Lecture 18

2024 Spring

July 16, 2024

Last Lecture

- (Cont') Central Limit Theorem
- Continuous Mapping Theorem & Slutsky's Theorem & Delta Method

$X_n \xrightarrow{P} X$, $g(X_n) \xrightarrow{P} g(X)$.

$X_n \xrightarrow{d} X$. $g(X_n) \xrightarrow{d} g(X)$

$g(\cdot)$ is continuous.

$\sqrt{n}(Y_n - \theta) \xrightarrow{d} Norm(0. \sigma^2)$

$\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{d} Norm(0, \sigma^2 \cdot \{g'(\theta)\}^2)$

This Lecture

- Finish the Delta method (Finish first half of Chapter 4)
- Estimation (Start the second half of Chapter 4)

(Cont') Delta Method (optional)

Big $O_p(\cdot)$ Small $o_p(\cdot)$ Notations: If X_n is a sequence of random variables, and a_n is a sequence of constants.

1. We say $X_n = o_p(a_n)$ if
$$\frac{X_n}{a_n} \xrightarrow{p} 0.$$
2. We say $X_n = O_p(a_n)$ if X_n/a_n is *stochastically bounded*: for any $\epsilon > 0$, there exists a finite constant M_ϵ and a finite constant N_ϵ such that
$$\Pr\left(\left|\frac{X_n}{a_n}\right| > M_\epsilon\right) < \epsilon.$$
for any $n > N_\epsilon$.

For example, the sample mean of an iid sample from a distribution with variance σ^2 and mean μ satisfies
$$\bar{X}_n - \mu \xrightarrow{p} 0,$$ By WLLN.

which implies that

$\bar{X}_n - \mu = O_p(1)$ $\quad \frac{\bar{X}_n - \mu}{1} = \bar{X}_n - \mu \xrightarrow{p} 0$ $\quad a_n = 1$

It also satisfies
$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \text{Norm}(0, \sigma^2)$$ By CLT.

which implies that

$\bar{X}_n - \mu = O_p\left(\frac{1}{\sqrt{n}}\right)$ $\quad \frac{\bar{X}_n - \mu}{1/\sqrt{n}} = \sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \text{Norm}(0, \sigma^2)$

satisfies stochastically bounded

Heuristic Explanation: We provide a heuristic explanation for the Delta method.

1. If we know Y_n satisfies the following

$$\sqrt{n}(Y_n - \theta) \xrightarrow{d} \text{Norm}(0, \sigma^2),$$

then we have

$$Y_n - \theta = O_p\left(\frac{1}{\sqrt{n}}\right)$$

2. Perform a Taylor series expansion of $g(Y_n)$ around θ gives

$$\frac{\text{Const}\cdot(Y_n-\theta)^3}{Y_n} \propto n\cdot(Y_n-\theta)^3 = \boxed{\{\sqrt{n}(Y_n-\theta)\}^2(Y_n-\theta)}$$

$\{\sqrt{n}(Y_n-\theta)\}^2 \xrightarrow{d}$ Scaled χ^2_1, $(Y_n-\theta) \xrightarrow{P} 0$

$\xrightarrow{P} 0 \Rightarrow o_p(\frac{1}{n})$

$$g(Y_n) = g(\theta) + g'(\theta)\underbrace{(Y_n - \theta)}_{O_p(\frac{1}{\sqrt{n}})} + \frac{1}{2}g''(\theta)\underbrace{(Y_n - \theta)^2}_{O_p(\frac{1}{n})} + o_p\left(\frac{1}{n}\right)$$

3. Thus, we have

$$\sqrt{n}\left\{g(Y_n) - g(\theta)\right\} = \underbrace{g'(\theta)\sqrt{n}(Y_n - \theta)}_{\xrightarrow{d} \text{Norm}(0, \{g'(\theta)\}^2\sigma^2)} + \frac{\sqrt{n}}{2}g''(\theta)(Y_n - \theta)^2 + \underbrace{o_p\left(\frac{1}{\sqrt{n}}\right)}_{\xrightarrow{P} 0}$$

$$\frac{g''(\theta)}{2}\cdot\underbrace{\sqrt{n}(Y_n-\theta)}_{\xrightarrow{d}\text{Norm}(0,\sigma^2)}\cdot\underbrace{(Y_n-\theta)}_{\xrightarrow{P}0} \xrightarrow{P} 0$$

Examples: ① $X_1 \cdots X_n \overset{iid}{\sim} Poi(\mu)$. Find the limiting distribution of $Z_n = \sqrt{n}(\sqrt{\bar{X}_n} - \sqrt{\mu})$

$E(X_1) = Var(X_1) = \mu$. By CLT $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{\mu}} \xrightarrow{d} Norm(0,1)$.

