Chapter 10: Asymptotic Evaluation

Tiandong Wang

Department of Statistics
Texas A&M University

April 12, 2021

Asymptotic Evaluation

Samples $X_1, \cdots, X_n$ i.i.d. $f(x|\theta)$, n large. We will see what happens if $n \to \infty$

- This assumption $n \to \infty$ generally makes life easier.
- Because limit theorems become available, distributions can be found approximately. Limiting distributions are much simpler than actual distributions

The behaviors (including its distribution, bias, variance) of an estimator under $n \to \infty$ are known as its *asymptotic* properties.

Consistency

Does the estimator converge/get closer to the parameter when n gets larger and larger?

Definition

Let $T_n = T_n(X_1, \cdots, X_n)$ be a sequence of estimators for $\tau(\theta)$. We say that T_n is **consistent** (weakly) for estimating $\tau(\theta)$ if

$$T_n \xrightarrow{p} \tau(\theta) \quad \text{under } P_\theta, \forall\theta$$

That is, given any $\epsilon > 0$, $\lim_{n\to\infty} P(|T_n - \tau(\theta)| > \epsilon) = 0$

Strongly consistent:

$$T_n \xrightarrow{a.s.} \tau(\theta) \quad \text{under } P_\theta, \forall\theta$$

Consistency

How to prove consistency of $\hat{\theta}$ for estimating θ?

- by definition (often complicated)
- Chebychev's Inequality

$$P(|T_n - \tau(\theta)| > \epsilon) \leq \frac{E[T_n - \tau(\theta)]^2}{\epsilon^2}$$

Consistent if $E[T_n - \tau(\theta)]^2 \to 0$, as $n \to \infty$.

Consistency

Theorem

If T_n is a sequence of estimators of $\tau(\theta)$ satisfying
$\lim_{n\to\infty} Bias_\theta(T_n) = 0$ *(asymptotically unbiased),*
$\lim_{n\to\infty} Var_\theta(T_n) = 0,$
for all θ, then T_n is consistent for $\tau(\theta)$.

-

Theorem

Let T_n be a consistent sequence of estimators of $\tau(\theta)$. Let a_n and b_n be a sequence constant satisfying $a_n \to 1$, $b_n \to 0$. Then the sequence $U_n = a_n T_n + b_n$ is a consistent estimator of $\tau(\theta)$

-

Convergence of Transformations:

Assume $X_n \xrightarrow{p} X$, $Y_n \xrightarrow{p} Y$, then

- $aX_n + bY_n \xrightarrow{p} aX + bY, X_n Y_n \xrightarrow{p} XY$
- $X_n / Y_n \xrightarrow{p} X/Y$ if $P(Y = 0) = 0$.
- Assume g is a continuous function. Then $g(X_n) \xrightarrow{p} g(X)$
- Assume h is a continuous function. Then $h(X_n, Y_n) \xrightarrow{p} h(X, Y)$

Invariance Principle of Consistency

- If T_n is consistent for θ and g is a continuous function, then $g(T_n)$ is consistent for $g(\theta)$.
- MME (method of moment estimator) is generally consistent.
- MLE (maximum likelihood estimator) is consistent
  Let $X_1, \cdots, X_n$ be i.i.d. $f(x|\theta)$ and let $\hat{\theta}_n$ be the MLE of θ. Then $\hat{\theta}_n$ is consistent for θ
- UMVUE (unbiased minimal variance estimator) is consistent.
  Let $X_1, \cdots, X_n$ be i.i.d. $f(x|\theta)$ and let T_n be the UMVUE of $\tau(\theta)$. Then T_n is consistent for $\tau(\theta)$

Consistency: Examples

Let $X_1, \ldots, X_n$ be i.i.d. Poisson(λ) observations. Is either the UMVUE or MLE of $e^{-\lambda}$ consistent?

Asymptotic normality

A statistic T_n is *asymptotically normal* if

$$\sqrt{n}\{T_n - \tau(\theta)\} \xrightarrow{d} N(0, v(\theta)), \forall\theta$$

- $\tau(\theta)$ is called the asymptotic mean;
- $v(\theta)$ is called the asymptotic variance

We write T_n as $AN(\tau(\theta), v(\theta)/n)$.

Remark: Asymptotic normality implies consistency! Why?

