☆ Convergence in Probability. $Z_n \to_p Z$.

$$P(|Z_n - Z| > \varepsilon) \to 0 \quad \text{as } n \to \infty. \ (\forall \varepsilon > 0)$$

property 2.5. $\begin{cases} X_n \to_p \mu. \\ Y_n \to_p \nu. \end{cases} \Rightarrow \left[X_n + Y_n \Rightarrow_p \mu + \nu\right].$

proof. $\forall \varepsilon > 0$. $P(|X_n - \mu| > \frac{\varepsilon}{2}) \to 0$

$$P(|Y_n - \nu| > \frac{\varepsilon}{2}) \to 0$$

$$|X_n - \mu + Y_n - \nu| \le |X_n - \mu| + |Y_n - \nu|$$

$$\Rightarrow \cancel{P\left(\left||X_n + Y_n - (\mu+\nu)\right|\right| > \varepsilon\right)}$$

$$P(|X_n + Y_n - (\mu+\nu)| > \varepsilon)$$

$$\le P(|X_n - \mu| + |Y_n - \nu| > \varepsilon)$$

$$\le P(|X_n - \mu| > \frac{\varepsilon}{2}) + P(|Y_n - \nu| > \frac{\varepsilon}{2})$$

$$\to 0.$$

P.0

chapter 3.

$\boxed{\text{population}}$ entire group that you want to draw conclusions about.

$\boxed{\text{Sample}}$ the specific group that you will collect data from.

population eg.

Undergraduate students in HKU.

$\boxed{\text{population parameter.}}$

A measure that describes the whole population.

eg. the mean political attitude rating of all undergraduate in HKU.

Sample. eg.

300 undergraduates randomly drawn.

$\boxed{\text{Sample statistic}}$

A measure that describes the sample.

eg. 300. students. Survey.

rating 1, 2, -- 7.

Average rating. 3.2. (sample mean).

Chapter 3.

random variable $X \sim$ pdf $f(x;\underset{\sim}{\theta})$

$\underset{\sim}{\theta} \in \mathbb{R}^s$. unknown parameter. $\underset{\sim}{\theta} \in \Omega$.

Find a good estimator of $\underset{\sim}{\theta}$.

A random sample $\underset{\sim}{X} = \{X_1, \cdots X_n\}$

☆ $\boxed{\text{Properey 3.1}}$

$MSE(\hat{\theta}) = E[(\hat{\theta}-\theta)^2]$

$= E\left(\left\{[\hat{\theta}-E(\hat{\theta})]+[E(\hat{\theta})-\theta]\right\}^2\right)$

$= E\left(\left\{[\hat{\theta}-E(\hat{\theta})]^2 + 2[\hat{\theta}-E(\hat{\theta})][E(\hat{\theta})-\theta] + [E(\hat{\theta})-\theta]^2\right\}\right)$

$= E([\hat{\theta}-E(\hat{\theta})]^2) + 2\underbrace{E[\hat{\theta}-E(\hat{\theta})]}[E(\hat{\theta})-\theta] + [E(\hat{\theta})-\theta]^2$

$= Var(\hat{\theta}) + 0 + [Bias(\hat{\theta})]^2$

$= Var(\hat{\theta}) + [Bias(\hat{\theta})]^2$

☆ $\boxed{\text{Ex 3.8}}$

$f(x_1, x_2, \cdots x_n; \theta)$

$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right\}$

$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right\}$

$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\sum_{i=1}^{n}\frac{(x_i-\bar{x}+\bar{x}-\mu)^2}{2\sigma^2}\right\}$

$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{i=1}^{n}2(x_i-\bar{x})(\bar{x}-\mu) + n(\bar{x}-\mu)^2\right]\right\}$

$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2\right)\right)$

$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right) \exp\left(-\frac{n(\bar{x}-\mu)^2}{2\sigma^2}\right)$

$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{n}{2\sigma^2}S^2\right) \exp\left(-\frac{n(\bar{x}-\mu)^2}{2\sigma^2}\right)$

☆ $S^2 = \overline{X^2} - (\overline{X})^2$.

proof. $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$

$= \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2\overline{X}X_i + (\overline{X})^2)$

$= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\frac{\overline{X}}{n}\sum_{i=1}^{n}X_i + \frac{1}{n}\cdot n(\overline{X})^2$

$= \overline{X^2} - 2(\overline{X})^2 + (\overline{X})^2$

$= \overline{X^2} - (\overline{X})^2$.

