Tutorial for 2020-10-13 and 2020-10-16

Maximum Likelihood Estimation (MLE) $\swarrow$ method $\searrow$ $\hat{\theta}_{MLE}$

- How to use MLE → method

$$\begin{cases} \theta = \{\theta_1, \theta_2, \dots, \theta_k\} \\ X_1, X_2, \dots, X_n \overset{iid}{\sim} f_\theta(x) \end{cases}$$

① To find likelihood function $L(\theta)$

$$L(\theta) = L(\theta_1, \theta_2, \dots, \theta_k \mid x) = \begin{cases} \prod_{i=1}^{n} f(x_i \mid \theta) & \text{Continuous.} \\ \prod_{i=1}^{n} P(x_i \mid \theta) & \text{discrete.} \end{cases}$$

② To find log likelihood fun.. $l(\theta)$

$$l(\theta) = \log[L(\theta)] = \begin{cases} \sum_{i=1}^{n} \log f(x_i \mid \theta) & \text{Continuous..} \\ \sum_{i=1}^{n} \log P(x_i \mid \theta) & \text{discrete.} \end{cases}$$

$a \times b \rightarrow \log(a \times b) \rightarrow \log a + \log a.$

③ Take derivatives

k parameter $\Leftarrow$ k times.

≥ 1 solution.

$$\begin{cases} \frac{\partial l(\theta)}{\partial \theta_1} = 0 \\ \frac{\partial l(\theta)}{\partial \theta_2} = 0 \\ \cdots \\ \frac{\partial l(\theta)}{\partial \theta_k} = 0 \end{cases}$$

k equations

k unknown param

MLE Candidate

1 solution.

$\Rightarrow \hat{\theta}_{MLE} = \{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k\}$

④ Double check.

$k=1$. 1 parameter $\theta = \{\theta_1\}$

$l''(\theta) = \frac{\partial^2 l(\theta)}{\partial \theta_i^2} < 0$.

- Invariance property of MLE.

$$\begin{cases} X \sim f_\theta(x) \\ g(\cdot) \end{cases}$$

$\hat{\theta} \xrightarrow{MLE} \theta$

$g(\theta) \xleftarrow{MLE} g(\hat{\theta})$

-

① MLE may not be unbiased

$$\begin{cases} E(\hat{\theta}) \mathrel{!=} \theta \rightarrow \text{biased.} \\ E(\hat{\theta}) = \theta \rightarrow \text{unbiased} \end{cases}$$

See it later!!!

Sorry T^T

② MLE may not be unique.

e.g. $X_1, X_2, \ldots X_n \sim U[\theta-\frac{1}{2}, \theta+\frac{1}{2})$

$\theta \in \mathbb{R}$

Q: $\hat{\theta}_{MLE}$?

$\underbrace{\theta \in [x_{(n)}-\frac{1}{2}, x_{(1)}+\frac{1}{2})}$

③ MLE × $\underbrace{\text{asymptotically}}_{\text{d}}$ normal.

approximate / Limited

$t>0$, t is very small.

$\underbrace{P(\frac{n(\theta-\hat{\theta})}{\theta} \le t)}_{\text{c}} \xrightarrow{d}$ Normal

e.g. $X_1, X_2, \ldots X_n \sim U(0, \theta)$

Q. MLE for θ

$\hat{\theta}_{MLE} = X_{(n)}$

$P(\frac{n(\theta-\hat{\theta})}{\theta} \le t) \longrightarrow 1-e^{-t}$

$\rightarrow \hat{\theta}$ is $\exp(1)$

Information Inequality: $\theta \Leftarrow \hat{\theta}$

$\theta \in R^k$

$\epsilon > 0$ any

$\lim P(|\hat{\theta} - \theta| > \epsilon) = 0$

$\hat{\theta}_n$ is { Consistent:

~~asympto~~tically unbiased { $E(\hat{\theta}) = \theta$ & Limitat. $(E(\hat{\theta}) - \theta) \xrightarrow{P} 0$

efficient

normal distributed.

