Econometrics 2 (2018) TA session 1*

Kenji Hatakenaka †

4 October 2018

目次

1 Maximum Likelihood Estimator

$X_1, X_2, \cdots, X_n$: i.i.d. random variables
$f(x|\theta)$: probability density function of X, θ is parameter
The likelihood function of X is

$$L(\theta|x) \equiv \prod_{i=1}^{n} f(x_i|\theta) \tag{1}$$

The log likelihood function is

$$\log L(\theta|x) \equiv \sum_{i=1}^{n} \log f(x_i|\theta) \tag{2}$$

* All comments welcome!

† E-mail: u626530i@ecs.osaka-u.ac.jp, Room 501

1.1 Definition of MLE

MLE($\hat{\theta}$) maximize the likelihood function. MLE satisfies the following condition

$$\frac{\partial \log L(\theta|x)}{\partial \theta} = 0 \tag{3}$$

$$\frac{\partial^2 \log L(\theta|x)}{\partial \theta \partial \theta'} \text{ is negative definite matrix.} \tag{4}$$

In other words,

$$\log L(\hat{\theta}) \geq \log L(\theta) \tag{5}$$

is satisfied for any θ.

1.2 Fisher's information matrix

Fisher's information matrix is defined as

$$I(\theta) \equiv -E\left[\frac{\partial^2 \log L(\theta|x)}{\partial \theta \partial \theta'}\right] \tag{6}$$

$$= V\left[\frac{\partial \log L(\theta|x)}{\partial \theta}\right] \tag{7}$$

$$\int L(\theta|x)dx = 1 \tag{8}$$

Take a derivation with respect to $\theta(k \times 1)$.

$$\begin{aligned}
&\frac{\partial}{\partial \theta}\int L(\theta|x)dx = 0 \\
\Longleftrightarrow &\int \frac{\partial}{\partial \theta} L(\theta|x)dx = 0 \\
\Longleftrightarrow &\int \frac{\partial \log L(\theta|x)}{\partial \theta} L(\theta|x)dx = 0 \\
\Longleftrightarrow &E[\frac{\partial \log L(\theta|x)}{\partial \theta}] = 0
\end{aligned} \tag{9}$$

Again, defferentiating (9) with respect to $\theta'(1 \times k)$.

$$\int \frac{\partial^2 \log L(\theta|x)}{\partial\theta\partial\theta'} L(\theta|x)dx + \int \frac{\partial \log L(\theta|x)}{\partial\theta} \frac{\partial \log L(\theta|x)}{\partial\theta'} L(\theta|x)dx = 0 \tag{10}$$

$$\Longrightarrow I(\theta) \equiv -E\left[\frac{\partial^2 \log L(\theta|x)}{\partial\theta\partial\theta'}\right] = V\left[\frac{\partial \log L(\theta|x)}{\partial\theta}\right] \tag{11}$$

1.3 Cramer-Rao lower bound

$s(X)$:unbiased estimator of θ ($E[s(X)] = \theta$)

$$V[s(X)] \geq I(\theta)^{-1} \tag{12}$$

Proof

For simplicity, let θ and $s(X)$ be scalar.

$$\begin{aligned}
E[s(X)] &= \int s(x)L(\theta|x)dx \\
\frac{\partial}{\partial\theta}E[s(X)] &= \int s(x)\frac{\partial \log L(\theta|x)}{\partial\theta}dx \\
&= E\left[s(X)\frac{\partial \log L(\theta|x)}{\partial\theta}\right] \\
&= cov\left(s(X), \frac{\partial \log L(\theta|x)}{\partial\theta}\right) \\
\Longleftrightarrow 1 &= cov\left(s(X), \frac{\partial \log L(\theta|x)}{\partial\theta}\right)
\end{aligned}$$