☆ Rao-Blackwell Theorem.

Example. $X_1 \text{---} X_n \overset{iid}{\sim}$ Uniform $[0, \theta]$.

θ is unknown parameter.

A sufficient statistic ~~of~~ for θ is $\max_{1\le i\le n} X_i$

Start with the unbiased estimator $\tilde{\theta} = 2X_1$

$E(\tilde{\theta}) = 2E(X_1) = 2\cdot\frac{\theta}{2} = \theta$.

Rao-Blackwell gives

$W(t) = E(\tilde{\theta} \mid T = t)$

$= E(2X_1 \mid \max_{1\le i\le n} X_i = t)$

$= 2E(X_1 \mid \max_{1\le i\le n} X_i = t)$

$= 2\left(\frac{1}{n}\cdot t + \frac{n-1}{n}\cdot\frac{t}{2}\right)$

$= \frac{n+1}{n}t$.

$X_1 = X_{(n)}$ w.p. $\frac{1}{n}$.

$X_1 < X_{(n)}$ w.p. $\frac{n-1}{n}$.

Then $E(X_1) = \frac{t}{2}$.

So $\hat{\theta}_* = W(T) = \frac{n+1}{n}X_{(n)}$ is also unbiased estimator of θ, and

$Var(\hat{\theta}_*) \le Var(\tilde{\theta})$

Page 31. $\begin{cases} Var(\hat{\theta}_*) = \frac{\theta^2}{n(n+2)}. \\ Var(\tilde{\theta}) = 4\times\frac{1}{12n} = \frac{1}{3n} \end{cases}$

☆ Example that UMVUE doesn't achieve CRLB.

$X_1, \dots, X_n \overset{iid}{\sim} N(\theta, 1)$.

The UMVUE for θ^2 is $(\bar{X})^2 - \frac{1}{n}$.

$$\begin{cases} I(\theta^2) = -E\left(\dfrac{\partial^2 \log f(x;\theta^2)}{\partial \theta^4} \right) \\ \text{CRLB} = \dfrac{1}{nI(\theta^2)}. \end{cases}$$

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

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

$$\frac{\partial^2 \log f(x;\theta^2)}{\partial \theta^4} = x \cdot \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot (\theta^2)^{-\frac{3}{2}} = -\frac{x}{4} (\theta^2)^{-\frac{3}{2}}.$$

$$I(\theta^2) = \frac{\theta}{4} \cdot (\theta^2)^{-\frac{3}{2}} = \frac{1}{4}\theta^{-2}.$$

$$\text{CRLB} = \frac{1}{n \cdot \frac{1}{4} \cdot \theta^{-2}} = \frac{4\theta^2}{n}.$$

$$\text{Var}\left((\bar{X})^2 - \frac{1}{n}\right) = \text{Var}\left((\bar{X})^2\right)$$

$$= E\left((\bar{X})^4\right) - \left[E(\bar{X})^2\right]^2$$

$$= \frac{4\theta^2}{n} + \frac{2}{n^2}.$$

$\bar{X} \sim N(\theta, \frac{1}{n})$.

$$\begin{cases} E(\bar{X})^2 = \theta^2 + \frac{1}{n}. \\ E(\bar{X})^4 = \theta^4 + 6 \cdot \theta^2 \cdot \frac{1}{n} + 3 \cdot \frac{1}{n^2} \end{cases}$$

So $\text{Var}\left((\bar{X})^2 - \frac{1}{n}\right) >$ CRLB of θ^2.

