1. **Asymptotic Normality of MLE:**

Let $\tilde{\theta}$ be MLE of θ.

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

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

where it is assumed that $\lim_{n\to\infty}\left(\frac{I(\theta)}{n}\right)$ 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, $\mathrm{V}(s(X))$ is approximately equal to $(I(\theta))^{-1}$.

Practically, we utilize the following approximated distribution:

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

Then, we can obtain the significance test and the confidence interval for θ

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

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

   Then, the central limit theorem is given by:

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

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

That is,

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

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

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

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

Note that $\mathrm{E}(\overline{X}) = \mu$ and $n\mathrm{V}(\overline{X}) = \Sigma$.

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

Assume:

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

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

The central limit theorem is given by:

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

Note that $\mathrm{E}(\overline{X}) = \mu$ and $n\mathrm{V}(\overline{X}) \longrightarrow \sigma^2$.

In the case where X_i is a vector of random variable with mean μ and variance Σ_i, the central limit theorem is given by:

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

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

Note that $\mathrm{E}(\overline{X}) = \mu$ and $n\mathrm{V}(\overline{X}) \longrightarrow \Sigma$.

[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.

- **Convergence in Distribution (分布収束)** $X_n \longrightarrow X$, i.e., X converges in distribution to X. The distribution of X_n converges to the distribution of X as n goes to infinity.

Some Formulas

X_n and Y_n : Convergence in Probability

Z_n : Convergence in Distribution

- If $X_n \longrightarrow a$, then $f(X_n) \longrightarrow f(a)$.
- If $X_n \longrightarrow a$ and $Y_n \longrightarrow b$, then $f(X_n Y_n) \longrightarrow f(ab)$.
- If $X_n \longrightarrow a$ and $Z_n \longrightarrow Z$, then $X_n Z_n \longrightarrow aZ$, i.e., aZ is distributed with mean $\mathrm{E}(aZ) = a\mathrm{E}(Z)$ and variance $\mathrm{V}(aZ) = a^2\mathrm{V}(Z)$.

[End of Review]

4. **Asymptotic Normality of MLE — Proof:**

The density (or probability) function of X_i is given by $f(x_i;\theta)$.

The likelihood function is: $L(\theta;x) \equiv f(x;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$,

where $x = (x_1, x_2, \cdots, x_n)$.

MLE of θ results in the following maximization problem:

$$\max_{\theta} \ \log L(\theta;x).$$

A solution of the above problem is given by MLE of θ, denoted by $\tilde{\theta}$.

That is, $\tilde{\theta}$ is given by the θ which satisfies the following equation:

$$\frac{\partial \log L(\theta;x)}{\partial\theta} = \sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial\theta} = 0.$$

Replacing x_i by the underlying random variable X_i, $\dfrac{\partial \log f(X_i;\theta)}{\partial\theta}$ is taken as the ith random variable, i.e., X_i in the **Central Limit Theorem II**.

Consider applying **Central Limit Theorem II**.

In this case, we need the following expectation and variance:

$$\mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \log f(X_i;\theta)}{\partial\theta}\Big) \qquad \text{and} \qquad \mathrm{V}\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \log f(X_i;\theta)}{\partial\theta}\Big).$$

Defining the variance:

$$\mathrm{V}\Big(\frac{\partial \log f(X_i;\theta)}{\partial\theta}\Big) = \Sigma_i,$$

we can rewrite the information matrix as follows:

$$\begin{aligned} I(\theta) &= \mathrm{V}\Big(\frac{\partial \log L(\theta;X)}{\partial\theta}\Big) = \mathrm{V}\Big(\sum_{i=1}^{n}\frac{\partial \log f(X_i;\theta)}{\partial\theta}\Big) \\ &= \sum_{i=1}^{n}\mathrm{V}\Big(\frac{\partial \log f(X_i;\theta)}{\partial\theta}\Big) = \sum_{i=1}^{n}\Sigma_i \end{aligned}$$

The third equality holds when $X_1, X_2, \cdots, X_n$ are mutually independent.

