Lecture 15: Asymptotic normality of the MLE

Some more theorems about convergence

Continuous mapping theorem: Let $X_1, X_2, \ldots$ be a sequence of random variables, and let g be a continuous function,

① If $X_n \xrightarrow{d} X$ then $g(X_n) \xrightarrow{d} g(X)$

② If $X_n \xrightarrow{p} X$ then $g(X_n) \xrightarrow{p} g(X)$ ← HW

③ If $X_n \xrightarrow{a.s.} X$ then $g(X_n) \xrightarrow{a.s.} g(X)$

Slutsky's theorem: Let $\{X_n\}$ and $\{Y_n\}$ be sequences of random variables, and suppose that $X_n \xrightarrow{d} X$, and $Y_n \xrightarrow{p} c$ (c is a constant). Then

- $X_n + Y_n \xrightarrow{d} X + c$
- $X_n Y_n \xrightarrow{d} X c$
- $\dfrac{X_n}{Y_n} \xrightarrow{d} \dfrac{X}{c}$ (provided c is invertible)

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

$$\mathcal{I}_1(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \log f(Y_i \mid \theta)\right]$$

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

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

(b) $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \mathcal{I}_1^{-1}(\theta))$ as $n \to \infty$

(asymptotic normality)

We will prove (b) when $d = 1$

Proof outline $(d=1)$:

① $\sqrt{n}(\hat{\theta}_n - \theta) \approx \dfrac{\frac{1}{\sqrt{n}} \ell_n'(\theta)}{-\frac{1}{n} \ell_n''(\theta)}$ (Taylor expansion)

② $\frac{1}{n} \ell_n''(\theta) \xrightarrow{P} -\mathcal{I}_1(\theta)$ (WLLN)

③ $\frac{1}{\sqrt{n}} \ell_n'(\theta) \xrightarrow{d} N(0, \mathcal{I}_1(\theta))$ (CLT)

④ Apply Slutsky's:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \frac{1}{\mathcal{I}_1(\theta)} N(0, \mathcal{I}_1(\theta))$$

$$= N(0, \mathcal{I}_1^{-1}(\theta))$$

$X \sim N(\mu, \sigma^2)$ $\quad aX \sim N(a\mu, a^2\sigma^2)$

$a \in \mathbb{R}$

CLT: $\dfrac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$ $\quad \bar{X}_n =$ mean of n iid observations

Intermediate steps

Using the WLLN and the CLT, argue that:

② • $\frac{1}{n}\ell_n''(\theta) \xrightarrow{p} -I_1(\theta)$

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

Pf of ②:

$$\ell_n(\theta) = \sum_{i=1}^{n} \log f(Y_i|\theta)$$

$$\frac{1}{n}\ell_n''(\theta) = \underbrace{\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \log f(Y_i|\theta)}_{\text{sample mean of } \frac{\partial^2}{\partial\theta^2} \log f(Y_i|\theta)}$$

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

$\Rightarrow$ WLLN: $\frac{1}{n}\ell_n''(\theta) \xrightarrow{p} -I_1(\theta)$

(or $-\frac{1}{n}\ell_n''(\theta) \xrightarrow{p} I_1(\theta)$) //