CHAPTER 3

PARAMETRIC POINT ESTIMATION

§3.1 METHODS OF FINDING ESTIMATORS

Let $X_1,\ldots,X_n$ be a random sample from a density $f(\cdot\,;\,\theta)$, where the form of f is known and the parameter θ is unknown. Assume that $\theta=(\theta_1,\ldots,\theta_k)$. Our goal in this section is to find statistics to estimate some functions, say, $\tau_1(\theta),\ldots,\tau_r(\theta)$ of $\theta=(\theta_1,\ldots,\theta_k)$.

Def 3.1 : Any statistic whose values are used to estimate some function of θ, say $\tau(\theta)$, is defined to be an estimator of $\tau(\theta)$.

- Note that an estimator is a statistic which is both a random variable and a function.
- Let $\mathrm{T} = \mathscr{t}\left(X_1, \ldots, X_n\right)$ be an estimator of $\tau\left(\theta\right)$.

  A value of the function $t = \mathscr{t}\left(x_1, \ldots, x_n\right)$ is always called an estimate of $\tau\left(\theta\right)$.

EX 3.1 : Estimators : $\bar{X}, S^2$; Estimates : $\bar{x}, s^2$

- The notation $\hat{\theta} = \left(\hat{\theta}_1, \ldots, \hat{\theta}_k\right)$ is used to denote the

estimate of $\theta = \left(\theta_1, \ldots, \theta_k\right)$, where $\hat{\theta}_j$ estimate θ_j, $j = 1, \ldots, k$. $\hat{\Theta} = \hat{\vartheta}\left(X_1, \ldots, X_n\right)$ is the corresponding estimator.

§3.1.1 METHODS OF MOMENTS

Let X be a random variable with density $f(\cdot\,;\,\theta_1,\ldots,\theta_k)$. Recall that μ'_r is the rth moment about 0; i.e., $\mu'_r = E(X^r)$. Usually, μ'_r is a function of $\theta_1,\ldots,\theta_k$; i.e., $\mu'_r = \mu'_r(\theta_1,\ldots,\theta_k)$. Let $x_1,\ldots,x_n$ be the observations of a random sample $X_1,\ldots,X_n$ from $f(\cdot\,;\,\theta_1,\ldots,\theta_k)$, and let $m'_j = \frac{1}{n}\sum_{i=1}^{n} x_i^j$ denote the jth sample moment of $x_1,\ldots,x_n$. Set

$$
\begin{aligned}
m'_1 &= \mu'_1(\theta_1,\ldots,\theta_k),\\
m'_2 &= \mu'_2(\theta_1,\ldots,\theta_k),\\
&\vdots\\
m'_k &= \mu'_k(\theta_1,\ldots,\theta_k).
\end{aligned}
$$

Let $\hat{\theta}_1,\ldots,\hat{\theta}_k$ be the solution of the k equations. Then the estimator $\left(\hat{\Theta}_1,\ldots,\hat{\Theta}_k\right)$ is called the method-of-moment estimator(MME) of $\left(\theta_1,\ldots,\theta_k\right)$.

EX3.2 : $X_1,\ldots,X_n \sim \ iid\ N\left(\mu\ ;\ \sigma^2\right)$

Recall that $\mu'_1 = \mu = E\left(X\right)$, $\sigma^2 = E\left(X^2\right) - \mu^2$

$\Rightarrow \mu'_2 = E\left(X^2\right) = \sigma^2 + \mu^2$

Set $\begin{cases} m'_1 = \mu'_1 = \mu \\ m'_2 = \mu'_2 = \sigma^2 + \mu^2 \end{cases}$

$$\Rightarrow \begin{cases} \mu = \bar{X} \\ \sigma^2 + \mu^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i^2 \end{cases} \Rightarrow \begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} (X_i - \bar{X})^2 \end{cases}$$

So the MMEs for μ and σ^2 are $\bar{X}$ and $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2$.

