Lecture 16: MLE Asymptotics

Back to estimation:

For a finite sample we looked at estimators that are (1) unbiased and (2) low variance.

Asymptotically, we also want estimators that are

(1) asymptotically unbiased [consistency]

(2) asymptotic variance to be small

Theorem: MLEs are consistent. (*) under some regularity conditions (works for Exp. fam)

If $\hat{\theta}$ is my MLE for $\tau(\theta)$

then $\hat{\theta} \xrightarrow{P} \tau(\theta)$.

Defn: Asymptotic Normality

We say that $\hat{\theta}_N$ is asymptotically normal w/

(1) asymptotic mean $\tau(\theta)$

and (2) asymptotic variance $\nu(\theta)$

if

$$\textcircled{*} \quad \sqrt{N}\left(\hat{\theta}_N - \tau(\theta)\right) \xrightarrow{d} N(0, v(\theta))$$

and we write notation for above

$$\hat{\theta}_N \sim AN\left(\tau(\theta), \, v(\theta)/N\right)$$

Defn: Asymptotic Relative Efficiency (ARE)

If T_N and W_N are est. for $\tau(\theta)$

and

$$T_N \sim AN(\tau(\theta), \sigma_T^2(\theta))$$

$$W_N \sim AN(\tau(\theta), \sigma_W^2(\theta))$$

then we define the ARE of W_N wrt T_N as

$$ARE(W_N, T_N) = \frac{\sigma_T^2(\theta)}{\sigma_W^2(\theta)}$$

idea: If ARE < 1 the we prefer T_N
ARE > 1 " W_N

Ex. $X_n \overset{iid}{\sim} Pois(\lambda)$ and let $\tau(\lambda) = e^{-\lambda}$

$$= \mathbb{P}(X_n = 0)$$

$$= \frac{\lambda^0 e^{-\lambda}}{0!}$$

Notice: $\bar{X}$ is the MLE for λ

and so $e^{-\bar{X}}$ is the MLE for $e^{-\lambda}$

Alt: If $Y_n = \mathbb{1}(X_n = 0) \sim \text{Bern}(p)$

$\hookrightarrow p = P(Y_n = 1)$

so $\mathbb{E}Y_n = p = e^{-\lambda}$

$= \mathbb{P}(X_n = 0)$

$= e^{-\lambda}$

so $\mathbb{E}\bar{Y} = p = e^{-\lambda}$

pct. of zeros among X_n

Q: which is better asymptotically?

(1) $e^{-\bar{X}}$ (2) $\bar{Y}$

---

(1) $\bar{X} \sim AN(\lambda, \lambda/N)$ by CLT

What about $e^{-\bar{X}}$? $g(\bar{X})$ when $g(x) = e^{-x}$

Use FO Δ-method; $g'(x) = -e^{-x}$

$$g(\bar{X}) \sim AN(g(\lambda), [g'(\lambda)]^2 \lambda/N)$$

i.e. $$\boxed{e^{-\bar{X}} \sim AN(e^{-\lambda}, e^{-2\lambda}\lambda/N)}$$

$Y_n = \mathbb{1}(X_n = 0)$ and $\bar{Y} = \frac{1}{N}\sum_n Y_n$

$\sim \text{Bern}(e^{-\lambda})$

p

So $\boxed{\bar{Y} \sim AN\left(e^{-\lambda}, \frac{e^{-\lambda}(1-e^{-\lambda})}{N}\right)}$ $\boxed{Z_n \sim Bern(p) \quad \bar{Z} \sim AN\left(p, \frac{p(1-p)}{N}\right)}$

by CLT

Which is better?

$$ARE(\bar{Y}, e^{-\bar{X}}) = \frac{\text{asymp. var } e^{-\bar{X}}}{\text{asymp. var. } \bar{Y}}$$

$$= \frac{\lambda e^{-2\lambda}/\cancel{N}}{e^{-\lambda}(1-e^{-\lambda})/\cancel{N}} \frac{e^{\lambda}}{e^{\lambda}}$$

$$e^{\lambda} = 1 + \frac{\lambda}{1} + \frac{\lambda^2}{2} + \frac{\lambda^3}{3!} + \frac{\lambda^4}{4!} + \cdots$$

$$= \frac{\lambda e^{-\lambda}}{1 - e^{-\lambda}} \frac{e^{\lambda}}{e^{\lambda}}$$

$$= \frac{\lambda}{e^{\lambda} - 1}$$

$$= \frac{\lambda}{\lambda + \lambda^2 + \frac{\lambda^3}{3!} + \frac{\lambda^4}{4!} + \cdots}$$

$$= \frac{\lambda}{\lambda + \underset{\text{pos.}}{\text{smthg}}} < 1$$

So asymp. var $\bar{Y}$ > asymp. var $e^{-\bar{X}}$

i.e. we prefer $e^{-\bar{X}}$.

