0)), \tag{7}$$

where $D_g(\theta)$ is a Jacobian matrtix of θ.

Suppose that there is a parameter $\bar{\theta} \in (\theta_0, \hat{\theta_n})$. Then, we can use a mean value expansion:

$$g(\hat{\theta_n}) = g(\theta_0) + D_g(\bar{\theta})(\hat{\theta_n} - \theta_0).$$

By using this expression, we can prove (7).

Proof. Because of the mean value expansion, we can state:

$$\begin{aligned}\sqrt{n}[g(\hat{\theta_n}) - g(\theta_0)] &= \sqrt{n}D_g(\bar{\theta})(\hat{\theta_n} - \theta_0) \\ &= \sqrt{n}D_g(\theta_0)(\hat{\theta_n} - \theta_0) + \sqrt{n}[D_g(\bar{\theta}) - D_g(\theta_0)](\hat{\theta_n} - \theta_0) \end{aligned} \tag{8}$$

Here, if $\hat{\theta_n} \xrightarrow[n\to\infty]{p} \theta_0$ is given, $\bar{\theta} \xrightarrow[n\to\infty]{p} \theta_0$ is established. Therefore, the second term of (RHS) in the (8) is calculated as:

$$[D_g(\bar{\theta}) - D_g(\theta_0)]\sqrt{n}(\hat{\theta_n} - \theta_0) = o_p(1)O_p(1) = o_p(1), \tag{9}$$

because $\sqrt{n}(\hat{\theta_n} - \theta_0)$ converges in distribution.[1] By the (8) and the (9), we can derive:

$$\sqrt{n}[g(\hat{\theta_n}) - g(\theta_0)] = \sqrt{n}D_g(\theta_0)(\hat{\theta_n} - \theta_0) + o_p(1). \tag{10}$$

Since we can apply the property of multivariate normal distribution in this equation, we can say $\sqrt{n}D_g(\theta_0)(\hat{\theta_n} - \theta_0) \xrightarrow[n\to\infty]{d} N_{\mathbb{R}^{dim(g)}}(0, D_g(\theta_0)\Sigma D'_g(\theta_0))$. We can now apply the asymptotic equivalence lemma of the main text.

Lemma 2.3. Let $\{x_n\}$ and $\{y_n\}$ be sequences of $n \times 1$ random vectors. If $z_n \xrightarrow[n\to\infty]{d} z$ and $x_n - z_n \xrightarrow[n\to\infty]{p} 0$, then $x_n \xrightarrow[n\to\infty]{d} z$.

By using this lemma, we can derive (7). □

[1]The lemma 3.5 in the main textbook implies that a $K \times 1$ vector x_n is $O_p(1)$ if x_n converges x in distribution.

2.3 Test Statistics

Suppose that $\{\hat{b}_n : n = 1, 2, \dots\}$ be a sequence of estimators which satisfies:

$$\sqrt{n}(\hat{b_n} - b) \xrightarrow[n\to\infty]{d} N_{\mathbb{R}^{dim(b^n)}}(0, V)$$

where $V > 0$ is the asymptotic variance covariance matrix of $\sqrt{n}R(\hat{b_n} - b)$ and $R \in \mathbb{R}^{q\times k}$ with $q \leq K$ and $\text{rank}(R) = q$. Then, following lemma is derived.

Lemma 2.4. In the above settings, $\sqrt{n}R(\hat{b} - b) \xrightarrow[n\to\infty]{d} N_{\mathbb{R}^q}(0, RVR')$ and:

$$[\sqrt{n}R(\hat{b_n} - b)]'(RVR')^{-1}[\sqrt{n}R(\hat{b_n} - b)] \xrightarrow[n\to\infty]{d} \chi^2(q).$$

In addition, if $\hat{V}$(the estimator of V) has consistency, then:

$$[\sqrt{n}R(\hat{b_n} - b)]'(R\hat{V}R')^{-1}[\sqrt{n}R(\hat{b_n} - b)] \xrightarrow[n\to\infty]{d} \chi^2(q).$$

Proof. [2] If $\sqrt{n}(\hat{b_n} - b) \xrightarrow{d} N(0, V)$ as $n \to \infty$, then $\sqrt{n}R(\hat{b_n} - b) \xrightarrow[n\to\infty]{d} N(0, RVR')$ is derived. Assume that x is written as follows:

$$x = [\sqrt{n}R(\hat{b_n} - b)]'Q_n^{-1}[\sqrt{n}R(\hat{b_n} - b)],$$

where $Q_n = R\hat{V}R'$($\hat{V}$ is a consistent estimator of V) and $c_n = \sqrt{n}R(\hat{b_n} - b)$. Then, $c_n \xrightarrow[n\to\infty]{d} c$ where $c \sim N_{\mathbb{R}^c}(0, RVR')$ and $Q_n \xrightarrow{d} Q$ where $Q = RVR'$. Because R is full rank and V is positive definite, Q is invertible. Therefore, $W \xrightarrow{d} c'Q^{-1}c \sim \chi^2(n)$ by the Theorem 3.1. □

2.4 Review of t Statistics

In Econometrics class, t distribution is applied to the statistical test and the confidence interval of OLSE. Therfore, we shortly review of this statistics.