☆ Review of Chapeer 3.

3.1. Maximum likelihood Estimator.

Definition. △ Likelihood function.

△ Maximum likelihood Estimator

△ Maximum likelihood Estimate.

eg1. Uniform distribution. $U[0,\beta]$.

pdf. $f(x;\beta) = \begin{cases} \frac{1}{\beta} & 0 \le x \le \beta \\ 0 & o/w. \end{cases}$

$$= \frac{1}{\beta} I(0 \le x \le \beta).$$

Likelihood function.

$$L(\beta) = \prod_{i=1}^{n} \frac{1}{\beta} I(0 \le X_i \le \beta) = \frac{1}{\beta^n} \prod_{i=1}^{n} I(0 \le X_i \le \beta)$$

MLE: $\hat{\beta} = X_{(n)}$

UE: $\frac{n+1}{n} X_{(n)}$ is unbiased est. of β.

UE: $2\bar{X}$.

$$Var\left(\frac{n+1}{n} Y\right) = \frac{\beta^2}{n(n+2)}.$$

$$Var(2\bar{X}) = \frac{\beta^2}{3n}.$$

$$Eff\left(\frac{n+1}{n} Y, 2\bar{X}\right) = \frac{n+2}{3}.$$

△ Unif (α, β).

$$f(x_1, \ldots x_n; \underset{\sim}{\theta}) = \left(\frac{1}{\beta-\alpha}\right)^n I(\alpha \le X_{(1)}) I(X_{(n)} \le \beta)$$

$(X_{(1)}, X_{(n)})$ is a sufficient statistic for (α,β)

△ Note uniform distribution is not belonging to exponential family!

3.2. Method of Moments estimator.

$\Theta \in \mathbb{R}^s$, $\quad \theta = h(\mu_1, \mu_2, \cdots \mu_k)$.

$h: \mathbb{R}^k \to \mathbb{R}^s$, $\quad \mu_r = EX^r$.

$m_r = \frac{1}{n} \sum_{i=1}^{n} x_i^r$

△ Methods of moments estimator (MME)

$\tilde{\theta} = h(m_1, m_2, \cdots m_k)$.

3.3. Estimator properties.

3.3.1. Unbiasedness.

△ $\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$.

△ Asymptotically unbiased: $\lim_{n\to\infty} \text{Bias}(\hat{\theta}) = 0$

3.3.2. Efficiency.

△ $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$

△ $MSE(\hat{\theta}) = Var(\hat{\theta}) + [\text{Bias}(\hat{\theta})]^2$

△ Efficiency. $Eff(\hat{\theta}, \tilde{\theta}) = \frac{Var(\tilde{\theta})}{Var(\hat{\theta})}$.

△ UMVUE: uniformly minimum variance estimator

$UMVUE = \underset{\tilde{\theta} \text{ is unbiased}}{\arg\min}\, Var(\tilde{\theta})$.

△ sufficient statistic.

- $f(x_1, \cdots x_n; \underset{\sim}{\theta}) = g(T(x_1, \cdots x_n); \underset{\sim}{\theta})\, h(x_1, \cdots x_n)$

  $T(x_1, \cdots x_n)$ is sufficient for $\underset{\sim}{\theta}$.

  $T(x_1, \cdots x_n)$ is sufficient statistic for $\underset{\sim}{\theta}$.

  $\left\{ \begin{array}{l} T(\underset{\sim}{x}) \text{ is sufficient statistic for } \underset{\sim}{\theta} \\ \nu(T) = \nu(T(\underset{\sim}{x})) \text{ is also a sufficient - } \underset{\sim}{\theta}. \\ \nu(\cdot) \text{ invertible}. \end{array} \right.$

- $f(x; \underset{\sim}{\theta}) = h(x)\, c(\underset{\sim}{\theta}) \exp\left\{ \sum_{i=1}^{s} p_i(\underset{\sim}{\theta})\, t_i(x) \right\}$.

