STAT 135 Lecture 5

Henry Liev

2 July 2025

Method 0.1

Delta Method:
$Y, E(Y) = \mu, Var(Y) = \sigma^2$
Approximate $Var(f(Y))$

Solution. $f(Y) - f(\mu) \approx f'(\mu)(Y - \mu)$
$Var(f(Y)) \approx f'(Y)^2 Var(\hat{p})$
$Var(Y) = f'(\mu)^2\sigma^2$ ■

Remark 0.2

$\sqrt{n}(f(\bar{X}) - f(\mu)) \approx \mathcal{N}(0, f'(\mu)^2\sigma^2)$

Remark 0.3

Large sample properties of MLE

1. MLE is consistent: $\hat{\theta} \to \theta$
2. MLE are asymptotically normal
3. MLE is asymptotically efficient (optimality)

Remark 0.4

Intuition about consistency:

$$L(\theta) = \sum_{i=1}^{n} log(f(x_i|\theta))$$

$$\frac{1}{n}\sum_{i=1}^{n} log(f(x_i|\theta)) \xrightarrow{p} E(log(f(x|\theta)))$$

The maximizing θ of the average of the log-likelihood tends towards the θ that maximizes the expected log-likelihood.

Definition 0.5

A sequence of estimators T_n of a parameter is **consistent** if

$$T_n \xrightarrow{p} \theta$$

Definition 0.6

Convergence in probability

$$X_n \xrightarrow{p} \theta, \forall \varepsilon > 0, P(|X_n - \theta| > \varepsilon) \to 0 \text{ as } n \to \infty$$

Definition 0.7

Information about θ (Fisher Information)

$$I(\theta) = E\left[\left[\frac{\partial}{\partial\theta}\log(f(x|\theta))\right]^2\right]$$

$\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} \mathcal{N}(0, \frac{1}{I(\theta)})$ Minimizes variance by maximizing the information

Lemma 0.8

$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log(f(x|\theta))\right]$

Definition 0.9

Relative rate of change of density at x

$$\left(\frac{f'(x)}{f(x)}\right)^2$$

Method 0.10

MLE Playbook:
How to use MLE:

1. Find $\hat{\theta}_{MLE}$, maximize the likelihood
   - Write down log likelihood
   - Take derivative and set to 0
   - Solve
2. Find $I(\theta)$
   Usually use $-E\left[\frac{\partial^2}{\partial\theta^2}\log(f(x|\theta))\right]$
3. Form approximate interval: $\hat{\theta}_{MLE} \pm 1.96\frac{1}{\sqrt{nI(\hat{\theta}_{MLE})}}$

Example 0.11

$X_1, \ldots, X_n \overset{i.i.d}{\sim} \exp(\lambda)$
$\hat{\theta}_{MLE} = \frac{1}{\bar{X}}$
$\frac{1}{\bar{X}} \pm 1.96 \frac{1}{\sqrt{nI(\lambda)}}$
$\log(f(x|\theta)) = log(\lambda) - \lambda x$
$\frac{\partial}{\partial \lambda} \log(f(x|\lambda)) = \frac{1}{\lambda} - x$
$\frac{\partial^2}{\partial \lambda^2} \log(f(x|\lambda)) = -\frac{1}{\lambda^2}$
$I(\lambda) = \frac{1}{\lambda^2}$
$\hat{\theta}_{MLE} \pm 1.96 \frac{1}{\sqrt{\frac{n}{\bar{X}^2}}}$

When can we get CIs?

1. $\overline{X}$ CLT
2. $f(\overline{X})$ is a smooth function of $\overline{X}$ (Delta Method)
3. MLE

Definition 0.12

Parametric Bootstrap
A way to get CIs with simulation instead of math

Method 0.13

$X_1, \ldots, X_n \overset{i.i.d}{\sim} f(x|\theta)$
$\hat{\theta}$ estimates θ_0 ('θ_0 is the true value of θ)
Think about the distribution of $\hat{\theta} - \theta_0$
Suppose we have a distribution such that 95% of the distribution is contained in the interval (a, b)
$P(a \leq \hat{\theta} - \theta_0 \leq b) = 0.95$
$P(-b \leq \theta_0 - \hat{\theta} \leq -a) = 0.95$ $P(\hat{\theta} - b \leq \theta_0 \leq \hat{\theta} - a) = 0.95$
$(\hat{\theta} - b, \hat{\theta} - a)$ is a 95% CI for θ_0
Don't know θ_0 or distribution of $\hat{\theta} - \theta_0$, estimate θ_0 using $\hat{\theta}$ and create an estimated density $f(x|\hat{\theta})$ by simulating from $f(x|\hat{\theta})$, with which we can find $\hat{a}$ and $\hat{b}$
Use $\hat{a}$ and $\hat{b}$ to get a CI

Example 0.14

$X_1, \ldots, X_{50} \overset{i.i.d}{\sim} \text{Poisson}(\lambda)$
$\lambda = 3$, but we don't know it
$\bar{X}$ as our estimator of λ

1. Compute $\bar{X}_{\hat{\theta}} \to 3.2$
2. Take 10000 samples of size 50 from Poisson(3.2)
3. Compute θ^* for each sample
4. Approximate CI $= (\hat{\theta} - \hat{b}, \hat{\theta} - \hat{a})$