EX3.3 : $X_1,\ldots,X_n \sim \ iid \ Exp(\theta)$

$$f(x\,;\,\theta)=\frac{1}{\theta}e^{\frac{-x}{\theta}},0\le x<\infty.$$

Set $m_1^{'}=\mu_1^{'}\Rightarrow \bar{X}=\theta$, So $\hat{\theta}=\bar{X}$ is an MME of θ.

Note that if we set $m_2^{'}=\mu_2^{'}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2=2\theta^2\right)$

$\Rightarrow \hat{\theta}=\sqrt{\frac{1}{2n}\sum_{i=1}^{n}X_i^2}$ We notice that the MMEs are not uniquely defined.

§3.1.2 METHODS OF MAXIMUM LIKELIHOOD

Def 3.2 : The likelihood function of n random variables $X_1,\ldots,X_n$ is defined by $L(\theta\,;\,x_1,\ldots,x_n) = f_{X_1,\ldots,X_n}(x_1,\ldots,x_n\,;\,\theta)$, which is considered to be a function of θ. Particularly, if $X_1,\ldots,X_n$ is a random sample from a density $f(x\,;\,\theta)$, then

$$L(\theta\,;\,{}_1,\ldots,x_n) = f(x_1;\theta)\cdots f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta).$$

Def 3.3 : Let $L(\theta) = L(\theta\,;\,x_1,\ldots,x_n)$ be the likelihood function for the random variables $X_1,\ldots,X_n$. If $\hat{\theta}\left(=\hat{\vartheta}(x_1,\ldots,x_n)\right)$ is the value of θ in Ω which maximizes $L(\theta)$, then $\hat{\Theta}=\hat{\vartheta}(X_1,\ldots,X_n)$ is called the maximum likelihood estimator(MLE) of θ.

- Let $X_1,\ldots,X_n$ be a random sample from a density $f(x;\theta)$. Assume that $\theta = (\theta_1,\ldots,\theta_k)$. If the likelihood function $L(\theta)$ $= L(\theta_1,\ldots,\theta_k) = \prod_{i=1}^{n} f(x_i\ ;\ \theta_1,\ldots,\theta_k)$ is differentiable (in θ_i), then possible maximum likelihood estimate of $(\theta_1,\ldots,\theta_k)$ is the solution of the k equations $\frac{\partial L(\theta_1,\ldots,\theta_k)}{\partial \theta_i} = 0,\ i = 1,\ldots,k.$

Note that $L(\theta)$ and $\log L(\theta)$ have their maxima at the same value of θ (since the log function is strictly increasing on $(0,\infty)$),it is sometimes easier to find the solution of the equations

$$\frac{\partial \log L(\theta_1,\ldots,\theta_k)}{\partial \theta_i} = 0,\ i = 1,\ldots,k.$$

EX3.4 : $X_1,\ldots,X_n \sim \ iid\ N\left(\mu,\sigma^2\right)$, where μ and σ^2 are unknown and $\theta=(\mu,\sigma^2)$.

Then $L\left(\mu,\sigma^2\right)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma}}e^{\frac{-1}{2\sigma^2}\left(x_i-\mu\right)^2}=\frac{1}{\left(2\pi\sigma^2\right)^{n/2}}e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2}$

$\text{Log}L\left(\mu,\sigma^2\right)=\frac{-n}{2}\log 2\pi-\frac{n}{2}\log\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$

Set $\begin{cases}\frac{\partial}{\partial\mu}\log L\left(\mu,\sigma^2\right)=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)=0\\ \frac{\partial}{\partial\sigma^2}\log L\left(\mu,\sigma^2\right)=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i-\mu)^2=0\end{cases}$

We get $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Since $\sum_{i=1}^{n}(X_i - \mu)^2 > \sum_{i=1}^{n}(X_i - \bar{X})^2$ for $\mu \neq \bar{X}$, for any value of σ^2.

$$\frac{1}{\left(2\pi\sigma^2\right)^{n/2}} e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2} \geq \frac{1}{\left(2\pi\sigma^2\right)^{n/2}} e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2}$$

Thus, the likelihood function achieves its global max. at $\left(\hat{\mu}, \hat{\sigma}^2\right)$.

