

# Econometrics II

(Thu., 8:50-10:20)

Room # 4 (法経講義棟)

- The prerequisites of this class are **Special Lectures in Economics (Statistical Analysis)**, 経済学特論（統計解析） (last semester) and **Econometrics I (エコノメトリックス I)** (graduate level, last semester).

# TA Session

by **Simoshimizu** (D2 下清水 慎, [shimosshi1q94 \[at\] gmail.com](mailto:shimosshi1q94@gmail.com))

and **Kudo** (D1 工藤 健太, [u035478k \[at\] ecs.osaka-u.ac.jp](mailto:u035478k@ecs.osaka-u.ac.jp))

**From** Oct. 8, 2019

**Tue.** 10:30 - 12:00

**Room** # 4 (法経講義棟)

- **Download the lecture notes from the following websites:**

<http://www2.econ.osaka-u.ac.jp/~tanizaki/class/2019/econome2/>

<http://stat.econ.osaka-u.ac.jp/~tanizaki/class/2019/econome2/>

# Contents

<b>1</b>	<b>Maximum Likelihood Estimation (MLE, 最尤法)</b> — Review	<b>7</b>
<b>2</b>	<b>Qualitative Dependent Variable (質的従属変数)</b>	<b>30</b>
2.1	Discrete Choice Model (離散選択モデル) . . . . .	31
2.1.1	Binary Choice Model (二値選択モデル) . . . . .	31
2.2	Limited Dependent Variable Model (制限従属変数モデル) . . . . .	52
2.3	Count Data Model (計数データモデル) . . . . .	61
<b>3</b>	<b>Panel Data</b>	<b>69</b>
3.1	GLS — Review . . . . .	69
3.2	Panel Model Basic . . . . .	73
3.2.1	Fixed Effect Model (固定効果モデル) . . . . .	74

3.2.2	Random Effect Model (ランダム効果モデル) . . . . .	82
3.3	Hausman's Specification Error (特定化誤差) Test . . . . .	86
3.4	Choice of Fixed Effect Model or Random Effect Model . . . . .	88
3.4.1	The Case where $X$ is Correlated with $u$ — Review . . . . .	88
3.4.2	Fixed Effect Model or Random Effect Model . . . . .	90
<b>4</b>	<b>Applications</b>	<b>92</b>
<b>5</b>	<b>Generalized Method of Moments (GMM, 一般化積率法)</b>	<b>123</b>
5.1	Method of Moments (MM, 積率法) . . . . .	123
5.2	Generalized Method of Moments (GMM, 一般化積率法) . . . . .	131
5.3	Generalized Method of Moments (GMM, 一般化積率法) II — Non-linear Case — . . . . .	153

<b>6</b>	<b>Time Series Analysis (時系列分析)</b>	<b>175</b>
6.1	Introduction . . . . .	175
6.2	Autoregressive Model (自己回帰モデル or AR モデル) . . . . .	180
6.3	MA Model . . . . .	209
6.4	ARMA Model . . . . .	224
6.5	ARIMA Model . . . . .	229
6.6	SARIMA Model . . . . .	230
6.7	Optimal Prediction . . . . .	230
6.8	Identification . . . . .	233
6.9	Example of SARIMA using Consumption Data . . . . .	236
<b>7</b>	<b>Unit Root (単位根) and Cointegration (共和分)</b>	<b>241</b>
7.1	Unit Root (単位根) Test (Dickey-Fuller (DF) Test) . . . . .	241
7.2	Serially Correlated Errors . . . . .	277

7.2.1	Augmented Dickey-Fuller (ADF) Test . . . . .	278
7.3	Cointegration (共和分) . . . . .	281
7.4	Testing Cointegration . . . . .	298
7.4.1	Engle-Granger Test . . . . .	298

# 1 Maximum Likelihood Estimation (MLE, 最尤法) —

## Review

1. We have random variables  $X_1, X_2, \dots, X_n$ , which are assumed to be mutually independently and identically distributed.
2. The distribution function of  $\{X_i\}_{i=1}^n$  is  $f(x; \theta)$ , where  $x = (x_1, x_2, \dots, x_n)$  and  $\theta = (\mu, \Sigma)$ .

Note that  $X$  is a vector of random variables and  $x$  is a vector of their realizations (i.e., observed data).

Likelihood function  $L(\cdot)$  is defined as  $L(\theta; x) = f(x; \theta)$ .

Note that  $f(x; \theta) = \prod_{i=1}^n f(x_i; \theta)$  when  $X_1, X_2, \dots, X_n$  are mutually indepen-

dently and identically distributed.

The maximum likelihood estimator (MLE) of  $\theta$  is  $\theta$  such that:

$$\max_{\theta} L(\theta; X). \quad \iff \quad \max_{\theta} \log L(\theta; X).$$

MLE satisfies the following two conditions:

(a)  $\frac{\partial \log L(\theta; X)}{\partial \theta} = 0.$

(b)  $\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}$  is a negative definite matrix.

