Not responsible for sufficient statistics or asymptotic tests.

Fisher Information $\theta \in \mathbb{R}^n$

$$I_1(\theta) = -E\left[\frac{\partial^2 \log f(x;\theta)}{\partial \theta}\right] = E\left[\left(\frac{\partial \log f(x;\theta)}{\partial \theta}\right)^2\right]$$

single observation

$$I_n(\theta) = -E\left[\sum_{i=1}^{n} \frac{\partial^2 \log f(x;\theta)}{\partial \theta^2}\right] = E\left[\sum_{i=1}^{n}\left(\frac{\partial \log f(x;\theta)}{\partial \theta}\right)^2\right]$$

$= n I_1(\theta)$ for n iid observations (random sample).

Cramer-Rao Inequality — This is crucial

If Y is an unbiased estimator of θ ($E[Y]=\theta$), then $Var(Y) \geq [I_n(\theta)]^{-1}$

Based on whole sample

If $\theta \in \mathbb{R}^n$, then $Var(Y) - [I_n(\theta)]^{-1}$ is non-negative definite.

An estimator is said to be efficient if it is unbiased and its variance equals the Cramer-Rao lower bound. Equivalently, $Y = U(X)$ is efficient for θ if $E[Y]=\theta$ and $Var[Y] = [I_n(\theta)]^{-1}$.

Example: $X_i \overset{iid}{\sim} N(0,\theta)$, $0<\theta<+\infty$. Show that $\hat{\theta}_{MLE}$ is efficient.

1 Find $\hat{\theta}_{MLE}$.

1a $$L(\theta, x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sqrt{\theta}} \exp\left\{-\frac{x_i^2}{2\theta}\right\}$$

$$= \frac{1}{(2\pi)^{n/2}} \frac{1}{(\theta)^{n/2}} \exp\left\{-\frac{\sum x_i^2}{2\theta}\right\}$$

1b $$\log L(\theta, x_1, \ldots, x_n) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{1}{2\theta}\sum_{i=1}^{n} x_i^2$$

1c Take FOC wrt θ:

$$(\theta): \ -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2 = \frac{n}{2\theta} \Rightarrow \frac{1}{\theta}\sum_{i=1}^{n} x_i^2 = n \Rightarrow \hat{\theta}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

2 Verify $\hat{\theta}_{MLE}$ is unbiased:

$$E[\hat{\theta}_{MLE}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i^2] = \frac{1}{n}\sum_{i=1}^{n}\theta = \frac{n}{n}\theta = \theta$$

3 Compute $I_1(\theta)$

3a Compute $\log f(x;\theta)$

$$\log f(x;\theta) = \log\left[\frac{1}{\sqrt{2\pi}\sqrt{\theta}}\exp\left\{-\frac{x^2}{2\theta}\right\}\right]$$

$$= -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\theta - \frac{x^2}{2\theta}$$

3b Compute $s(x;\theta)$

$$s(x;\theta) = \frac{\partial \log f(x;\theta)}{\partial\theta} = -\frac{1}{2\theta} + \frac{x^2}{2\theta^2}$$

3c Compute $\frac{\partial s(x;\theta)}{\partial\theta}$

$$\frac{\partial s(x;\theta)}{\partial\theta} = \frac{1}{2\theta^2} - \frac{x^2}{\theta^3}$$

3d $$I_1(\theta) = -E\left[\frac{\partial s(X;\theta)}{\partial\theta}\right] = E\left[\frac{X^2}{\theta^3} - \frac{1}{2\theta^2}\right] = \frac{E[X]^2}{\theta^3} - \frac{1}{2\theta^2}$$

$$= \frac{\theta}{\theta^3} - \frac{1}{2\theta^2} = \frac{1}{2\theta^2}$$

4 Compute $I_n(\theta)$

$$I_n(\theta) = n\,I_1(\theta) = n\cdot\frac{1}{2\theta^2} = \frac{n}{2\theta^2}$$