So $\bar{X}$ and $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ are the MLEs.

EX3.5 : $X_1,\ldots,X_n \sim \ iid\ U\left(\theta-\frac{1}{2},\theta+\frac{1}{2}\right),\theta\in\mathbb{R}$

$$f(x\ ;\ \theta)=\begin{cases}1,\ x\in\left[\theta-\frac{1}{2},\theta+\frac{1}{2}\right]=I_{\left[\theta-\frac{1}{2},\theta+\frac{1}{2}\right]}(x)\\ 0,\ \text{o.w}\end{cases}$$

Then $L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}I_{\left[\theta-\frac{1}{2},\theta+\frac{1}{2}\right]}(x_i)=I_{\left[y_n-\frac{1}{2},y_n+\frac{1}{2}\right]}(\theta)$

So any statistic with value $\hat{\theta}$ saytisfying

$y_n-\frac{1}{2}\le\hat{\theta}\le y_1+\frac{1}{2}$ is a MLE.

EX3.6 : $X_1,\ldots,X_n \sim iid\ U(0,\theta),\ 0<\theta<\infty.$

$$f(x\,;\,\theta)=\frac{1}{\theta}I_{[0,\theta]}(x)$$

Then $L(\theta)=\prod_{i=1}^{n}\frac{1}{\theta}I_{[0,\theta]}(x_i)=\frac{1}{\theta^n}I_{[-\infty,y_n]}(\theta).$

Note that $L(\theta)$ is a decreasing function of θ.

Since $\theta \geq y_n$. $L(\theta)$ is maximized at $\hat{\theta}=y_n$.

Thus, Y_n is the unique MLE of θ.

EX 3.7 : $X_1,\ldots,X_n \sim iid\ Cauchy(\theta,1),\ \theta\in\mathbb{R}$

$$f(\theta\,;x)=\frac{1}{\pi\left[1+(x-\theta)^2\right]},\ \theta\in\mathbb{R}$$

$$L(\theta)=\prod_{i=1}^{n}\frac{1}{\pi\left[1+(x_i-\theta)^2\right]},\ \theta\in\mathbb{R}\ =\pi^{-n}\prod_{i=1}^{n}\left[1+(x_i-\theta)^2\right]^{-1}$$

$$\log L(\theta)=-n\log\pi-\sum_{i=1}^{n}\log\left[1+(x_i-\theta)^2\right]$$

Set $\frac{d}{d\theta}\log L(\theta)=0$, we get $\sum_{i=1}^{n}\frac{1}{\left[1+(x_i-\theta)^2\right]}\cdot 2(x_i-\theta)=0$

There are $(2n-1)$ solutions for θ, which are not easy to solve!

Numercial methods are necessary!

EX3.8 : $X_1,\ldots,X_n \sim iid\ N(\theta,1),\ -\infty<\theta<\infty$

Then $L(\theta)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x_i-\theta)^2}=\frac{1}{(2\pi)^{n/2}}e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2}$.

$$\ln L(\theta)=\ln(2\pi)^{-n/2}-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2$$

Set $\frac{d}{d\theta}\ln L(\theta)=\sum_{i=1}^{n}(x_i-\theta)=0\Rightarrow\hat{\theta}=\bar{X}$.

To verify that $\bar{X}$ is a global max. of the likelihood function, recall that $\sum_{i=1}^{n}(x_i-a)^2>\sum_{i=1}^{n}(x_i-\bar{x})^2$ for any number $a\neq\bar{x}$.

This implies that $e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2} < e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ for any $\theta \neq \bar{x}$.

Hence $\bar{X}$ is the MLE. If $\theta \geq 0$, then $L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\theta)^2}$

$= (2\pi)^{-\frac{n}{2}} e^{\left\{\frac{-n}{2}(\bar{x}-\theta)^2 - \frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}}$. If $\bar{X} < 0$, then $L(\theta)$ is a decreasing function for $\theta \geq 0$, and is maximized at $\theta = 0$.

