Lecture 14: Asymptotic properties of the MLE

Ciaran Evans

Recap: Convergence in probability

Definition: A sequence of random variables $X_1, X_2, \dots$ *converges in probability* to a random variable X if, for every $\varepsilon > 0$,

$$\lim_{n \to \infty} P(|X_n - X| \geq \varepsilon) = 0$$

We write $X_n \xrightarrow{p} X$.

Convergence in distribution

Definition: A sequence of random variables $X_1, X_2, \ldots$ *converges in distribution* to a random variable X if

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

at all points where $F_X(x)$ is continuous. We write $X_n \xrightarrow{d} X$.

Convergence of the MLE

Suppose that we observe $Y_1, Y_2, Y_3, \ldots$ iid from a distribution with probability function $f(y|\theta)$, where $\theta \in \mathbb{R}^d$ is the parameter(s) we are trying to estimate. Let

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

$$\widehat{\theta}_n = \text{argmax}_\theta \ \ell_n(\theta)$$

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

Theorem: Under certain regularity conditions (to be discussed later),

(a) $\widehat{\theta}_n \xrightarrow{p} \theta$

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

Asymptotic normality: proof approach

Let $\ell_n'(\theta) = \frac{\partial}{\partial\theta}\ell_n(\theta)$, $\ell_n''(\theta) = \frac{\partial^2}{\partial\theta^2}\ell_n(\theta)$

Begin with a Taylor expansion of ℓ_n' around θ:

$$\ell_n'(\widehat{\theta}_n) =$$

Asymptotic normality: proof approach

Using the Taylor expansion,

$$\sqrt{n}(\widehat{\theta}_n - \theta) \approx \frac{\frac{1}{\sqrt{n}}\ell'_n(\theta)}{-\frac{1}{n}\ell''_n(\theta)}$$

Next, look at limits for the numerator and denominator:

- $\frac{1}{\sqrt{n}}\ell'_n(\theta)$
- $-\frac{1}{n}\ell''_n(\theta)$

Asymptotic normality: the numerator

Want to show: $\frac{1}{\sqrt{n}}\ell'_n(\theta) \xrightarrow{d} N(0, \mathcal{I}_1(\theta))$

- CLT: for iid $X_1, X_2, \ldots$, under mild conditions
  $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X_i]\right) \xrightarrow{d} N(0, Var(X_i))$
- $\ell'_n(\theta) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(Y_i|\theta)$

Applying CLT to $\ell'_n(\theta)$:

Asymptotic normality: the numerator

Want to show: $\frac{1}{\sqrt{n}}\ell_n'(\theta) \xrightarrow{d} N(0, \mathcal{I}_1(\theta))$

CLT gives

$\sqrt{n}\left(\frac{1}{n}\ell_n'(\theta) - \mathbb{E}\left[\frac{\partial}{\partial\theta}\log f(Y_i|\theta)\right]\right) \xrightarrow{d} N\left(0, Var\left(\frac{\partial}{\partial\theta}\log f(Y_i|\theta)\right)\right)$

Need to show:

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

The expected score

Claim: Under regularity conditions,

$$\mathbb{E}\left[\frac{\partial}{\partial\theta}\log f(Y_i|\theta)\right] = 0$$

Fisher information

Claim: Under regularity conditions,

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

Numerator: putting everything together

Want to show: $\frac{1}{\sqrt{n}}\ell'_n(\theta) \xrightarrow{d} N(0, \mathcal{I}_1(\theta))$

- CLT gives

$$\sqrt{n}\left(\frac{1}{n}\ell'_n(\theta) - \mathbb{E}\left[\frac{\partial}{\partial\theta}\log f(Y_i|\theta)\right]\right) \xrightarrow{d} N\left(0, Var\left(\frac{\partial}{\partial\theta}\log f(Y_i|\theta)\right)\right)$$

- Under regularity conditions,

$$\mathbb{E}\left[\frac{\partial}{\partial\theta}\log f(Y_i|\theta)\right] = 0$$

- Under regularity conditions,

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

Now the denominator

Want to show: $-\frac{1}{n}\ell_n''(\theta) \xrightarrow{p} \mathcal{I}_1(\theta)$

Question: What big theorem do we have for convergence in probability?

The denominator: WLLN

Want to show: $-\frac{1}{n}\ell_n''(\theta) \xrightarrow{p} \mathcal{I}_1(\theta)$

- WLLN: For iid $X_1, X_2, ...,$ under mild conditions $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mathbb{E}[X_i]$
- $-\frac{1}{n}\ell_n''(\theta) = \frac{1}{n}\sum_{i=1}^{n} -\frac{\partial^2}{\partial\theta^2}\log f(Y_i|\theta)$

Applying WLLN to $-\frac{1}{n}\ell_n''(\theta)$: