Lecture 16: Asymptotic normality of the MLE

Convergence of the MLE

Suppose that $Y_1, Y_2, Y_3, \ldots$ are iid with probability function $f(y \mid \theta)$, $\theta \in \mathbb{R}^d$

Let $\ell_n(\theta) = \sum_{i=1}^{n} \log f(Y_i \mid \theta)$, and $\hat{\theta}_n$ be the MLE of θ using first n observations. Let

$$I_1(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \log f(Y_i \mid \theta)\right]$$

<u>Theorem</u>: Under regularity conditions,

(a) $\hat{\theta}_n \xrightarrow{p} \theta$ as $n \to \infty$ (consistency)

(b) $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I_1^{-1}(\theta))$ as $n \to \infty$

(asymptotic normality)

Intermediate steps

$\ell_n(\theta) = \sum_{i=1} \log f(Y_i \mid \theta)$

Using the WLLN and the CLT, argue that:

CLT: $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \xrightarrow{d} N(0,1)$

② • $\frac{1}{n}\ell''_n(\theta) \xrightarrow{p} -I_1(\theta)$ ✓ (last time)

③ • $\frac{1}{\sqrt{n}}\ell'_n(\theta) \xrightarrow{d} N(0, I_1(\theta))$

pf ③: $\ell'_n(\theta) = \sum_{i=1} \frac{\partial}{\partial\theta} \log f(Y_i \mid \theta)$

(CLT)

$$\sqrt{n}\left(\frac{1}{n}\ell'_n(\theta) - \underbrace{E\left[\frac{\partial}{\partial\theta}\log f(Y_i \mid \theta)\right]}_{0 \text{ (under regularity conditions)}}\right) \xrightarrow{d} N\Big(0, \underbrace{Var\left(\frac{\partial}{\partial\theta}\log f(Y_i \mid \theta)\right)}_{I_1(\theta)}\Big)$$

$I_1(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log f(Y_i \mid \theta)\right]$ (under regularity conditions)

$$\sqrt{n}\cdot\frac{1}{n}\ell'_n(\theta) \xrightarrow{d} N(0, I_1(\theta))$$

$$\frac{1}{\sqrt{n}}\ell'_n(\theta) \xrightarrow{d} N(0, I_1(\theta)) \quad //$$

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad Var(X_i) = \sigma^2$$

$$Var(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i) = \frac{\sigma^2}{n}$$

$$\to 0 \quad \text{as } n \to \infty$$

($\Rightarrow$ convergence in probability)

$$\frac{\bar{X}_n - \mu}{\sigma} \xrightarrow{P} 0$$

$$Var(\sqrt{n}\,\bar{X}_n) = (\sqrt{n})^2 Var(\bar{X}_n) = n \cdot \frac{\sigma^2}{n} = \sigma^2$$

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$$

$$\mathbb{E}[\text{"} \quad \text{"}] = 0$$

$$Var(\text{"} \quad \text{"}) = 1$$

Putting everything together

Full proof:

① If $\hat{\theta}_n$ is the MLE $\Rightarrow$ $l_n'(\hat{\theta}_n) = 0$

Also know that $\hat{\theta}_n \xrightarrow{P} \theta$ (by part (a) of the Theorem)

$\Rightarrow$ for large n, $\hat{\theta}_n \approx \theta$

$$0 = l_n'(\hat{\theta}_n) \approx l_n'(\theta) + (\hat{\theta}_n - \theta)\, l_n''(\theta)$$

$$\Rightarrow \quad \hat{\theta}_n - \theta \approx \frac{l_n'(\theta)}{-l_n''(\theta)}$$

$$\Rightarrow \quad \sqrt{n}(\hat{\theta}_n - \theta) \approx \frac{\sqrt{n}\, l_n'(\theta)}{-l_n''(\theta)} = \frac{\frac{1}{\sqrt{n}} l_n'(\theta)}{-\frac{1}{n} l_n''(\theta)}$$

$$\frac{1}{\sqrt{n}} l_n'(\theta) \xrightarrow{d} N(0, I_1(\theta))$$

$$-\frac{1}{n} l_n''(\theta) \xrightarrow{P} I_1(\theta)$$

(Slutsky's) $\Rightarrow$

$$\frac{\frac{1}{\sqrt{n}} l_n'(\theta)}{-\frac{1}{n} l_n''(\theta)} \xrightarrow{d} \frac{1}{I_1(\theta)} N(0, I_1(\theta)) = N(0, I_1^{-1}(\theta))$$

//

Regularity conditions

Some sufficient regularity conditions:

- The dimension of θ does not change with n
- $f(y \mid \theta)$ is a sufficiently smooth function of θ
- we can swap integration and differentiation
- θ is not on the boundary of the parameter space
- θ is identifiable (basically, $f(y \mid \theta_1) \neq f(y \mid \theta_2)$ if $\theta_1 \neq \theta_2$)

e.g., for Bernoulli(θ)
θ cannot be 0 or 1

for $N(\mu, \sigma^2)$
σ^2 cannot be 0, etc.

Counterexample: $X_1, X_2, X_3, \ldots \overset{iid}{\sim} \text{Uniform}(0, \theta)$

$$\hat{\theta}_n = X_{(n)} \qquad \hat{\theta}_n \xrightarrow{p} \theta$$

(previous class activity)

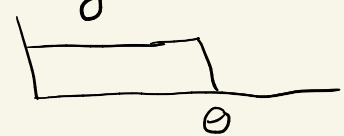
$$n(\hat{\theta}_n - \theta) \xrightarrow{d} -\text{Exponential}\left(\frac{1}{\theta}\right)$$

$$\sqrt{n}(\hat{\theta}_n - \theta) = \underbrace{\frac{1}{\sqrt{n}}}_{\to 0} \cdot \underbrace{n(\hat{\theta}_n - \theta)}_{\xrightarrow{d} -\text{Exp}\left(\frac{1}{\theta}\right)} \xrightarrow{p} 0 \quad \text{(Slutsky's)}$$

not to a Normal

Regularity conditions violated:

- θ is on the boundary, so $f(x|\theta)$ is not smooth at θ



$\Rightarrow$ derivatives are not defined at θ

$\Rightarrow \mathcal{I}_1(\theta)$ is not even defined

- can't exchange integration & differentiation (domain of integration depends on θ)

Counterexample $X_1, X_2, X_3, \ldots \overset{iid}{\sim}$ Bernoulli(p)

$p = 0$ or $p = 1$

$\downarrow$ $\downarrow$

$\hat{p} = \bar{X} \equiv 0$ $\hat{p} = \bar{X} \equiv 1$

$\sqrt{n}(\bar{X} - p) = 0 \overset{d}{\not\to}$ Normal $\sqrt{n}(\bar{X} - p) = 0 \overset{d}{\not\to}$ Normal

regularity condition violated: p is on the boundary of parameter space

More generally: can't converge in dist. if $\mathrm{Var}(\hat{\theta}_n) \equiv 0$