3. **Fisher's information matrix** (フィッシャーの情報行列) is defined as:

$$I(\theta) = -E\left(\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}\right),$$

where we have the following equality:

$$-E\left(\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}\right) = E\left(\frac{\partial \log L(\theta; X)}{\partial \theta} \frac{\partial \log L(\theta; X)}{\partial \theta'}\right) = V\left(\frac{\partial \log L(\theta; X)}{\partial \theta}\right)$$



**Proof of the above equality:**

$$\int L(\theta; x)dx = 1$$

Take a derivative with respect to  $\theta$ .

$$\int \frac{\partial L(\theta; x)}{\partial \theta} dx = 0$$

(We assume that (i) the domain of  $x$  does not depend on  $\theta$  and (ii) the derivative  $\frac{\partial L(\theta; x)}{\partial \theta}$  exists.)

Rewriting the above equation, we obtain:

$$\int \frac{\partial \log L(\theta; x)}{\partial \theta} L(\theta; x) dx = 0,$$

i.e.,

$$E\left(\frac{\partial \log L(\theta; X)}{\partial \theta}\right) = 0.$$

Again, differentiating the above with respect to  $\theta$ , we obtain:

$$\begin{aligned}
 & \int \frac{\partial^2 \log L(\theta; x)}{\partial \theta \partial \theta'} L(\theta; x) dx + \int \frac{\partial \log L(\theta; x)}{\partial \theta} \frac{\partial L(\theta; x)}{\partial \theta'} dx \\
 &= \int \frac{\partial^2 \log L(\theta; x)}{\partial \theta \partial \theta'} L(\theta; x) dx + \int \frac{\partial \log L(\theta; x)}{\partial \theta} \frac{\partial \log L(\theta; x)}{\partial \theta'} L(\theta; x) dx \\
 &= E\left(\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}\right) + E\left(\frac{\partial \log L(\theta; X)}{\partial \theta} \frac{\partial \log L(\theta; X)}{\partial \theta'}\right) = 0.
 \end{aligned}$$

Therefore, we can derive the following equality:

$$-E\left(\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}\right) = E\left(\frac{\partial \log L(\theta; X)}{\partial \theta} \frac{\partial \log L(\theta; X)}{\partial \theta'}\right) = V\left(\frac{\partial \log L(\theta; X)}{\partial \theta}\right),$$

where the second equality utilizes  $E\left(\frac{\partial \log L(\theta; X)}{\partial \theta}\right) = 0$ .

4. **Cramer-Rao Lower Bound** (クラメール・ラオの下限):  $(I(\theta))^{-1}$

Suppose that an unbiased estimator of  $\theta$  is given by  $s(X)$ .

Then, we have the following:

$$V(s(X)) \geq (I(\theta))^{-1}$$

**Proof:**

The expectation of  $s(X)$  is:

$$E(s(X)) = \int s(x)L(\theta; x)dx.$$

Differentiating the above with respect to  $\theta$ ,

$$\begin{aligned} \frac{\partial E(s(X))}{\partial \theta} &= \int s(x) \frac{\partial L(\theta; x)}{\partial \theta} dx = \int s(x) \frac{\partial \log L(\theta; x)}{\partial \theta} L(\theta; x) dx \\ &= \text{Cov} \left( s(X), \frac{\partial \log L(\theta; X)}{\partial \theta} \right) \end{aligned}$$

For simplicity, let  $s(X)$  and  $\theta$  be scalars.

Then,

$$\begin{aligned} \left( \frac{\partial \mathbb{E}(s(X))}{\partial \theta} \right)^2 &= \left( \text{Cov} \left( s(X), \frac{\partial \log L(\theta; X)}{\partial \theta} \right) \right)^2 = \rho^2 \mathbb{V}(s(X)) \mathbb{V} \left( \frac{\partial \log L(\theta; X)}{\partial \theta} \right) \\ &\leq \mathbb{V}(s(X)) \mathbb{V} \left( \frac{\partial \log L(\theta; X)}{\partial \theta} \right), \end{aligned}$$

where  $\rho$  denotes the correlation coefficient between  $s(X)$  and  $\frac{\partial \log L(\theta; X)}{\partial \theta}$ , i.e.,

$$\rho = \frac{\text{Cov} \left( s(X), \frac{\partial \log L(\theta; X)}{\partial \theta} \right)}{\sqrt{\mathbb{V}(s(X))} \sqrt{\mathbb{V} \left( \frac{\partial \log L(\theta; X)}{\partial \theta} \right)}}.$$

Note that  $|\rho| \leq 1$ .

