Lecture 14

Thursday, October 24, 2024 11:05 AM

Ex. Let $X_n \overset{iid}{\sim} \text{Bern}(p)$

Consider $g(x) = x\log(x/p) - (1-x)\log\left(\frac{1-x}{1-p}\right)$

KL divergence between two Bernoulli

What can I say about $g(\bar{X})$?

$\psi^2 = \psi^2(p)$

CLT: $\sqrt{N}(\bar{X}-p) \xrightarrow{d} N(0, P(1-P))$

notice: $g'(x) = \log(x/1-x) - \log(P/1-p)$

so $g'(p) = 0$

Can't use FO Δ-method

However, $g''(x) = \frac{1}{x(1-x)}$

and $g''(p) \neq 0$ for $0 < p < 1$

Can use SO Δ-method:

$$N\left(g(\bar{x})-g(p)\right) \xrightarrow{d} \underbrace{\frac{\psi^2 g''(p)}{2}}_{\frac{p(1-p)\frac{1}{p(1-p)}}{2}=\frac{1}{2}} \chi^2(1)$$

i.e.

$$N\left(g(\bar{x})-g(p)\right) \xrightarrow{d} \frac{1}{2}\chi^2(1)$$

Pf. Second order taylor expansion of g about θ

$$g(x) \approx g(\theta)+g'(\theta)(x-\theta)+\frac{1}{2}g''(\theta)(x-\theta)^2$$

if $g'(\theta)=0$ then

$$g(x) \approx g(\theta)+\frac{1}{2}g''(\theta)(x-\theta)^2$$

i.e.

$$g(x) - g(\theta) \approx \frac{1}{2} g''(\theta)(x-\theta)^2$$

$$\Rightarrow \quad N\left(g(x) - g(\theta)\right) \approx \frac{1}{2} g''(\theta) \left[\sqrt{N}(x-\theta)\right]^2$$

Know: $\boxed{\sqrt{N}(Y_N - \theta) \xrightarrow{d} N(0, \psi^2)}$

$$\underline{N\left(g(Y_N) - g(\theta)\right)} \approx \frac{1}{2} g''(\theta) \left[\sqrt{N}(Y_N - \theta)\right]^2$$

$$\to \psi N(0,1)$$

$$\to \underbrace{\frac{1}{2} g''(\theta)\left(\psi N(0,1)\right)^2}$$

$$\boxed{\frac{1}{2} g''(\theta) \psi^2 \chi^2(1)}$$

Back to estimation

Finite samples, considered:

① bias

① bias
② Variance

For large samples (asymptotically) we also want ests that
① are asymptotically unbiased
(consistency)

② are asymptotically low variance

<u>Theorem</u>: MLEs are consistent

(*) need some tech. conds, works for exp. fams

If $\hat{\theta}_{MLE}$ is the MLE for θ then

$$\hat{\theta}_{MLE} \xrightarrow{P} \theta.$$

<u>Defn: Asymptotically Normal</u>

We say $\hat{\theta}$ is asympt. normal w/

We say θ is asympt. normal w/

1. asymptotic mean $\tau(\theta)$
2. asymptotic variance $v(\theta)$

if

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

Notation:

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

Defn: Asymptotic Relative Efficiency (ARE)

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

and

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

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

then the ARE of W_N w.r.t. T_N is

$$ARE(W_N, T_N) = \sigma_T^2 / \sigma_W^2$$

$$\text{ARE}(W_N, T_N) = \sigma_T^2 / \sigma_W^2$$

If ARE < 1 then we prefer T_N

$\quad\quad > 1 \quad\quad$ " $\quad\quad W_N$

Ex. Let $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$

and let $\tau(\lambda) = e^{-\lambda} = P(X_n = 0) = \dfrac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda}$

Way 1: $\bar{X}$ is the MLE for λ

hence $\boxed{e^{-\bar{X}}}$ is the MLE for $e^{-\lambda}$

Way 2: $Y_n = \mathbb{1}(X_n = 0)$

could consider $\boxed{\bar{Y}}$ = pct. of X_n that are zero.

Q: which is better ① $e^{-\bar{X}}$ or ② $\bar{Y}$?

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

What abot $e^{-\bar{X}}$? Use Δ-method.

$$g(x) = e^{-x}, \quad g'(x) = -e^{-x}$$

$$g'(\lambda) = -e^{-\lambda} \neq 0$$

So use FO Δ-method.

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

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

---

(2) $Y_n = \mathbb{1}(X_n = 0) \sim Bern(p)$

$$p = P(Y_n = 1)$$
$$= P(X_n = 0) = e^{-\lambda}$$

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

CLT: $\boxed{\bar{Y} \sim AN(e^{-\lambda}, e^{-\lambda}(1-e^{-\lambda})/N)}$

Which do we prefer?

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

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

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

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

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

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

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

So we prefer $e^{-\bar{x}}$.

Defn: Asymptotic Efficiency

We say $\hat{\theta}$ is asymp. efficient for $\tau(\theta)$ if

$$\hat{\theta} \sim AN(\tau(\theta), \underbrace{B(\theta)}_{\text{CRLB}})$$

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

---

eg, $e^{-\bar{x}} \sim AN\left(e^{-\lambda}, \boxed{e^{-2\lambda} \lambda / N}\right)$

What's the CRLB?

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

$\rightarrow f_\lambda(x) = \lambda^x e^{-\lambda} / x!$

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

$\rightarrow \frac{\partial}{\partial x}[\cdots] = x/\lambda - 1$

$\rightarrow \frac{\partial^2}{\partial \lambda^2}[\cdots] = -\frac{x}{\lambda^2}$

$\rightarrow I(\lambda) = -E[\quad] = -E[-X/\lambda^2] = \frac{1}{\lambda^2} E X$

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

$I_N(\lambda) = N/\lambda$, $\tau(\lambda) = e^{-\lambda}$, $\tau'(\lambda) = -e^{-\lambda}$

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

Theorem: MLEs are asymptotically efficient (*)

(*) (1) likelihood diff'able

(2) $I(\theta) \neq 0$

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

<u>EXAM 2</u> φP 5–8

SP 5–9

lecture 7 — now

Hypothesis testing

A hypothesis is a statement about parameters:

$$H_0 : \theta \in \Theta_0 \quad v. \quad H_a : \theta \in \Theta_a$$

null hypothesis alt. hypothesis

call $\Theta = \Theta_0 \cup \Theta_a$

<u>Constraint</u>: $\Theta_0 \cap \Theta_a = \emptyset$.