By the fact that

$$\begin{aligned}
-1 &\leq \frac{cov\left(s(X), \frac{\partial \log L(\theta|x)}{\partial\theta}\right)}{\sqrt{V[s(X)]}\sqrt{V[\frac{\partial \log L(\theta|x)}{\partial\theta}]}} \leq 1 \\
\Longleftrightarrow -1 &\leq \frac{1}{\sqrt{V[s(X)]}\sqrt{V[\frac{\partial \log L(\theta|x)}{\partial\theta}]}} \leq 1,
\end{aligned}$$

we can derive that

$$
\begin{aligned}
V[s(X)] &\geq V\left[\frac{\partial \log L(\theta|x)}{\partial\theta}\right]^{-1} \\
&= I(\theta)^{-1}
\end{aligned}
$$

1.4 Asymptotic distribution of MLE

$\hat{\theta}$: MLE , θ_0 :true value of parameter

$$
\sqrt{n}(\hat{\theta}-\theta) \to N\left(0, \lim_{n\to\infty}\left(\frac{1}{n}I(\theta)\right)^{-1}\right) \tag{13}
$$

Expanding $\frac{\partial \log L(\hat{\theta}|x)}{\partial\theta} = 0$ around $\hat{\theta} = \theta$,

$$
\begin{aligned}
0 = \frac{\partial \log L(\hat{\theta}|x)}{\partial\theta} &\approx \frac{\partial \log L(\theta|x)}{\partial\theta} + \frac{\partial^2 \log L(\theta|x)}{\partial\theta\partial\theta'}(\hat{\theta}-\theta) \\
\Longleftrightarrow \hat{\theta}-\theta &= \left(-\frac{\partial^2 \log L(\theta|x)}{\partial\theta\partial\theta'}\right)^{-1}\frac{\partial \log L(\theta|x)}{\partial\theta} \\
\Longleftrightarrow \sqrt{n}(\hat{\theta}-\theta) &= \left(-\frac{1}{n}\frac{\partial^2 \log L(\theta|x)}{\partial\theta\partial\theta'}\right)^{-1}\frac{1}{\sqrt{n}}\frac{\partial \log L(\theta|x)}{\partial\theta}
\end{aligned}
$$

From the result that

$$
\begin{aligned}
-\frac{1}{n}\frac{\partial^2 \log L(\theta|x)}{\partial\theta\partial\theta'} &\xrightarrow{p} \lim_{n\to\infty}\left\{-\frac{1}{n}E[\frac{\partial^2 \log L(\theta|x)}{\partial\theta\partial\theta'}]\right\} \\
&= \lim_{n\to\infty}\left\{\frac{1}{n}I(\theta)\right\} \\
\frac{1}{\sqrt{n}}\frac{\partial \log L(\theta|x)}{\partial\theta} &\xrightarrow{d} N(0, \lim_{n\to\infty}\frac{1}{n}V[\frac{\partial \log L(\theta|x)}{\partial\theta}]) \\
&= N(0, \lim_{n\to\infty}\left\{\frac{1}{n}I(\theta)\right\}),
\end{aligned}
$$

So we can find that

$$
\sqrt{n}(\hat{\theta}-\theta) \xrightarrow{d} N(0, \lim_{n\to\infty}\frac{1}{n}I(\theta))
$$

1.4.1 Review(Multivariate Lindeberg-Feller CLT)

$x_1, \cdots, x_n$:sample of random vectors such that $E[x_i] = \mu_i$,$V[x_i] = Q_i$

$$\sqrt{n}(\bar{x}_n - \bar{\mu}_n) \xrightarrow{d} N(0, Q) \tag{14}$$

where

$$\lim_{n\to\infty} \overline{Q}_n = Q, \tag{15}$$

if the conditions are satisfied

- for every i,

$$\lim_{n\to\infty} (n\overline{Q}_n)^{-1} Q_i = \lim_{n\to\infty} \left(\sum_{i=1}^{n} Q_i \right)^{-1} Q_i = 0 \tag{16}$$

- all mixed third moments of the multivariate distribution are finite