5] Compute Cramer-Rao lower bound for the sample

$$I^{-1}(\theta) = \frac{1}{\frac{n}{2\theta^2}} = \frac{2\theta^2}{n}$$

6] Compute $Var(\hat{\theta}_{MLE})$

$$Var(\hat{\theta}_{MLE}) = Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2\right] = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[E[X_i^4] - [E[X_i^2]]^2\right]$$

This step is not obvious. Need to verify $E[X_i^4]=3\theta^2$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\left[3\theta^2 - \theta^2\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\left[2\theta^2\right] = \frac{2\theta^2}{n}$$

7] Compare $Var(\hat{\theta}_{MLE})$ to $I^{-1}(\theta)$

$$Var(\hat{\theta}_{MLE}) = \frac{2\theta^2}{n} = \frac{2\theta^2}{n} = I_n(\theta)$$

Therefore, $\hat{\theta}_{MLE}$ is an efficient estimator of θ.

Theorem: If $\hat{\theta}_{MLE}$ is an MLE of θ, then $\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, [I_1(\theta)]^{-1})$

single observation

Example: If $X_i \overset{iid}{\sim} N(0, \theta^2)$ and $\hat{\theta}_{MLE}$ is the MLE of θ. asymptotic distribution of $\sqrt{n}(\hat{\theta}_{MLE} - \theta)$.

$$\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, [I_1(\theta)]^{-1})$$

Just compute this

Test of statistical hypothesis.

Test $\Leftrightarrow$ Critical region C

- Decision rule: If $(x_1, \ldots, x_n) \in C$, reject H_0
- Significance level: $Pr[(x_1, \ldots, x_n) \in C \mid H_0] = \alpha \Rightarrow \alpha$ is the significance level for the test C.

	H_0	H_1
A		II
R	I	

$$power(\theta, C) = Pr[(x_1, \ldots, x_n) \in C \mid \theta]$$

$$power(H_1, C) = Pr[(x_1, \ldots, x_n) \in C \mid H_1]$$
$$= 1 - \beta$$

where $\beta = Pr[(x_1, \ldots, x_n) \notin C \mid H_1] = Pr[\text{Type II error}]$

$$\alpha = \max_{\theta \in H_0} Pr[(x_1, \ldots, x_n) \in C]$$

① Size of test
② Significance level of the test
③ Maximum probability of committing type I error
④ Maximum of the power of the test when H_0 is true.

Example:

$X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, σ^2 known, $\alpha = 0.05$

$H_0: \mu = \mu_0 \quad H_1: \mu > \mu_0$

Statistic: $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

Under H_0: $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

Under H_1: $Z_1 = \dfrac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \sim N(0,1)$

Decision rule:

Reject H_0 if $Z_0 > 1.645$

Power function $P(H_1, C) = Pr[\underbrace{Z_0 > 1.645}_{\text{reject the null}} | H_1]$

$$= Pr\left[\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > 1.645 \,|\, H_1\right]$$

$$= Pr\left[\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} > \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + 1.645 \,|\, H_1\right]$$

$$= Pr\left[Z_1 > \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + 1.645 \,|\, H_1\right]$$

$$= 1 - \Phi\left(\frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + 1.645\right)$$

probability of correctly rejecting the null when H_1 is true.

$$= \Phi\left(\frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}} - 1.645\right)$$

when $\mu_1 \uparrow$, $P(H_1, C) \uparrow$

Best Critical Region (Most powerful)

- Neyman-Pearson
  - Simple H_0, Simple H_1
- L-R criteria
  - Composite H_0 <u>or</u> Composite H_1.

Fall 04 Q5

$$\begin{bmatrix} Y \\ X \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_Y \\ \mu_X \end{bmatrix}, \begin{bmatrix} \sigma_Y^2 & \rho\sigma_Y\sigma_X \\ \rho\sigma_Y\sigma_X & \sigma_X^2 \end{bmatrix} \right)$$

a) Let $u = (Y - \mu_Y) - \frac{\rho\sigma_Y}{\sigma_X}(X - \mu_X)$

$var(ax+by) = a^2 var(x) + b^2 var y + 2ab\, cov(x,y)$.

