Defn: Asymptotic Normality

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

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

if

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

and write

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

Defn: Asymptotic Relative Efficiency (ARE)

Let T_N and W_N be estimators 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 the ARE of W_N w.r.t. T_N is

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

Idea: If ARE < 1 then we prefer T_N

if ARE > 1 " W_N

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

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

Know: $\bar{X}$ is the MLE of λ so

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

one way

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

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

So $\mathbb{E}[Y_n] = p = e^{-\lambda}$

So $\mathbb{E}[\bar{Y}] = e^{-\lambda}$ another way of est $\tau(\lambda)$

Q: Which is better (asymptotically?)

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

① $e^{-\bar{X}}$, ② $\bar{Y}$.

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

$\mathbb{E}\bar{X} = \lambda$

$Var(\bar{X}) = \lambda/N$

So what about $e^{-\bar{X}}$? $\leftarrow g(\bar{X}) = e^{-\bar{X}}$

$\Downarrow g(\lambda) = e^{-\lambda}$

$g'(\lambda) = -e^{-\lambda} \Rightarrow (g'(\lambda))^2 = e^{-2\lambda}$

Use Δ-method:

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

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

② $\bar{Y}$?

$\bar{Y} = \frac{1}{N}\sum_n Y_n$; $Y_n \sim Bern(\underbrace{e^{-\lambda}}_{p})$

$$\bar{Y} \sim AN\left(p, \frac{p(1-p)}{N}\right)$$

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

Calculate ARE

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

$$\mathrm{ARE}(Y, e \quad) = \quad \overline{\text{asymp. var } \bar{Y}}$$

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

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

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

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

$$= \frac{\lambda}{\lambda + \underbrace{\frac{\lambda^2}{2} + \frac{\lambda^3}{3!} + \frac{\lambda^4}{4!} + \cdots}_{>0}}$$

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

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

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

Defn: Asymptotic Efficiency

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

infinite sample UMVUE

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

CRLB

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

Prev. Ex.

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

Q: is this asymp. efficient?

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

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

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

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

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

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

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

Yes: Asymp. efficient b/c asymp var = CRLB.

Theorem: MLEs are asymptotically efficient. (*)

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

MLE for $\tau(\theta)$

(*) under some conditions

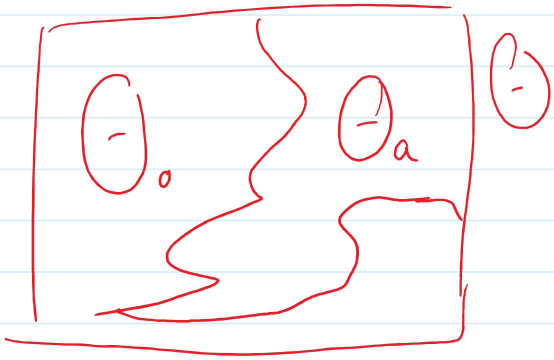
Defn: Hypothesis

a hypothesis is a statement about a parameter,

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

constraint:

$$\left.\begin{array}{l} ① \ \Theta_0 \Theta_a = \emptyset \\ ② \ \Theta = \Theta_0 \cup \Theta_a \end{array}\right\} \text{partition}$$



Ex. Let θ be the propn of defective items in some production process

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

might test:

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

$$\left[H_0 : \theta \in [0, .1] \quad v. \quad H_a : \theta \in (.1, 1] \right]$$

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

Might test

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

$$\left[\Theta = \mathbb{R}, \; \Theta_0 = \{0\}, \; \Theta_a = \mathbb{R} \setminus \{0\} \right]$$

$$\boxed{\Theta = \mathbb{R}, \ \Theta_0 = \{0\}, \ \Theta_a = \mathbb{R} \setminus \{0\}}$$

If θ is a 1-d parameter (e.g. $\theta \in \mathbb{R}$) then a test of the form

(1) $H_0: \theta \leq c$ v. $H_a: \theta > c$
or $H_0: \theta < c$ v. $H_a: \theta \geq c$
or $H_0: \theta \geq c$ v. $H_a: \theta < c$
or $H_0: \theta > c$ v. $H_a: \theta \leq c$

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

(2) $H_0: \theta = c$ v. $H_a: \theta \neq c$
or $H_0: \theta \neq c$ v. $H_a: \theta = c$

is called <u>a two sided hypothesis</u>

(3) $H_0: \theta = c_0$ v. $H_a: \theta = c_a$

one value in Θ_0 and Θ_a

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

<u>Defn: Hypothesis Testing Procedure</u>

Idea: want to determine if $\theta \in \Theta_0$ or $\theta \in \Theta_a$ is more plausible/consistent w/ the data I see.