2.4.1 General Definition

Suppose the case that the sequence of random variables $X_i(i = 1, \cdots, n)$. Each X_i follows $i.i.d.$ normal distribution such that $N(\mu, \sigma^2)$. By applying CLT, we have:

$$\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \xrightarrow[n\to\infty]{d} Z, \tag{11}$$

where Z follows a standard normal distribution.If we do not know the true variance σ^2, we use an estimator of the sample variance, s^2:

$$t = \frac{\bar{X} - \mu}{\sqrt{s^2/n}}, \tag{12}$$

[2]This proof is explained in Chapter 2 of Fumio, Hayashi(2000) "ECONOMETRICS", PRINCETON UNIVERSITY PRESS.

where $s^2 = \frac{1}{n-1}[(X_1 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2]$.Now, we can rewrite (12) as follows:

$$t = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} / \sqrt{\frac{(n-1)s^2}{\sigma^2} / (n-1)}. \tag{13}$$

The numerator follows standard normal distribution and a part of denominator, $\frac{(n-1)s^2}{\sigma^2}$, follows χ^2 distribution with $n-1$ freedom.

2.4.2 t Statistics for OLSE

In the case of OLSE of (2), as we learn in Section 1, the following relationship is established:

$$\frac{\hat{\beta} - \beta}{\sigma / \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \xrightarrow[n \to \infty]{d} N_{\mathbb{R}}(0, 1).$$

In this case, replacing σ by its estimator $\hat{\sigma}$, we obtain t statistics such as:

$$t = \frac{\hat{\beta} - \beta}{\hat{\sigma} / \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \sim t(n-2). \tag{14}$$

We will explain the reason why we can derive above equation. At first, we must consider how to derive $\hat{\sigma}^2$.

Theorem 2.5. Under the assumption of the classical OLS model, the (unbiased) estimator of σ^2 is given as follows:

$$\hat{\sigma^2} = \frac{1}{n-2} \sum_{i=1}^{n} \hat{u_i}^2. \tag{15}$$

Proof. In the (15), we can say $\sum_{i=1}^{n} \hat{u_i}^2 = \hat{u}'\hat{u}$ and the residual is rewritten as follows:

$$\hat{u} = M_x \hat{u} = [I_z - x(x'x)^{-1}x']u. \tag{16}$$

Therefore, $\hat{u}'\hat{u}$ is calculated as follows:

$$\begin{aligned} \hat{u}'\hat{u} &= u' M_z u \\ &= tr(u' M_x u) \\ &= tr(M_x u' u) \end{aligned} \tag{17}$$

Then, the expectation of (17) is given as follows:

$$\begin{aligned} \mathbb{E}(\hat{u}'\hat{u}) &= E[tr(M_x u'u)] \\ &= tr[\mathbb{E}(M_x u'u)] \\ &= tr[\mathbb{E}(\mathbb{E}(M_x u'u|x))] \\ &= tr[\mathbb{E}(M_x \mathbb{E}(uu'|x))] \\ &= \sigma^2 tr[\mathbb{E}(M_x)] \end{aligned} \tag{18}$$

In the above equation, $tr[\mathbb{E}(M_z)]$ is represented as follows:

$$
\begin{aligned}
tr[\mathbb{E}(M_x)] &= tr[I_n - \mathbb{E}(x(x'x)^{-1}x')] \\
&= n - tr[\mathbb{E}(x(x'x)^{-1}x')] \\
&= n - tr[(x'x)^{-1}x'x] \\
&= n - tr[I_x] \\
&= n - 2. \qquad (19)
\end{aligned}
$$

From these equations, above theorem is proven. Note that this proof follows the exact same steps as in the case of K covariates. $\square$

Finally, we can confirm that t statistic in (14) follows a t distribution whose degrees of freedom is equal to $n-2$.

$$
t = \frac{\hat{\beta} - \beta}{\sqrt{\sigma^2 / \sum_{i=1}^{n}(x_i - \bar{x})^2}} / \sqrt{\frac{(n-2)\hat{\sigma}^2}{\sigma^2} / (n-2)} \qquad (20)
$$

The numerator follows standard normal distribution and a part of denominator, $\frac{(n-2)\hat{\sigma}^2}{\sigma^2}$, follows χ^2 distribution whose degrees of freedom is equal to $n-2$.

2.5 R Exercise

In this subsection, we will explain how to use R. Today, we use data of the speed of cars and the distances taken to stop recorded in the 1920s. Consider the following regression model:

$$(distance)_i = a + b(speed)_i + u_i.$$

The result of this estimation is easily outputted by the stargazer package. This package makes a table of estimation by tex.

Table 1:

	Dependent variable:
	dist
speed	3.932^{***}
	(0.416)
	t = 9.464
	p = 0.000
Constant	-17.579^{**}
	(6.758)
	t = −2.601
	p = 0.013
Observations	50
R^2	0.651
Adjusted R^2	0.644
Residual Std. Error	15.380 (df = 48)
F Statistic	89.567^{***} (df = 1; 48) (p = 0.000)
Note:	*p<0.1; **p<0.05; ***p<0.01

We can make a figure of a regression line by pdf file. R code is given in the Appendix. The lm function is a default function of R.

3 Appendix

3.1 R code

```
library(stargazer)
data(cars)
#Let us check the single regression model by using "cars" data set.
fix(cars)

speed<-cars[,1]
dist<-cars[,2]

cars.lm<-lm(dist~speed)
stargazer(cars.lm,style="all",type="latex")
plot(cars)
abline(cars.lm, lwd=1, col="blue")

#A command to make pdf file.
pdf("carsdata.pdf")
plot(cars)
par(new=T)
abline(cars.lm, lwd=1, col="blue")
dev.off()
```