Theorem: Central Limit Theorem

If X_n are iid w/ $E[X_n] = \mu$ and $Var(X_n) = \sigma^2 < \infty$

then

$$\sqrt{N}\left(\frac{\bar{X}_N - \mu}{\sigma}\right) \xrightarrow{d} N(0,1).$$

Theorem: Δ-method

If (Y_n) is a seq of RVs where

$$\sqrt{N}(Y_n - \theta) \xrightarrow{d} N(0, \psi^2)$$

then if g is differentiable and $g'(\theta) \neq 0$

then

$$\sqrt{N}(g(Y_n) - g(\theta)) \xrightarrow{d} N(0, [g'(\theta)]^2 \psi^2)$$

Ex. Variance-Stabilizing Transformation

$$Y \sim AN(\theta, \psi^2/N)$$

ψ^2 may depend on θ

Q: Is there a transformation g so that

$$g(Y) \sim AN(g(\theta), \sim)$$

doesn't depend on θ

depend on θ

Soln: Use Δ-method b/c it says:

$$g(Y) \sim AN\left(g(\theta), \boxed{[g'(\theta)]^2 \varphi^2_{(\theta)}}\right)$$

want to find a g so that

$$[g'(\theta)]^2 \varphi(\theta)^2 = \text{constant}$$

So we get a condition that says choose g s.t.

$$\underbrace{[g'(\theta)]^2 \varphi(\theta)^2 = c}_{\text{ODE}}$$

some constant.

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

then the CLT says:

$\sqrt{N}\left(\dfrac{\bar{X}_N - \lambda}{\sqrt{\lambda}}\right) \xrightarrow{d} N(0,1)$	$\underset{\overset{\shortparallel}{Y}}{\bar{X}_N} \sim AN\left(\underset{\theta}{\lambda}, \overbrace{\lambda/N}^{\varphi^2_{(\lambda)}}\right)$

Δ-method says:

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

choose g to make this constant!

Sol. ODE:

Choose g to make this constant!

Set ODE:

$$\left(\frac{dg}{dx}\right)^2 \frac{\lambda}{N} = c$$

$$\Rightarrow \frac{dg}{dx} = \sqrt{\frac{cN}{\lambda}}$$

$$\Rightarrow dg = \sqrt{\frac{cN}{\lambda}}\, dx$$

$$\Rightarrow g = \int dg \propto \int \frac{1}{\sqrt{x}}\, dx$$

$$\propto \sqrt{x}$$

So $\boxed{g(x) = \sqrt{x}}$ $\longrightarrow \frac{dg}{dx} = \frac{1}{2\sqrt{x}}$

So my variance will be

$$[g'(x)]^2 \frac{\lambda}{N} = \left(\frac{1}{2\sqrt{\lambda}}\right)^2 \frac{\lambda}{N} = \frac{1}{4N} \quad \checkmark$$

i.e. $\boxed{\sqrt{\bar{X}} \sim AN\left(\sqrt{\lambda}, \frac{1}{4N}\right)}$

Theorem: Second-Order Δ-method

If $\sqrt{N}(Y_n - \theta) \xrightarrow{d} N(0, \psi^2)$

and g is twice-differentiable, but $g'(\theta) = 0$.

then

$$\boxed{N\left(g(Y_N) - g(\theta)\right) \xrightarrow{d} \frac{\varphi^2 g''(\theta)}{2} \underbrace{\chi^2(1)}_{N(0,1)^2}}$$

Ex. Let $X_n \overset{iid}{\sim}$ Bernoulli(p) and let

$$g(t) = \underbrace{t \log(t/p) - (1-t)\log\left(\frac{1-t}{1-p}\right)}_{\text{KL-divergence}}$$

KL-divergence dist btwn Bern(p) and Ben(t)

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

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

notice that

$$g'(t) = \log\left(\frac{t}{1-t}\right) - \log\left(p/_{1-p}\right)$$

and $g'(p) = \log\left(p/_{1-p}\right) - \log\left(p/_{1-p}\right) = 0$.

$g'(p) = 0$! No delta method!

Use second order method:

$$g''(t) = \frac{1}{t} + \frac{1}{1-t} = \frac{1}{t(1-t)}$$

Second Order Δ-method says

$$n\left(g(\bar{X}) - g(p)\right) \xrightarrow{d} \frac{\psi^2 g''(p)}{2} \chi^2_{(1)}$$

here $\psi^2 = p(1-p)$

$$g''(p) = \frac{1}{p(1-p)}$$

so

$$\cdots \xrightarrow{d} \frac{1}{2}\chi^2_{(1)}$$

<u>proof of SO-ΔM</u>

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

Taylor Expansion:

$$g(Y_N) \approx g(\theta) + \cancel{g'}^{\,0}(\theta)(Y_n - \theta) + \frac{1}{2}g''(\theta)(Y_n - \theta)^2 + \cdots$$

so

$$n\left(g(Y_n) - g(\theta)\right) \approx \frac{1}{2}g''(\theta)(\sqrt{n}(Y_n - \theta))^2$$

$$\xrightarrow{d} \varphi N(0,1)$$

$$\xrightarrow{d} \frac{1}{2} g''(\theta) (\varphi N(0,1))^2$$

$$\xrightarrow{d} \frac{1}{2} g''(\theta) \varphi^2 \chi^2(1).$$

Back to estimation ....