Hence the MLE of θ is $\hat{\Theta} = \begin{cases} \bar{X}, & \bar{X} \geq 0 \\ 0, & \bar{X} < 0 \end{cases}$

Thm 3.1 : (Invariance property of MLE)

If $\hat{\Theta} = \hat{\vartheta}\left(X_1, \ldots, X_n\right)$ is the MLE of θ, then for any function $\tau(\cdot)$, the MLE of $\tau(\theta)$ is $\tau\left(\hat{\Theta}\right)$.

Remark 3.1 : The invariance property of MLE also holds in the multivariate case. That is, if the MLE of $\left(\theta_1, \ldots, \theta_k\right)$ is $\left(\hat{\Theta}_1, \ldots, \hat{\Theta}_k\right)$, then the MLE of any function $\tau\left(\theta_1, \ldots, \theta_k\right)$ is $\tau\left(\hat{\Theta}_1, \ldots, \hat{\Theta}_k\right)$.

EX 3.9 : $X_1, \ldots, X_n \sim \ iid \ N\left(\mu, \sigma^2\right)$

Let $\tau(\theta) = \mu + Z_q \sigma$, where Z_q is given by $\Phi\left(Z_q\right) = q$.

The MLE of $\tau(\theta)$ is $\bar{X} + Z_q \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}$.

Thm 3.2 : Let $X_1,\ldots,X_n$ be a random sample from a density $f(x;\theta)$. If a sufficient statistic $T = \ell(X_1,\ldots,X_n)$ for θ exists and if a MLE $\hat{\Theta}$ of θ also exists uniquely, then $\hat{\Theta}$ is a function of $T = \ell(X_1,\ldots,X_n)$.

Remark 3.2 : If the MLE $\hat{\Theta}$ in Thm3.2 is also sufficient, then by definition, it must be a minimal sufficient statistic.

EX 3.10 : The MLE $\hat{\Theta} = Y_n$ of θ in $U(0,\theta)$ is a minimal sufficient statistic for θ.

EX 3.11 : $X_1,\ldots,X_n \sim \; iid \; U\left(\theta-\frac{1}{2},\theta+\frac{1}{2}\right)$. In EX3.5, we have shown that any statistic with value $\hat{\theta}$ satisfying $y_n-\frac{1}{2}\le\hat{\theta}\le y_1+\frac{1}{2}$ is a MLE .It is easy to show that $\hat{\Theta}=\left(Y_n-\frac{1}{2}\right)+\cos^2 x\left(Y_1-Y_n+1\right)$ is a MLE of θ, but $\hat{\Theta}$ is not a function of Y_1 and Y_n. Also note that the MLE $\frac{Y_1+Y_n}{2}$ is a function of the sufficient statistics Y_1 and Y_n, however, it is not sufficient.

§3.2 METHODS OF EVALUATING ESTIMATORS

In this section, we will introduce some basic criteria, whichan estimator may or may not possess, that will help us in evaluating estimators.

§3.1.1 MEAN SQUARED ERROR

<u>Def 3.4</u> : Let $T = \ell\left(X_1,\ldots,X_n\right)$ be an estimator of $\tau\left(\theta\right)$. The maen squared error (MSE) of T, denoted by $MSE_{\ell}\left(\theta\right)$, is defined to be

$$MSE_{\ell}\left(\theta\right) = E_\theta\left[\left(T-\tau\left(\theta\right)\right)^2\right].$$

Remark 3.3 :

(i) If $f\left(x;\theta\right)$ is the pdf from which the sample was selected, then $E_\theta\left[\left(\mathrm{T}-\tau\left(\theta\right)\right)^2\right]$

$$=\int\cdots\int\left[t\left(x_1,\ldots,x_n\right)-\tau\left(\theta\right)\right]^2\prod_{i=1}^{n}f\left(x_i;\theta\right)dx_1\cdots dx_n.$$

(ii) Since MSE is a function of θ, there will not be one "best" estimator. Often, the MSEs of two estimators will cross each other.