  Then $T(\underset{\sim}{x}) = \left(\sum_{j=1}^{n} t_1(x_j), \cdots, \sum_{j=1}^{n} t_s(x_j) \right)$

  is a sufficient statistic for $\underset{\sim}{\theta}$.

  <Exponential family>

- Rao-Blackwell Theorem.

$\tilde{\theta}$ unbiased for θ.

$T(\underset{\sim}{x})$ sufficient for θ.

$w(t) = E(\tilde{\theta} \mid T = t)$.

Then $W(T)$ is unbiased for θ and

$Var(W(T)) \le Var(\tilde{\theta})$.

- If $\hat{\theta}$ is UMVUE of θ, then $\hat{\theta}$ is unique.

△ Complete Statistic.

For a given sample $\underset{\sim}{x}$, $T(\underset{\sim}{x})$ is a complete statistic of $\underset{\sim}{\theta}$ if

$E[Z(T)] = 0$ for all $\underset{\sim}{\theta}$ $\Rightarrow P(Z(T)=0) = 1$ $\forall \theta$.

for any measurable function. $Z(\cdot)$.

△ property 3.4. Exponential family.

$f(x;\theta) = h(x)\, c(\theta) \exp\{\sum_{i=1}^{s} p_i(\theta)\, t_i(x)\}$.

If parameter space Θ contains an open set in $\mathbb{R}^s$, then

$T(\underset{\sim}{x}) = \left(\sum_{j=1}^{n} t_1(x_j), \cdots, \sum_{j=1}^{n} t_s(x_j)\right)$ is

a complete statistic for $\underset{\sim}{\theta}$.

△ Theorem. $T(\underset{\sim}{X})$ complete & sufficient.

$\varphi(T)$ is the unique UMVUE of $E[\varphi(T)]$

☆ Example Normal distribution $N(\theta_1, \theta_2)$.

△ $L(\underset{\sim}{\theta}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} \exp\left(-\frac{(x_i-\theta_1)^2}{2\theta_2}\right)$

$l(\underset{\sim}{\theta}) = -\frac{n}{2}\log(2\pi\theta_2) - \frac{1}{2\theta_2}\sum_{i=1}^{n}(X_i-\theta_1)^2$

△ MLE: $\hat{\theta}_1 = \bar{X}$, $\hat{\theta}_2 = S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$.

△ $E(\bar{X}) = \theta_1$.

$E(\hat{\theta}_2) = E(S^2) = \frac{n-1}{n}\theta_2$.

△ Sufficient statistic for (θ_1, θ_2)

$\begin{cases} (\bar{X}, S^2) \\ (\bar{X}, \overline{X^2}) \end{cases}$.

△ complete statistic for (θ_1, θ_2)

$(\bar{X}, S^2)$

△ $\begin{cases} \text{UMVUE of } \theta_1 \text{ is } \bar{X}. \\ \text{UMVUE of } \theta_2 \text{ is } \frac{n}{n-1}S^2. \end{cases}$

△ $\underset{\sim}{X}$ iid poisson(λ)

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

$T(\underset{\sim}{X}) = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is sufficient & complete

$\bar{X} \to$ UMVUE of λ.

△ $\underset{\sim}{X} \sim$ Bernoulli distribution. (p).

$L(p) = \prod_{i=1}^{n}\left[p^{x_i}(1-p)^{1-x_i}\right]$

MLE: $\hat{p} = \bar{X}$.

UE: $E(\hat{p}) = p$.

Sufficient & Complete: $\bar{X}$.

UMVUE: $\bar{X}$.

☆ Example of Sufficient but NOT Complete

Let $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, \theta^2)$. $\theta > 0$.

$T = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is sufficient statistic of (θ, θ^2).

