STAT 3100 Lecture Note

Week Eleven (Nov 22 & 24, 2022)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2022/Nov/25 19:36:12

Asymptotic properties of MLE (con'd)

Consistency (or consistence, CB Sec 10.1.1)

- $T_n = T_n(X_1, \dots, X_n)$ is consistent for θ iff $T_n \xrightarrow{p} \theta$ as $n \to \infty$
  - A sufficient condition for consistency: $\mathrm{E}(T_n \mid \theta) \to \theta$ and $\mathrm{var}(T_n \mid \theta) \to 0$ as $n \to \infty$

CB Example 5.5.3

- Suppose that iid $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$. Prove that
  - $S_n^2 = (n-1)^{-1} \sum_i (X_i - \bar{X}_n)^2$ is consistent for σ^2;
  - $\widehat{\sigma^2}_{\mathrm{ML}} = n^{-1} \sum_i (X_i - \bar{X}_n)^2$ is consistent for σ^2 too.

to θ; - Violated by, e.g., $\mathrm{Unif}(0, \theta)$; - θ_0 is an interior point of parameter space Θ.

Example of inconsistent MLE

There are independent $X_{i1}, X_{i2} \sim \mathcal{N}(\mu_i, \sigma^2)$, $i = 1, \dots, n$. Then $\widehat{\sigma^2}_{\mathrm{ML}}$ is NOT consistent for σ^2.

Examples of consistent MLE with the regularity conditions violated

- iid $X_1, \dots, X_n \sim \mathrm{Ber}(1)$
- iid $X_1, \dots, X_n \sim \mathrm{Unif}(0, \theta)$

Efficiency

- (HMC Def 6.2.2) For an estimator, say T_n, unbiased for $\tau(\theta)$, the (finite-sample) *efficiency* of T_n is the ratio of the CRLB to $\mathrm{var}(T_n)$, i.e., $[\{\tau'(\theta)\}^2 / I_n(\theta)] / \mathrm{var}(T_n \mid \theta)$.
  - The higher efficiency the better;
  - the efficiency $= 1 \iff$ an efficient estimator.
- (CB Def 10.1.9) If $k_n\{T_n - \tau(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma^2)$, then σ^2 is the *asymptotic variance* of T_n.
- (CB Def 10.1.11) T_n is *asymptotically efficient* for $\tau(\theta) \iff \sqrt{n}\{T_n - \tau(\theta)\} \xrightarrow{d} \mathcal{N}(0, \{\tau'(\theta)\}^2 / I_1(\theta))$, where

  $$I_1(\theta) = -\mathrm{E}\left\{\frac{\partial^2}{\partial \theta^2} \ln f(X_i \mid \theta) \mid \theta\right\} = -\mathrm{E}\{H(\theta; X_i) \mid \theta\} \text{ is the Fisher information of one single observation.}$$

  - i.e., the asymptotic variance of T_n is $\{\tau'(\theta)\}^2 / I_1(\theta)$, attaining the CRLB

- (CB Def 10.1.16 & HMC Def 6.2.3(c)) Denote by T_n and W_n two estimators for $\tau(\theta)$. Suppose that $\sqrt{n}\{T_n - \tau(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma_T^2)$ and $\sqrt{n}\{W_n - \tau(\theta)\} \xrightarrow{d} \mathcal{N}(0, \sigma_W^2)$. The *asymptotic relative efficiency* (ARE) of T_n with respect to W_n is defined as

$$\mathrm{ARE}(T_n, W_n) = \sigma_W^2 / \sigma_T^2.$$

  - T_n is asymptotically more efficient than $W_n \iff \mathrm{ARE}(T_n, W_n) > 1$
  - T_n is asymptotically efficient $\iff \{\tau'(\theta)\}^2 / \{I_1(\theta)\sigma_T^2\} = 1$

Asymptotic efficiency of MLE (CB Thm 10.1.12 & Ex. 10.7)

- $\sqrt{n}(\hat{\theta}_{\mathrm{ML}} - \theta_0) \xrightarrow{d} \mathcal{N}(0, 1/I_1(\theta_0))$, provided that $\hat{\theta}_{\mathrm{ML}}$ is the MLE for θ_0, we have the previous four regularity conditions (for the consistency of MLE) plus the following two more (CB Sec 10.6.2):
  - For each $x \in \mathrm{supp}(X)$, $f(x \mid \theta)$ is three time continuously differentiable with respect to θ; and $\int f(x \mid \theta) \mathrm{d}x$ can be differentiated three times under the integral sign;
  - for each $\theta \in \Theta$, there exists $c(\theta) > 0$ and $M(x, \theta)$ such that $|\frac{\partial^3}{\partial \theta^3} \ln f_X(x \mid \theta)| \le M(x, \theta)$ for all $x \in \mathrm{supp}(X)$ and $\theta \in (\theta - c(\theta), \theta + c(\theta))$.
- In practice,
  - $nI_1(\theta_0) = I_n(\theta_0) \approx I_n(\hat{\theta}_{\mathrm{ML}}) \approx \hat{I}_n(\hat{\theta}_{\mathrm{ML}})$
    * (Expected) Fisher information (number) $I_n(\theta_0) = -\mathrm{E}\{H(\theta_0; \mathbf{X})\}$
    * Observed Fisher information (number) $\hat{I}_n(\hat{\theta}_{\mathrm{ML}}) = -\frac{\partial^2}{\partial \theta^2} \ln L(\theta; \boldsymbol{x})\big|_{\theta = \hat{\theta}_{\mathrm{ML}}} = -H(\hat{\theta}_{\mathrm{ML}}; \boldsymbol{x})$
  - Hence $\mathrm{var}(\hat{\theta}_{\mathrm{ML}}) \approx 1/I_n(\theta_0) \approx 1/I_n(\hat{\theta}_{\mathrm{ML}}) \approx 1/\hat{I}_n(\hat{\theta}_{\mathrm{ML}})$

Take-home exercises (NOT to be submitted; to be potentially covered in labs)

- CB Ex. 10.3, 10.17(a–c)