Asymptotic normality: Central Limit Theorem

Theorem

Assume $X_1, \cdots, X_n$ is iid $f(x|\theta)$, with finite mean $\mu = \mu(\theta)$ and variance $\sigma^2 = \sigma^2(\theta)$. Then

$$\sqrt{n}(\bar{X}_n - \mu(\theta)) \xrightarrow{d} N(0, \sigma^2(\theta)).$$

Asymptotic normality: Delta method

Assume $\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, v(\theta))$. If a function g satisfies that $g'(\theta) \neq 0$, then

$$\sqrt{n}(g(T_n) - g(\theta)) \xrightarrow{d} N\Big(0, [g'(\theta)]^2 v(\theta)\Big)$$

Examples:

- Let $X_1, \cdots, X_n$ be iid $N(\mu, \sigma^2)$, $\mu \neq 0$. Show the MLE of μ^2 is AN.
- Let $X_1, \cdots, X_n$ be iid from Bernoulli(p), where $p \neq 1/2$. Find the asymptotic mean and variance of $\bar{X}(1 - \bar{X})$.

Asymptotic Relative Efficiency (ARE)

If two estimators are both asymptotically unbiased and normal. Which is better? Can compare asymptotic variances.

Definition

If two estimators T_n and S_n satisfy

$$\sqrt{n}\{T_n - \tau(\theta)\} \xrightarrow{d} N(0, \sigma_T^2)$$

$$\sqrt{n}\{S_n - \tau(\theta)\} \xrightarrow{d} N(0, \sigma_S^2)$$

The asymptotic relative efficiency (ARE) of T_n with respect to S_n is $ARE(T_n, S_n) = \frac{\sigma_S^2}{\sigma_T^2}$

If $ARE(T_n, S_n) \leq 1, \forall\theta$, then S_n is asymptotically more efficient than T_n.

Asymptotic Relative Efficiency (ARE): Example

Example: $X_1, \cdots, X_n$ iid Poisson(λ). Consider two estimator of $P_\lambda(X = 0) = e^{-\lambda}$. One estimator is $Y_n = \dfrac{1}{n} \sum_{i=1}^{n} I(X_i = 0)$, and the other is the MLE.

Asymptotic Efficiency

Definition

Definition: A sequence T_n is asymptotically efficient for $\tau(\theta)$ if for all $\theta \in \Theta$.

$$\sqrt{n}(T_n - \tau(\theta)) \xrightarrow{d} N(0, \frac{[\tau'(\theta)]^2}{I(\theta)})$$

The asymptotic variance of T_n achieves the Cramér-Rao lower bound.

Asymptotic Efficiency of MLEs

Let $X_1, \cdots, X_n$ be iid $f(x|\theta)$, and let $\hat{\theta}$ be the MLE of θ. Under some regularity conditions, $\hat{\theta}$ is $AN(\theta, 1/(nI(\theta)))$ or $AN(\theta, 1/(I_n(\theta)))$, i.e.

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

for all $\theta \in \Theta$. Assume $\tau(\theta)$ is continuous and differentiable in θ, then

$$\sqrt{n}(\tau(\hat{\theta}) - \tau(\theta)) \xrightarrow{d} N(0, \frac{[\tau'(\theta)]^2}{I(\theta)}),$$

That is, $\tau(\hat{\theta})$ is a consistent and asymptotic efficient estimator of $\tau(\theta)$.

Approximating the variance

For finite sample size n, let $\hat{\theta}$ be the MLE. The variance of $\tau(\hat{\theta})$ can be approximated as

$$
\begin{aligned}
Var(\tau(\hat{\theta})) &\approx \frac{[\tau'(\theta)]^2}{I(\theta)} \quad \text{(asymptotic variance)} \\
&\approx \frac{[\tau'(\hat{\theta})]^2}{I(\hat{\theta})} \quad \text{(asymptotic variance)}
\end{aligned}
$$

$$
\begin{aligned}
Var(\tau(\hat{\theta})) &\approx \frac{[\tau'(\theta)]^2}{E_\theta(-\frac{\partial^2}{\partial\theta^2}\log(L(\theta|\mathbf{X})))} \\
&\approx \frac{[\tau'(\hat{\theta})]^2}{(-\frac{\partial^2}{\partial\theta^2}\log(L(\theta|\mathbf{X})))|_{\theta=\hat{\theta}}} \quad \text{(asymptotic variance)}
\end{aligned}
$$

The denominator is called observed information number.

Approximating the variance: Example

Example: (Approximate Binomial Variance)
$X_1, \cdots, X_n$ iid from $Bin(1, p)$.

- (1) Calculate the variance of the MLE of p.
- (2) Calculate the variance of the MLE of the odds $p/(1-p)$.