Let $\mathcal{X}$ be the support of $\underset{\sim}{X}$

[typically $\mathcal{X} \subset \mathbb{R}^N$]

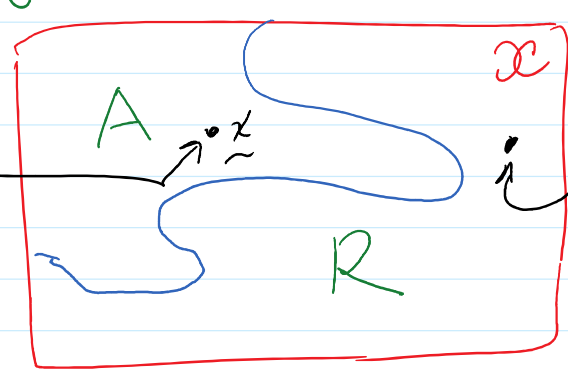
A HT procedure is simply a rule that partitions $\mathcal{X}$ into

$$\mathcal{X} = A \cup R$$

acceptance region : null plausible

reject region : null implausible

data consistent w/ null



data is inconsistent w/ null

We "reject H_0" if $\underset{\sim}{x} \in R$

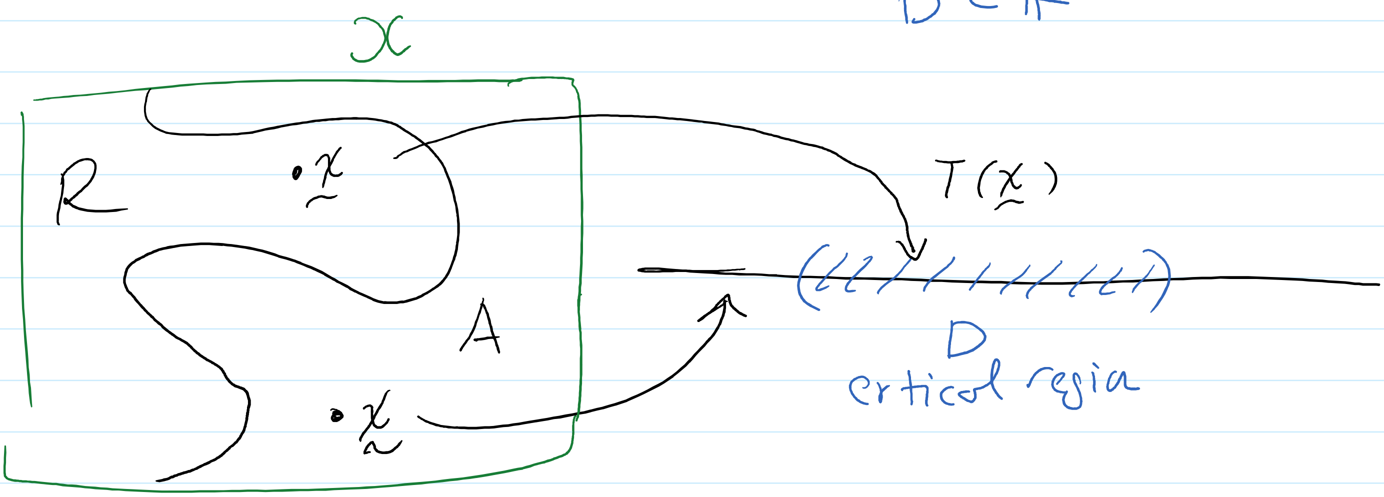
We "fail to reject H_0" if $\underset{\sim}{x} \in A$

Often we can define R (equiv. A) through a "test statistic" so that our HT is

$$R = \{\underset{\sim}{x} \mid T(\underset{\sim}{x}) \in D\}$$

critical region

$D \subset \mathbb{R}^N$



ex, $T = \bar{X}$ and $D = (5, \infty)$

my HT is

$$R = \{\underset{\sim}{x} \mid \bar{X} > 5\}$$