But T is not a complete statistic.

proof. ① $f(x; \theta) = \frac{1}{\sqrt{2\pi\theta^2}} \exp\{-\frac{(x-\theta)^2}{2\theta^2}\}$

$$= \frac{1}{\sqrt{2\pi\theta^2}} \exp\{-\frac{x^2}{2\theta^2} - \frac{\theta^2}{2\theta^2} + \frac{x\theta}{2\theta^2}\}$$

$$= \frac{1}{\sqrt{2\pi\theta^2}} \exp\{-x^2 \cdot \frac{1}{2\theta^2} + x \cdot \frac{1}{2\theta}\} \exp\{-\frac{1}{2}\}.$$

By Property 3.2.

$T(\underset{\sim}{x}) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ is sufficient statistic for (θ, θ^2).

② Parameter Space. $\Omega = \{(\theta, \theta^2) : \theta \neq 0\}$.

parameter space does not contain open set.

So property 3.4 does not apply.

In fact, we should prove this using the definition of completeness.

i.e. we should find a function g of two variables with zero expectation $E[g(T)] = 0$ for all θ and s.t. $P(g(T) = 0) \neq 1$.

How to find it?

First consider. $E[\sum_{i=1}^n X_i] = n\theta$.

$$E[(\sum_{i=1}^n X_i)^2] = Var(\sum_{i=1}^n X_i) + [E(\sum_{i=1}^n X_i)]^2$$

$$= n \cdot \theta^2 + (n\theta)^2 = n\theta^2(1+n) \quad (*)$$

$$E[\sum_{i=1}^n X_i^2] = \sum_{i=1}^n E(X_i^2) = n(\theta^2 + \theta^2) = 2n\theta^2. \quad (\#)$$

If we divide (*) by $n+1$ and substract

(#) divided by 2, we obtain zero for any θ.

So the function $g(x, y)$ can be $\frac{x^2}{n+1} - \frac{y}{2}$.

$$g(T) = \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n+1} - \frac{\sum_{i=1}^{n} X_i^2}{2}.$$

Then $E(g(T)) = 0$, while

$$P(g(T) = 0) = P\left(\frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n+1} - \frac{\sum_{i=1}^{n} X_i^2}{2} = 0\right) \neq 1.$$

Why?

(Hint: $\frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n+1} - \frac{\sum_{i=1}^{n} X_i^2}{2}$ is a random variable

(continuous). $P(\text{r.v.} = 0) = 0$.)

☆ (Fisher Information)

$$I_n(\theta) = E\left[\left(\frac{\partial l(\theta)}{\partial \theta}\right)^2\right].$$

☆ Theorem 3.3. $\underset{\sim}{X}$. sample $f(x;\theta)$

① $I_n(\theta) = nI(\theta)$,

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

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

③ Cramer-Rao Inequality.

$$Var(\hat{\theta}) \geq \frac{1}{I_n(\theta)}.$$

where $\hat{\theta}$ is UE of θ.

proof. ①. $E\left[\frac{\partial \log f(X;\theta)}{\partial \theta}\right] = E\left[\frac{\frac{\partial}{\partial \theta} f(X;\theta)}{f(X;\theta)}\right]$

$$= \int \frac{\frac{\partial}{\partial \theta} f(x;\theta)}{f(x;\theta)} \cdot f(x;\theta)\, dx$$

$$= \int \frac{\partial}{\partial \theta} f(x;\theta)\, dx$$

$$= \frac{\partial}{\partial \theta} \int f(x;\theta)\, dx = \frac{\partial}{\partial \theta} 1 = 0.$$

$$I_n(\theta) = E\left[\left(\frac{\partial l(\theta)}{\partial \theta}\right)^2\right]$$

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

$$= E\left[\sum_{i=1}^n \left(\frac{\partial \log f(X_i;\theta)}{\partial \theta}\right)^2\right] + E\left[\sum_{\substack{i=1\\j=1}}^n \sum_{i\neq j} \frac{\partial \log f(X_i;\theta)}{\partial \theta} \frac{\partial \log f(X_j;\theta)}{\partial \theta}\right]$$