$g(\mu) = \sqrt{\mu}$. $g'(\mu) = \frac{1}{2} \cdot \frac{1}{\sqrt{\mu}}$

By the Delta method. $\{g'(\mu)\}^2 \cdot \mu = \frac{1}{4} \cdot \frac{1}{\mu} \cdot \mu = \frac{1}{4}$

$Z_n \xrightarrow{d} Norm(0, \frac{1}{4})$

② $X_1 \cdots X_n \overset{iid}{\sim} Exp(\theta)$. with mean θ.

Find the limiting distributions of (a). $\bar{X}_n$ (b) $Z_n = \sqrt{n}(\bar{X}_n - \theta)/\bar{X}_n$ (c) $U_n = \sqrt{n}(\bar{X}_n - \theta)$ (d). $V_n = \sqrt{n}\{\log(\bar{X}_n) - \log(\theta)\}$

(a). $\bar{X}_n \xrightarrow{P} \theta$, $F(x) = \begin{cases} 0 & , x < \theta \\ 1 & , x \geq \theta \end{cases}$

(b) $Z_n = \frac{\sqrt{n}(\bar{X}_n - \theta)}{\bar{X}_n} = \underbrace{\frac{\sqrt{n}(\bar{X}_n - \theta)}{\sqrt{\theta^2}}}_{\xrightarrow{d} Norm(0,1)} \cdot \frac{\sqrt{\theta^2}}{\bar{X}_n} = \frac{\theta}{\bar{X}_n} \xrightarrow{P} 1$ By CMT.

$\xrightarrow{d} Norm(0,1)$

(c). $U_n = \sqrt{n}(\bar{X}_n - \theta) = \frac{\sqrt{n}(\bar{X}_n - \theta)}{\sqrt{\theta^2}} \cdot \sqrt{\theta^2} = \theta \cdot \frac{\sqrt{n}(\bar{X}_n - \theta)}{\theta} \xrightarrow{d} \theta Z$, $Z \sim Norm(0,1)$
$= Norm(0, \theta^2)$

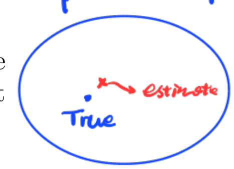
(d). $V_n = \sqrt{n}\{\log(\bar{X}_n) - \log(\theta)\} \xrightarrow{d} Norm(0, \theta^2 \cdot \{g'(\theta)\}^2)$
$= Norm(0, 1)$. using the Delta method.

Point Estimation

Setting: We have an iid sample $X_1, \ldots, X_n$ from a distribution $f(x;\theta)$ (we assume iid for simplicity). We usually say that the distribution is indexed by θ. The parameter θ could be a scalar, or it could also be a vector $\boldsymbol{\theta}^{\mathrm{T}} = (\theta_1, \theta_2)$.

Model: The parameter(s) θ is usually unknown. Thus, we provide a family of models by assuming "a" model $f(x;\theta)$. A model is a data-generating mechanism, or you can think of this as a pdf/cdf, etc.

parameter space Θ



Parameter Space: Describe the family of models you choose. You believe the true data-generating mechanism lives in the chosen family, which is equivalent to the true value of θ lives in the parameter space.

"Assume $X_1 \cdots X_n$ has a normal distribution"

$\Updownarrow$

The true model is within a family of distribution $\Leftrightarrow$ The true parameters (μ_0, σ_0^2) lives in a parameter space. $\{(\mu, \sigma^2) \mid \mu \in \mathbb{R}, \sigma^2 > 0\}$

All normal distributions with $\mu \in \mathbb{R}$, and $\sigma^2 > 0$

Estimation: Based on the sample, what is your "guess" about the true value of the parameter?

Statistic: A function of the random sample that does not depend on any parameters.

$T(X_1 \cdots X_n)$.

e.g. $\bar{X}_n$ $\quad \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{n-1}$

$\frac{\bar{X}_n - \mu}{\sigma}$ not a statistic

Estimator & Estimate:

Estimator: a statistic used to estimate parameters.
If it is used to estimate θ, usually we use $\hat{\theta}$ to denote the estimator.
$\hat{\theta} = \hat{\theta}(X_1 \cdots X_n)$ is a random variable.

Estimate: is the realized version of the estimate.
noun. $\hat{\theta} = \hat{\theta}(x_1, \cdots, x_n)$ is non-random (numbers)
lower case.

How to evaluate an estimator?

① Unbiasedness: if $E(\hat{\theta}_n) = \theta$, then $\hat{\theta}$ is unbiased.

asympotically unbias: $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$.

② Consistency: if $\hat{\theta}_n \xrightarrow{P} \theta$, the $\hat{\theta}_n$ is consistent.

③. Mean Square Error (MSE): $E\{(\hat{\theta}_n - \theta)^2\}$

MSE can be decomposed into two parts.