Note that $\mathrm{E}\Big(\dfrac{\partial \log L(\theta; X)}{\partial \theta}\Big) = 0$ and $\mathrm{V}\Big(\dfrac{\partial \log L(\theta; X)}{\partial \theta}\Big) = I(\theta)$.

$$\frac{1}{n}\frac{\partial \log L(\theta; X)}{\partial \theta} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial \log f(X_i; \theta)}{\partial \theta}$$

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \log f(X_i; \theta)}{\partial \theta} - \mathrm{E}\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \log f(X_i; \theta)}{\partial \theta}\Big)\right) \longrightarrow N(0, \Sigma),$$

where

$$n\mathrm{V}\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \log f(X_i; \theta)}{\partial \theta}\Big) = \frac{1}{n}\mathrm{V}\Big(\sum_{i=1}^{n}\frac{\partial \log f(X_i; \theta)}{\partial \theta}\Big) = \frac{1}{n}\mathrm{V}\Big(\frac{\partial \log L(\theta; X)}{\partial \theta}\Big)$$

$$= \frac{1}{n}I(\theta) \longrightarrow \Sigma.$$

That is,

$$\frac{1}{\sqrt{n}}\frac{\partial \log L(\theta; X)}{\partial \theta} \longrightarrow N(0, \Sigma),$$

where $X = (X_1, X_2, \cdots, X_n)$.

Now, consider replacing θ by $\tilde{\theta}$, i.e.,

$$\frac{1}{\sqrt{n}}\frac{\partial \log L(\tilde{\theta}; X)}{\partial \theta},$$

which is expanded around $\tilde{\theta} = \theta$ as follows:

$$0 = \frac{1}{\sqrt{n}}\frac{\partial \log L(\tilde{\theta}; X)}{\partial \theta} \approx \frac{1}{\sqrt{n}}\frac{\partial \log L(\theta; X)}{\partial \theta} + \frac{1}{\sqrt{n}}\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}(\tilde{\theta} - \theta).$$

Therefore,

$$\frac{1}{\sqrt{n}}\frac{\partial^2 \log L(\theta; X)}{\partial \theta \partial \theta'}(\tilde{\theta} - \theta) \approx -\frac{1}{\sqrt{n}}\frac{\partial \log L(\theta; X)}{\partial \theta} \longrightarrow N(0, \Sigma).$$

The left-hand side is rewritten as:

$$\frac{1}{\sqrt{n}}\frac{\partial^2 \log L(\theta; X)}{\partial\theta\partial\theta'} = \sqrt{n}\frac{1}{n}\frac{\partial^2 \log L(\theta; X)}{\partial\theta\partial\theta'}(\tilde{\theta} - \theta).$$

Then,

$$\begin{aligned}\sqrt{n}(\tilde{\theta} - \theta) &\approx -\Big(\frac{1}{n}\frac{\partial^2 \log L(\theta; X)}{\partial\theta\partial\theta'}\Big)^{-1}\Big(\frac{1}{\sqrt{n}}\frac{\partial \log L(\theta; X)}{\partial\theta}\Big) \\ &\longrightarrow N(0, \Sigma^{-1}\Sigma\Sigma^{-1}) = N(0, \Sigma^{-1}).\end{aligned}$$

Note that

$$\frac{1}{n}\frac{\partial^2 \log L(\theta; X)}{\partial\theta\partial\theta'} \longrightarrow \lim_{n\to\infty}\frac{1}{n}\mathrm{E}\Big(\frac{\partial^2 \log L(\theta; X)}{\partial\theta\partial\theta'}\Big) = \Sigma,$$

and $\Big(\frac{1}{n}\frac{\partial^2 \log L(\theta; X)}{\partial\theta\partial\theta'}\Big)^{-1}\Big(\frac{1}{\sqrt{n}}\frac{\partial \log L(\theta; X)}{\partial\theta}\Big)$ has the same asymptotic distribution as $\Sigma^{-1}\Big(\frac{1}{\sqrt{n}}\frac{\partial \log L(\theta; X)}{\partial\theta}\Big)$.

Exam — Aug. 4, 2016 (AM8:50-10:20)

- 60 - 70% from two homeworks including optional an additional questions (2 つの宿題から 60 - 70%)
- 30 - 40% of new questions (30 - 40% の新しい問題)
- Questions are written in English, and answers should be in English or Japanese. (出題は英語，解答は英語または日本語)
- With no carrying in (持ち込みなし)