① Show $u \perp\!\!\!\perp X - \mu_X$

Hint: write $Y = \mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(X - \mu_X) + u$

Then $(Y \mid X = x) \sim N\left(\mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(x - \mu_X), Var(u)\right)$

② Compute $E[Y^2 \mid X = x]$

$$\begin{bmatrix} u \\ X - \mu_X \end{bmatrix} = \begin{bmatrix} 1 & -\rho\frac{\sigma_Y}{\sigma_X} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Y - \mu_Y \\ X - \mu_X \end{bmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & -\rho\frac{\sigma_Y}{\sigma_X} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_Y^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_X^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\rho\frac{\sigma_Y}{\sigma_X} & 1 \end{bmatrix} \right)$$

$$= N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_Y^2 - \rho^2\sigma_Y^2 & 0 \\ 0 & \sigma_X^2 \end{bmatrix} \right)$$

② $E[Y^2 \mid X = x] = Var(u) + (E[Y])^2$

$= (1-\rho^2)\sigma_Y^2 + \mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(X - \mu_X)$

Qualifying 503 Q4

$X_i \overset{iid}{\sim} N(\mu, 1)$ $\quad H_0: \mu = 0$
$\quad H_1: \mu > 0$

You are suggested a testing strategy where H_0 is rejected iff $S = \frac{1}{n}\sum_{i=1}^{n} 1(X_i > 0) - \frac{1}{2} \geq \frac{1.96}{\sqrt{n}} \cdot \frac{1}{2}$

① What is the exact probability of type I error when $n=4$?

② What is the limit of probability of type I error when $n \to \infty$?

Soln:

1] $\Pr[\text{type I}] = \Pr[\text{reject } H_0 \mid H_0 \text{ is true}]$

$$= \Pr\left[\frac{1}{n}\sum_{i=1}^{n} 1(X_i > 0) - \frac{1}{2} \geq \frac{1.96}{\sqrt{n}} \cdot \frac{1}{2} \,\middle|\, H_0\right]$$

$$= \Pr\left[\frac{1}{4}\sum_{i=1}^{4} 1(X_i > 0) \geq 0.49 + \frac{1}{2} \,\middle|\, H_0\right]$$

$$= \Pr\left[\sum_{i=1}^{4} 1(X_i > 0) \geq 3.96 \,\middle|\, H_0\right]$$

$$= \Pr\left[(X_1 > 0) \cap (X_2 > 0) \cap (X_3 > 0) \cap (X_4 > 0) \mid H_0\right]$$

$$= \frac{1}{16}$$

2] $\Pr[\text{type I}] = \Pr\left[\frac{1}{n}\sum_{i=1}^{n} 1(X_i > 0) \geq \frac{1.96}{\sqrt{n}} \cdot \frac{1}{2} + \frac{1}{2} \,\middle|\, H_0\right]$

$$\lim_{n \to \infty} \Pr[\text{type I}] = \lim_{n \to \infty} \Pr\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} 1(X_i > 0) - \frac{1}{2}\sqrt{n} \geq 1.96 \times \frac{1}{2} \,\middle|\, H_0\right]$$

$$\sqrt{n}\left(\Pr(X > 0) - \frac{1}{2}\right) \xrightarrow{d} N\left(0, \frac{1}{4}\right)$$

$$\Rightarrow \frac{\sqrt{n}\left(\Pr(X > 0) - \frac{1}{2}\right)}{\frac{1}{2}} \xrightarrow{d} N(0, 1)$$

$$\Rightarrow \lim_{n \to \infty} \Pr\left[\frac{\sqrt{n}\left(\Pr(X > 0) - \frac{1}{2}\right)}{\frac{1}{2}} \geq 1.96 \,\middle|\, H_0\right] = 0.025$$