PH 712 Probability and Statistical Inference

Part VII: Evaluating Estimators II

Zhiyang Zhou (zhou67@uwm.edu, zhiyanggeezhou.github.io)

2024/10/26 17:12:24

Consistency (or consistence, CB Sec 10.1.1)

- A statistic T_n is consistent for $g(\theta)$ if and only if $T_n \approx g(\theta)$ as $n \to \infty$
  - The "$\approx$" notation is abused here and, rigorously speaking, is supposed to be "$\xrightarrow{p}$" (convergence in probability);
  - $T_n \xrightarrow{p} g(\theta) \Leftrightarrow$ for each $\varepsilon > 0$, $\lim_{n\to\infty} \Pr(|T_n - g(\theta)| > \varepsilon) = 0$;
  - A sufficient condition for $T_n \xrightarrow{p} g(\theta)$: as $n \to \infty$, $\lim_{n\to\infty} \mathrm{E}(T_n) = g(\theta)$ and $\lim_{n\to\infty} \mathrm{var}(T_n) = 0$.
- (CB Example 5.5.3) Suppose that $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Then
  - $S^2 = (n-1)^{-1} \sum_i (X_i - \bar{X})^2$ is consistent for σ^2;
  - $\widehat{\sigma^2}_{\mathrm{ML}} = n^{-1} \sum_i (X_i - \bar{X})^2$ is consistent for σ^2 too.

Asymptotic efficiency

- (CB Def 10.1.11) T_n is *asymptotically efficient* for $g(\theta)$ if and only if $\sqrt{n}\{T_n - g(\theta)\} \approx \mathcal{N}(0, \{g'(\theta)\}^2/I_1(\theta))$
  - Where $I_1(\theta)$ is the Fisher information with $n = 1$
    - For an iid sample, $I_1(\theta) = n^{-1} I_n(\theta)$, no longer a function of n
  - The "$\approx$" notation is abused here and, rigorously speaking, is supposed to be "$\xrightarrow{d}$" (convergence in distribution);
  - $\sqrt{n}\{T_n - g(\theta)\} \xrightarrow{d} \mathcal{N}(0, I_1^{-1}(\theta)\{g'(\theta)\}^2)$ means that the limiting distribution of $\sqrt{n}\{T_n - g(\theta)\}$ is $\mathcal{N}(0, I_1^{-1}(\theta)\{g'(\theta)\}^2)$
- (CB Def 10.1.16 & HMC Def 6.2.3(c)) Denote by T_n and W_n two estimators for $g(\theta)$. Suppose that $\sqrt{n}\{T_n - g(\theta)\} \approx \mathcal{N}(0, \sigma_T^2)$ and $\sqrt{n}\{W_n - g(\theta)\} \approx \mathcal{N}(0, \sigma_W^2)$. 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 if and only if $\mathrm{ARE}(T_n, W_n) > 1$