Defn: Asymptotic Efficiency

We say $\hat{\theta}_N$ is asymptotically efficient for $\tau(\theta)$ if

$$\hat{\theta}_N \sim AN(\tau(\theta), B(\theta))$$

CRLB

$$B(\theta) = \left(\frac{\partial \tau}{\partial \theta}\right)^2 / I_N(\theta)$$

Prev. Ex. $\quad e^{-\bar{X}} \sim AN\left(e^{-\lambda}, \frac{(e^{-\lambda})^2 \lambda}{N}\right)$

Q: is this asympt. eff.?

Calc CRLB: $\quad f(x) = \lambda^x e^{-\lambda} / x!$

$$\rightarrow \log f = x \log \lambda - \lambda - \log(x!)$$

$$\rightarrow \frac{\partial \log f}{\partial \lambda} = x/\lambda - 1$$

$$\rightarrow \frac{\partial^2 \log f}{\partial \lambda^2} = -\frac{x}{\lambda^2}$$

$$\rightarrow I(\lambda) = -E\left[\frac{\partial^2}{\partial \lambda} \cdots\right] = \frac{1}{\lambda^2} E[X]$$

$$= \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

So $I_N(\lambda) = N I(\lambda) = N/\lambda$

So $\tau(\lambda) = e^{-\lambda}$ then $\left(\frac{\partial \tau}{\partial \lambda}\right)^2 = \left(-e^{-\lambda}\right)^2 = \left(e^{-\lambda}\right)^2$

and $$B = \left(\frac{\partial \tau}{\partial \lambda}\right)^2 \Big/ I_N(\lambda) = \boxed{\frac{\left(e^{-\lambda}\right)^2 \lambda}{N}}$$

So my est. $e^{-\bar{x}}$ is asymp. eff. for $e^{-\lambda}$.

Theorem: MLEs are asymptotically efficient! (*) under some reg. conditions

$$\hat{\theta}_{MLE} \sim AN\left(\tau(\theta), \frac{(\partial \tau / \partial \theta)^2}{I_N(\theta)}\right)$$

MLE for $\tau(\theta)$

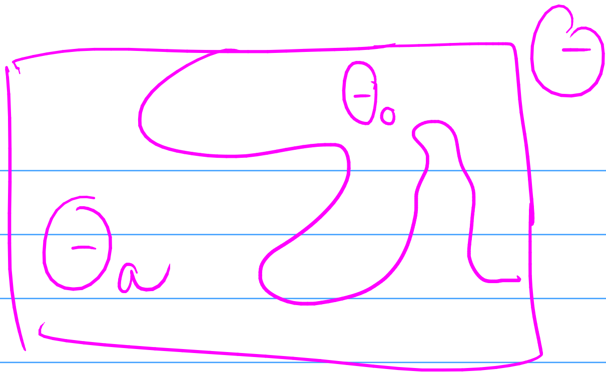
Hypothesis testing

Defn: Hypothesis a hypothesis is a statement about a parameter

$$\underbrace{H_0 : \theta \in \Theta_0}_{\text{null hypothesis}} \quad \text{v.} \quad \underbrace{H_a : \theta \in \Theta_a}_{\text{alt. hypothesis}}$$

Constraint: (1) $\Theta_0 \Theta_a = \emptyset$ (2) $\Theta = \Theta_0 \cup \Theta_a$

i.e. partition



Ex. Let θ be pct. of defective items in my manufacturing process.

$$\Theta = [0, 1]$$

might test:

$$H_0: \theta \le .1 \quad v. \quad H_a: \theta > .1$$

$$\Theta_0 = [0, .1], \quad \Theta_a = (.1, 1]$$

Ex. Let θ denote the change in BP after taking some medicine.

might test

$$H_0: \theta = 0 \quad v. \quad H_a: \theta \neq 0$$

$$\Theta = \mathbb{R}; \quad \Theta_0 = \{0\}; \quad \Theta_0 = \mathbb{R} \setminus \{0\}$$

If θ is a 1-D parameter $\theta \in \mathbb{R}$ then

1. a test of the form

$$H_0: \theta \leq c \quad v. \quad H_a: \theta > c$$

$$H_0: \theta < c \quad v. \quad H_a: \theta \geq c$$

$$\vdots \quad > \quad \vdots \quad \leq \quad \vdots$$

$$\vdots \quad \geq \quad \vdots \quad < \quad \vdots$$

is called a <u>one-sided hypothesis test</u>.

2. A test of the form

$$H_0: \theta = c \quad v. \quad H_a: \theta \neq c$$

$$H_0: \theta \neq c \quad v. \quad H_a: \theta = c$$

is called a <u>two-sided</u> test.

3. A test of the form

$$H_0: \theta = a \quad v. \quad H_a: \theta = b$$

is called a <u>simple hypothesis test</u>.

Defn: Hypothesis Testing Procedure

idea: want to determine if $\theta \in \Theta_0$ v. $\theta \in \Theta_a$ based on some data I collect.