Let $\mu = E(\hat{\theta}_n)$, $E\{(\hat{\theta}_n - \mu + \mu - \theta)^2\}$

$$MSE = E\{(\hat{\theta}_n - \mu)^2\} + (\mu-\theta)^2 + 2E\{(\hat{\theta}_n - \mu)(\mu-\theta)\}$$

$$= \underline{Var(\hat{\theta}_n)} + \underline{(bias)^2}$$

If $\hat{\theta}_n$ is unbiased, then $MSE = Var(\hat{\theta}_n)$

Method of Moments

Setting: Suppose $X_1, \ldots, X_n$ are iid with pdf/pmf $f(x; \boldsymbol{\theta})$. We need to estimate $\boldsymbol{\theta}^{\mathrm{T}} = (\theta_1, \ldots, \theta_k)$.

Method: Method of Moments (MM) estimator.

1. Moments.

$$\mu_1 = E(X_1)$$
$$\mu_2 = E(X_1^2)$$
$$\vdots$$
$$\mu_k = E(X_1^k)$$

2. Sample moments

$$\hat{\mu}_1 = \bar{X}_n$$
$$\hat{\mu}_2 = \frac{\sum_{i=1}^{n} X_i^2}{n}$$
$$\vdots$$
$$\hat{\mu}_k = \frac{\sum_{i=1}^{n} X_i^k}{n}.$$

3. How to derive the MM estimators

Step 1: moment as a function of θ; Step 2: invert the transformation

$$\begin{cases} \mu_1 = \mu_1(\theta_1 \cdots \theta_k) \\ \vdots \\ \mu_k = \mu_k(\theta_1 \cdots \theta_k) \end{cases} \qquad \begin{cases} \theta_1 = \theta_1(\mu_1 \cdots \mu_k) \\ \vdots \\ \theta_k = \theta_k(\mu_1 \cdots \mu_k) \end{cases}$$

Step 3:
replace $\mu_1 \cdots \mu_k$ with the sample versions
$\hat{\theta}_i = \theta_i(\hat{\mu}_1, \cdots \hat{\mu}_k)$
for $i = 1, 2 \cdots k$.

4. If $\hat{\theta}$ is the MM estimator for θ,

$\Rightarrow \tau(\hat{\theta})$ is the MM estimator for $\tau(\theta)$

e.g. $\theta = (\mu, \sigma^2)$, $\hat{\theta}$
$\tau(\theta) = \mu^2 + \sigma^2$
$\tau(\hat{\theta})$ is the MM estimator for $\tau(\theta)$

Examples: ① $X_1 \cdots X_n \overset{iid}{\sim} Poi(\theta)$, a) Find the MM estimator of θ.

Step 1. $\mu_1 = E(X_1) = \theta$

Step 2. $\theta = \mu_1$

Step 3. $\hat{\theta} = \overline{X_n}$

b). Find the MM estimator for θ^2

$$\underset{\substack{\uparrow \\ \text{estimator} \\ \text{for } \theta^2}}{\widehat{\theta^2}} = (\underset{\substack{\uparrow \\ \text{estimator} \\ \text{for } \theta}}{\hat{\theta}})^2 = (\overline{X}_n)^2$$

②. $X_1 \cdots X_n \overset{iid}{\sim} \begin{cases} \text{(a) } Exp(\theta) \\ \text{(b) } Unif(0,\theta) \\ \text{(c) } f(x) = \theta \cdot x^{\theta-1}, \ x \in (0,1), \ \theta > 0. \end{cases}$ Find the MM estimator of θ.

(a). Step 1. $\mu_1 = E(X_1) = \theta$

Step 2. $\theta = \mu_1$

Step 3: $\hat{\theta} = \overline{X_n}$

(b). Step 1. $\mu_1 = E(X_1) = \frac{\theta}{2}$

Step 2: $\theta = 2\mu_1$

Step 3: $\hat{\theta} = 2\overline{X_n}$

(c) Step 1: $\mu_1 = E(X_1) = \int_0^1 x \cdot \theta x^{\theta-1} dx = \frac{\theta}{1+\theta}$

Step 2: $\theta = \frac{\mu_1}{1-\mu_1}$

Step 3: $\hat{\theta} = \frac{\overline{X}_n}{1-\overline{X}_n}$

③. $X_1 \cdots X_n \overset{iid}{\sim} f(x; \mu, \sigma^2)$. $\mu = E(X)$. $\sigma^2 = var(X)$. Find the MM estimators of (μ, σ^2).

Step 1: $\begin{cases} \mu_1 = E(X) = \mu \\ \mu_2 = E(X^2) = Var(X) + \{E(X)\}^2 = \sigma^2 + \mu^2 \end{cases}$

Step 2. $\begin{cases} \mu = \mu_1 \\ \sigma^2 = \mu_2 - \mu_1^2 \end{cases}$

Step 3: $\begin{cases} \hat{\mu} = \overline{X_n} \\ \hat{\sigma}^2 = \frac{\sum_{i=1}^n X_i^2}{n} - (\overline{X_n})^2 = \frac{1}{n} \cdot \sum_{i=1}^n (X_i - \overline{X}_n)^2 \end{cases}$