For a finite sample

Q: what is the best estimator?

looked for unbiased ests. w/ small variances

Asymptotically ... same Q

Consider:

1. asymptotic unbiasedness (consistency)
2. asymptotic variance (this point)

Defn: Consistency

We say $\hat{\theta}$ is consistent for θ if

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

Theorem: MLEs are consistent. (*) some regularity needed (works for

If $\hat{\theta}$ is the MLE for $\tau(\theta)$ then (... exp. families)

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

Defn: Asymptotic Normality

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

w/ ① asymptotic mean $\tau(\theta)$

② asymptotic variance $v(\theta)$

if

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

or

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

Defn: Asymptotic Relative Efficiency (ARE)

Let T_N and W_N are estimators 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

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

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

$ARE(W_N, T_N) < 1$ then we prefer T b/c $\sigma_T^2 < \sigma_W^2$.

$ARE(W_N, T_N) > 1$ we prefer W since $\sigma_T^2 > \sigma_W^2$.

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

let

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

Know: $\bar{X}$ is MLE for λ.

So $\boxed{e^{-\bar{X}}}$ is the MLE for $e^{-\lambda} = \tau(\lambda)$.

one way

Alt: $\tau(\lambda) = P(X_n = 0)$

let $Y_n = \mathbb{1}(X_n = 0)$ then $E[Y_n] = P(X_n = 0)$

furthermore $Y_n \sim Bernoulli(p)$

where $p = \mathbb{P}(X_n = 0) = e^{-\lambda}$

consider

$$\boxed{\overline{Y} = \frac{1}{N}\sum_{n=1}^{N} Y_n},$$

a competing estimator (MLE of Y_ns)

notice $\mathbb{E}[\overline{Y}] = \mathbb{E}(Y_n) = e^{-\lambda}$

and $Var(\overline{Y}) = \frac{p(1-p)}{N} = \frac{e^{-\lambda}(1-e^{-\lambda})}{N}$

know $\overline{Y} \xrightarrow{P} T(\lambda)$ by WLLN

which is better?

1. $e^{-\overline{X}}$?

2. $\overline{Y}$?

Know:

(i) $\Rightarrow \overline{X} \sim AN(\lambda, \lambda/N)$ by CLT

so by Δ-method $g(x) = e^{-x} \Rightarrow g'(x) = -e^{-x}$

(ii) $\Rightarrow e^{-\overline{X}} \sim AN(e^{-\lambda}, (-e^{-\lambda})^2 \frac{\lambda}{N})$

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

(iii) $\overline{Y} \sim AN(p, \frac{p(1-p)}{N})$ where $p = e^{-\lambda}$

(iii) $\bar{Y} \sim AN\left(p, \frac{p(1-p)}{N}\right)$ when $p = e^{-\lambda}$

$$\boxed{\sim AN\left(e^{-\lambda}, \frac{e^{-\lambda}(1-e^{-\lambda})}{N}\right)}$$

ARE?

$$ARE(\bar{Y}, e^{-\bar{X}}) = \frac{\frac{(e^{-\lambda})^{\cancel{2}}\lambda}{\cancel{N}}}{\frac{\cancel{e^{-\lambda}}(1-e^{-\lambda})}{\cancel{N}}}$$

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

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

Taylor Expansion of

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

$$e^{x} - 1 = x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$$

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

λ + something pos

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

Defn: We say $\hat{\theta}$ is asymptotically efficient for $T(\theta)$ if

basically the best estimators

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

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

the best estimators asymptotically

when

$$B(\theta) = \frac{\left(\frac{\partial \tau}{\partial \theta}\right)^2}{N I(\theta)} = CRLB.$$

Ex. Prev. Example

$$e^{-\bar{X}} \sim AN\Big(\underbrace{\boxed{e^{-\lambda}}}_{\tau(\lambda)}, \boxed{(e^{-\lambda})^2 \frac{\lambda}{N}}\Big)$$

is this asymptotically efficient?

$I(\lambda)$?

$$\log(f(x)) = \log\left(\frac{e^{-\lambda}\lambda^x}{x!}\right) = -\lambda + x\log\lambda - \log x!$$

$$\frac{\partial}{\partial \lambda}(\ldots) = -1 + \frac{x}{\lambda}$$

$$\frac{\partial^2}{\partial \lambda^2}(\ldots) = -\frac{x}{\lambda^2}$$

$$I(\lambda) = -E\left[\frac{\partial^2 \log f(x)}{\partial \lambda^2}\right] = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

So $I_N(\lambda) = N \frac{1}{\lambda}$

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

So the CRLB for $\tau(\lambda) = e^{-\lambda}$ is

$$B = \frac{(e^{-\lambda})^2}{I_N(\lambda)} = \boxed{\frac{(e^{-\lambda})^2 \lambda}{N}}$$

✓

$$LS = \frac{\quad}{I_N(\lambda)} = \frac{\quad}{N}$$

So $e^{-\bar{X}}$ is **Asymptotically efficient.**

<u>Theorem</u>: MLEs are asymptotically efficient. ✱

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

MLE for $\tau(\theta)$.