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

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

$$= nI(\theta).$$

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

$$= \int \left\{\frac{\frac{\partial^2}{\partial \theta^2} f(x;\theta)}{f(x;\theta)} - \frac{1}{f^2(x;\theta)}\left[\frac{\partial f(x;\theta)}{\partial \theta}\right]^2\right\} f(x;\theta)\, dx$$

$$= \int \frac{\partial^2}{\partial \theta^2} f(x;\theta)\, dx - \int \frac{1}{f(x;\theta)}\left[\frac{\partial f(x;\theta)}{\partial \theta}\right]^2 dx$$

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

$$= 0 - I(\theta)$$

$$= -I(\theta)$$

$f(\underline{\theta}) = f(x_1, x_2, \cdots x_n; \theta) = f(x_1;\theta) \cdots f(x_n;\theta)$

is the joint p.d.f of $\underline{X}$.

For any unbiased estimator $\hat{\theta}$,

$\hat{\theta} = g(\underline{X}) = g(x_1, \cdots, x_n)$ for some functional g.

then $0 = E(\hat{\theta} - \theta) = \int \cdots \int [g(x_1, \cdots x_n) - \theta] f(x_1, \cdots x_n; \theta)\, dx_1 \cdots dx_n$

$\frac{\partial}{\partial\theta}$. $0 = \int \cdots \int \left\{ (-1) f(\underline{\theta}) + [g(x_1, \cdots, x_n) - \theta] \frac{\partial f(\theta)}{\partial \theta} \right\} dx_1 \cdots dx_n$.

$1 = \int \cdots \int [g(x_1, \cdots, x_n) - \theta] \frac{\partial f(\theta)}{\partial \theta}\, dx_1 \cdots dx_n$.

Cauchy-Schwarz. Inequliey. $\left[\int s_1(x) s_2(x) dx\right]^2 \leq \int s_1^2(x)\, dx \int s_2^2(x)\, dx$.

$$1 = \left[\int \cdots \int [g(x_1, \cdots x_n) - \theta] \frac{\partial f(\theta)}{\partial \theta}\, dx_1 \cdots dx_n \right]^2$$

$$= \left\{ \int \cdots \int [g(x_1, \cdots x_n) - \theta] \sqrt{f(\theta)} \left[\frac{1}{\sqrt{f(\theta)}} \frac{\partial f(\theta)}{\partial \theta} \right] dx_1 \cdots dx_n \right\}^2$$

$$= \left\{ \int \cdots \int [g(x_1, \cdots x_n) - \theta] \sqrt{f(\theta)} \left[\sqrt{f(\theta)} \frac{\partial \lg f(\theta)}{\partial \theta} \right] dx_1 \cdots dx_n \right\}^2$$

$$\leq \int \cdots \int (g(x_1 \cdots x_n) - \theta)^2 f(\theta)\, dx_1 \cdots dx_n \int \cdots \int f(\theta) \left[\frac{\partial \lg f(\theta)}{\partial \theta} \right]^2 dx_1 \cdots dx_n$$

$$= Var(\hat{\theta}) \cdot E\left[\frac{\partial \lg f(\theta)}{\partial \theta} \right]^2 = Var(\hat{\theta}) \cdot I_n(\theta) \quad \Rightarrow Var(\hat{\theta}) \geq \frac{1}{I_n(\theta)}.$$

Remark 3.4

"=" Holds iff there exists a constant A such that

$$A[g(x_1, - x_n) - \theta] \sqrt{f(\theta)} = \sqrt{f(\theta)} \frac{\partial \log f(\theta)}{\partial \theta}.$$

$$\Leftrightarrow A[g(x_1, - x_n) - \theta] = \frac{\partial \log f(x_1, x_2, - x_n; \theta)}{\partial \theta} \quad \text{w.p. } 1.$$

$$\Leftrightarrow A[\hat{\theta} - \theta] = \frac{\partial \log L(\theta)}{\partial \theta}. \quad \text{w.p. } 1$$

attainable condition for CRLB.

If $\hat{\theta}$ can achieve CRLB, then $\hat{\theta}$ must satisfies above attainable condition.