$\Rightarrow$

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N_k\left(0, \frac{1}{I_X(\theta)}\right)$$

Fisher Information matrix

$k \times k$ matrix

$$I_X(\theta) := E\left[\frac{\partial \log f_X(X|\theta)}{\partial \theta_i} \cdot \frac{\partial \log f_X(X|\theta)}{\partial \theta_j}\right].$$

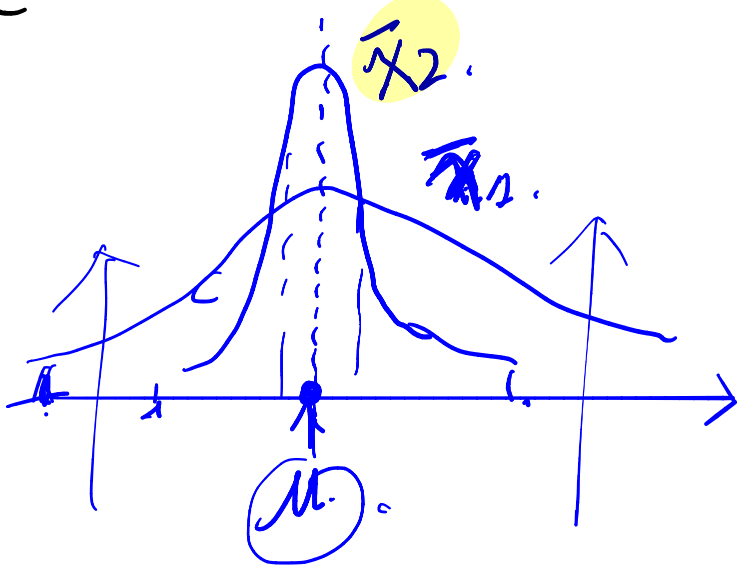
$i = 1, 2, \ldots k.$ $\quad j = 1, 2, \ldots k.$

$$= E\left[\underline{l'(\theta)}^T \cdot \underline{l'(\theta)}_{1 \times k}\right]$$

Efficient: (without asymptotically)

$\hat{\theta} \rightarrow$

$$\begin{cases} Var(\hat{\theta}) = \dfrac{1}{I_X(\theta)} \rightarrow \hat{\theta} \text{ is efficient} \\ \cdots > \dfrac{1}{I_X(\theta)} \rightarrow \hat{\theta} \text{ is sub-efficient} \\ \cdots < \cdots \rightarrow \hat{\theta} \text{ is super-efficient} \end{cases}$$



$X_1, X_2, \ldots X_n \sim N(\mu, \frac{\sigma^2}{n})$

σ^2 known

$E(\hat{\mu}) = \mu \qquad \frac{\sigma^2}{n}$

$\hat{\mu} = \bar{x} \sim N(\mu, Var)$

$Var(\hat{\theta}) = Var(\bar{x})$

$$\frac{1}{I_X(\theta)} = \frac{\sigma^2}{n}$$

$$I_X(\theta) = \frac{n}{\sigma^2}$$

Example:

Let $X_1, X_2, \ldots X_n \sim Bin(1, P)$ $P \in (0,1)$

Q MLE for P

pdf $p^{x_i}(1-p)^{1-x_i}$

Solution: $\theta = p$

① $L(p) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$

② $l(p) = \log L(p) = n[\bar{x} \log p + (1-\bar{x}) \cdot \log(1-p)]$

$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

③ $l'(p) = \frac{d\,l(p)}{dp} = n\left(\frac{\bar{x}}{p} - \frac{1-\bar{x}}{1-p}\right) = 0$

$\Rightarrow \hat{p} = \bar{x}$

④ $l''(p) = \left(-\frac{\bar{x}}{p^2} - \frac{(1-\bar{x})}{(1-p)^2}\right) < 0.$

$\bar{x}$ is a MLE for p //

$p^3 \to g(x) = x^3$

Q2: p^2, $p(1+p)$

$p \cdot (p+1) \cdot \log p$

invariance. $p^2 =: g(p)$

$\begin{cases} \bar{x} \xrightarrow{MLE} p \\ g(p) = p^2 \end{cases} \to g(\bar{x}) = (\bar{x})^2 \xrightarrow{MLE} g(p) = p^2$