Therefore, we have the following inequality:

$$\left(\frac{\partial \mathbb{E}(s(X))}{\partial \theta}\right)^2 \leq \mathbb{V}(s(X)) \mathbb{V}\left(\frac{\partial \log L(\theta; X)}{\partial \theta}\right),$$

i.e.,

$$\mathbb{V}(s(X)) \geq \frac{\left(\frac{\partial \mathbb{E}(s(X))}{\partial \theta}\right)^2}{\mathbb{V}\left(\frac{\partial \log L(\theta; X)}{\partial \theta}\right)}$$

Especially, when  $\mathbb{E}(s(X)) = \theta$ ,

$$\mathbb{V}(s(X)) \geq \frac{1}{-\mathbb{E}\left(\frac{\partial^2 \log L(\theta; X)}{\partial \theta^2}\right)} = (I(\theta))^{-1}.$$

Even in the case where  $s(X)$  is a vector, the following inequality holds.

$$\mathbb{V}(s(X)) \geq (I(\theta))^{-1},$$

where  $I(\theta)$  is defined as:

$$\begin{aligned} I(\theta) &= -\mathbb{E}\left(\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}\right) \\ &= \mathbb{E}\left(\frac{\partial \log L(\theta; X)}{\partial \theta} \frac{\partial \log L(\theta; X)}{\partial \theta'}\right) = \mathbb{V}\left(\frac{\partial \log L(\theta; X)}{\partial \theta}\right). \end{aligned}$$

The variance of any unbiased estimator of  $\theta$  is larger than or equal to  $(I(\theta))^{-1}$ .

## 5. Asymptotic Normality of MLE:

Let  $\tilde{\theta}$  be MLE of  $\theta$ .

As  $n$  goes to infinity, we have the following result:

$$\sqrt{n}(\tilde{\theta} - \theta) \longrightarrow N\left(0, \lim_{n \rightarrow \infty} \left(\frac{I(\theta)}{n}\right)^{-1}\right),$$

where it is assumed that  $\lim_{n \rightarrow \infty} \left(\frac{I(\theta)}{n}\right)$  converges.

That is, when  $n$  is large,  $\tilde{\theta}$  is approximately distributed as follows:

$$\tilde{\theta} \sim N\left(\theta, (I(\theta))^{-1}\right).$$

Suppose that  $s(X) = \tilde{\theta}$ .

When  $n$  is large,  $V(s(X))$  is approximately equal to  $(I(\theta))^{-1}$ .

Practically, we utilize the following approximated distribution:

$$\tilde{\theta} \sim N\left(\theta, (I(\tilde{\theta}))^{-1}\right).$$

Then, we can obtain the significance test and the confidence interval for  $\theta$

6. **Central Limit Theorem:** Let  $X_1, X_2, \dots, X_n$  be mutually independently distributed random variables with mean  $E(X_i) = \mu$  and variance  $V(X_i) = \sigma^2 < \infty$  for  $i = 1, 2, \dots, n$ .

Define  $\bar{X} = (1/n) \sum_{i=1}^n X_i$ .

Then, the central limit theorem is given by:

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \longrightarrow N(0, 1).$$

Note that  $E(\bar{X}) = \mu$  and  $V(\bar{X}) = \sigma^2/n$ .



That is,

$$\sqrt{n}(\bar{X} - \mu) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \mu) \longrightarrow N(0, \sigma^2).$$

Note that  $E(\bar{X}) = \mu$  and  $nV(\bar{X}) = \sigma^2$ .

In the case where  $X_i$  is a vector of random variable with mean  $\mu$  and variance  $\Sigma < \infty$ , the central limit theorem is given by:

$$\sqrt{n}(\bar{X} - \mu) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \mu) \longrightarrow N(0, \Sigma).$$

Note that  $E(\bar{X}) = \mu$  and  $nV(\bar{X}) = \Sigma$ .

7. **Central Limit Theorem II:** Let  $X_1, X_2, \dots, X_n$  be mutually independently distributed random variables with mean  $E(X_i) = \mu$  and variance  $V(X_i) = \sigma_i^2$  for  $i = 1, 2, \dots, n$ .

Assume:

$$\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \sigma_i^2 < \infty.$$

Define  $\bar{X} = (1/n) \sum_{i=1}^n X_i$ .

The central limit theorem is given by:

$$\sqrt{n}(\bar{X} - \mu) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \mu) \longrightarrow N(0, \sigma^2).$$

Note that  $E(\bar{X}) = \mu$  and  $nV(\bar{X}) \longrightarrow \sigma^2$ .

In the case where  $X_i$  is a vector of random variable with mean  $\mu$  and variance  $\Sigma_i$ , the central limit theorem is given by:

$$\sqrt{n}(\bar{X} - \mu) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \mu) \longrightarrow N(0, \Sigma),$$

where  $\Sigma = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Sigma_i < \infty$ .

Note that  $E(\bar{X}) = \mu$  and  $nV(\bar{X}) \longrightarrow \Sigma$ .

### [Review of Asymptotic Theories]

- **Convergence in Probability (確率収束)**  $X_n \longrightarrow a$ , i.e.,  $X$  converges in probability to  $a$ , where  $a$  is a fixed number.