Econometrics II TA Session

Wang Xin

Oct 16, 2023

About this TA session

- This TA session is about Q&A, reviews and supplements of Econometrics 2, and explanation of homework (if this is any).
- The aim of this TA session is to help you understand the knowledge taught in formal class, so the supplements (which are not included in the Prof's notes) will not appear in exams, and this session will not last long, depending on the content.
- We will not have class on Nov 27, but on Nov 29. I will inform you again a week before.
- You are encouraged to email us your question before this class.

In last two weeks ...

- Definition and expression
- Fisher's information matrix
- Cramer-Rao Lower Bound
- Central Limit Theorem and Law of Large Numbers
- Asymptotic Normality of MLE

Contents

Definition

We consider i.i.d. random variables $X_1, \cdot \cdot \cdot, X_n$. A method for finding an estimator $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(X_1, \cdot \cdot \cdot, X_n)$ is to maximize a criterion function of the type

$$\underset{\boldsymbol{\theta}}{max} \; M_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} l(X_i; \boldsymbol{\theta})$$

$l(X_i; \boldsymbol{\theta}) : X_i \in \mathbf{R}$ is a known function and corresponds to the statistical model of interest.

- Least square vs Maximum likelihood

Suppose $X_1, \cdot \cdot \cdot, X_n$ have a probability density function $p_{\boldsymbol{\theta}}(X_i)$. Then the maximum likelihood estimator maximizes the likelihood function $\prod_{i=1}^{n} p_{\boldsymbol{\theta}}(X_i)$, or equivalently the log likelihood

$$\underset{\boldsymbol{\theta}}{max} \ \ M_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} l(X_i; \boldsymbol{\theta}), \ l(X_i; \boldsymbol{\theta}) = \log(p_{\boldsymbol{\theta}}(X_i)).$$

The MLE satisfies the statistical criterion

$$\hat{\boldsymbol{\theta}} = arg \ \underset{\boldsymbol{\theta}}{max} \ \ M_n(\boldsymbol{\theta}),$$

and thus satisfies the orthogonality condition

$$\frac{\partial M_n(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = 0.$$

We assume that $\frac{\partial^2 M_n(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} < 0$.

Theoretical Properties

Finite sample properties of the MLE (n is fixed)

- the MLE is generally biased but it is feasible to obtain an unbiased one under proper transformations.
- the MLE is efficient, that it matches with an estimator that attains the Cramer-Rao lower bound.

Large sample properties of the MLE (n goes to infinite)

- Consistency The MLE satisfies $\hat{\boldsymbol{\theta}} \underset{n\to\infty}{\to} \boldsymbol{\theta}$, the true parameter.
- Normality $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \underset{n\to\infty}{\to} \mathcal{N}(0, \Sigma)$

Theorem

Asymptotic distribution: when n is large enough, $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \Sigma_\theta)$.
The variance covariance matrix is defined as

$$\Sigma_\theta = (I(\theta))^{-1} = -E\left(\frac{\partial^2 M_n(\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)^{-1} = -\left(\sum_{i=1}^{n} E[\partial^2_{\theta\theta'} \log(p(X_i))]\right)^{-1}$$

Examples

1. Random variables with normal distribution

The random variables $X_1, \cdot \cdot \cdot, X_n$ are assumed independent and each of them follows $X_i \sim \mathcal{N}(\mu, \sigma^2)$.
The probability density function of $\mathcal{N}(\mu, \sigma^2)$ is given as

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

$$\begin{aligned} \Rightarrow M_n(\theta) &= \sum_{i=1}^{n} \log(p_\theta(X_i)) \\ &= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x-\mu)^2 \end{aligned}$$

1. Random variables with normal distribution

The random variables $X_1, \cdots, X_n$ are assumed independent and each of them follows $X_i \sim \mathcal{N}(\mu, \sigma^2)$.
The probability density function of $\mathcal{N}(\mu, \sigma^2)$ is given as

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

- σ^2 is known $\quad \Rightarrow \quad \mu \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$

1. Random variables with normal distribution

The random variables $X_1, \cdot\cdot\cdot, X_n$ are assumed independent and each of them follows $X_i \sim \mathcal{N}(\mu, \sigma^2)$.
The probability density function of $\mathcal{N}(\mu, \sigma^2)$ is given as

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

- σ^2 is unknown

$$\Rightarrow \quad \begin{pmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix}\right)$$

2. Simple Regression

We consider the simple regression model

$$Y_i = \alpha + \beta X_i + u_i$$

We assume $u_i | X_i \sim \mathcal{N}(0, \sigma^2)$. As a consequence, the density is

$$p_{u;\theta}(u_i|X_i) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{u_i^2}{2\sigma^2}\right).$$

$$\Rightarrow \qquad \boldsymbol{\theta} = \left(\alpha \ \beta \ \sigma^2\right)'$$

$$\begin{aligned} M_n(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \log p_\theta(Y_i, X_i) \\ &= \sum_{i=1}^{n} \log p_\theta(Y_i|X_i) + \sum_{i=1}^{n} \log p_\theta(X_i) = \sum_{i=1}^{n} \log p_\theta(Y_i|X_i) \end{aligned}$$

2. Simple Regression

We consider the simple regression model

$$Y_i = \alpha + \beta X_i + u_i$$

We assume $u_i | X_i \sim \mathcal{N}(0, \sigma^2)$. As a consequence, the density is

$$p_{u;\theta}(u_i|X_i) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{u_i^2}{2\sigma^2}\right).$$

$$\begin{aligned}
M_n(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \log p_\theta(Y_i|X_i) \\
&= \sum_{i=1}^{n} \log\left(p_{u;\theta}(f^{-1}(Y_i)|X_i)\left|\frac{\partial f^{-1}(Y_i)}{\partial Y_i}\right|\right) \\
&= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \alpha - \beta X_i)^2
\end{aligned}$$

2. Simple Regression

We consider the simple regression model

$$Y_i = \alpha + \beta X_i + u_i$$

We assume $u_i|X_i \sim \mathcal{N}(0, \sigma^2)$. As a consequence, the density is

$$p_{u;\theta}(u_i|X_i) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{u_i^2}{2\sigma^2}\right).$$

$$\left(\begin{array}{c} \hat{\alpha} \\ \hat{\beta} \\ \hat{\sigma}^2 \end{array}\right) \sim \mathcal{N}\left(\left(\begin{array}{c} \alpha \\ \beta \\ \sigma^2 \end{array}\right), \left(\begin{array}{ccc} \frac{n}{\sigma^2} & \frac{nE[X_i]}{\sigma^2} & 0 \\ \frac{nE[X_i]}{\sigma^2} & \frac{nE[X_i^2]}{\sigma^2} & 0 \\ 0 & 0 & \frac{n}{2\sigma^4} \end{array}\right)^{-1}\right)$$