Def 3.5 : An estimator $\mathrm{T}=t\left(X_1,\ldots,X_n\right)$ is defined to be an unbiased estimator of $\tau\left(\theta\right)$ if $E_\theta\left(\mathrm{T}\right)=\tau\left(\theta\right)$, for all $\theta\in\Omega$.

Remark 3.4 : MSE incorporates two components, one measuring the variability of the estimator (precision) and the other measuring its bias (accuracy). More precisely,

$$MSE_{\theta}(\theta) = Var(T) + \left[E_{\theta}(T) - \tau(\theta)\right]^2 = Var(T) + b_{\theta}^2(\theta),$$

where $b_{\theta}(\theta) = E_{\theta}(T) - \tau(\theta)$ is the bias of the estimator T.

If T is an unbiased estimator, then $MSE_{\theta}(\theta) = Var(T)$.

EX 3.12 : $X_1,\ldots,X_n \sim N\left(\mu,\sigma^2\right)$. Recall that the MLEs of

μ and σ^2 are $\bar{X}$ and $\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$, respectively.

$\because E\left(\bar{X}\right)=\mu$. $\therefore \bar{X}$ is an unbiased estimator of μ.

$MSE_{\bar{X}}\left(\mu\right)=E\left[\left(\bar{X}-\mu\right)^2\right]=Var\left(\bar{X}\right)=\frac{\sigma^2}{n}$. Since $E_\theta\left(S^2\right)=\sigma^2$,

$$E_\theta\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=E_\theta\left(\frac{1}{n}\cdot n-1\cdot\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=E_\theta\left(\frac{n-1}{n}S^2\right)$$

$$=\frac{n-1}{n}E_\theta\left(S^2\right)=\frac{n-1}{n}\sigma^2<\sigma^2.$$

So the MLE of σ^2 is not unbiased

$$MSE_{\hat{\sigma}^2}\left(\sigma^2\right) = E_\theta\left\{\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 - \sigma^2\right]\right\}$$

$$= Var\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) + \left\{E_\theta\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) - \sigma^2\right\}^2$$

$$= Var\left(\frac{n-1}{n}S^2\right) + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right) = \left(\frac{n-1}{n}\right)^2 Var\left(S^2\right) + \frac{\sigma^4}{n^2}$$

$$= \frac{\left(n-1\right)^2}{n^2}\cdot\frac{2\sigma^4}{n-1} + \frac{\sigma^4}{n^2} = \frac{\left(2n-1\right)\sigma^4}{n^2}.$$

§3.2.2 LOSS AND RISK FUNCTION

Def 3.2 : Let t denote an estimate of $\tau(\theta)$. The loss function, denoted by $\ell(t;\theta)$, is defined to be a real-valued function satisfying, (i) $\ell(t;\theta) \geq 0$ for all possible estimates t and all $\theta \in \Omega$, and (ii) $\ell(t;\theta) = 0$ for $t = \tau(\theta)$.

EX 3.13 : Some possible loss-functions:

(i) $\ell_1(t;\theta) = \left[t - \tau(\theta)\right]^2 \Rightarrow$ square error

(ii) $\ell_2(t;\theta) = \left[t - \tau(\theta)\right] \Rightarrow$ absolute error

(iii) $\ell_3(t;\theta) = \begin{cases} A, \text{ if } \left|t - \tau(\theta)\right| > \varepsilon \\ 0, \text{ if } \left|t - \tau(\theta)\right| < \varepsilon \end{cases}$ $(A > 0)$

(iv) $\ell_4(t;\theta) = p(\theta)\left|t - \tau(\theta)\right|^r, p(\theta) \geq 0$ and $r > 0$.

- Note that the loss function depends on the estimate $t = \ell\left(x_1, \ldots, x_n\right)$ and hence the samples $X_1, \ldots, X_n$. Thus, we hope to select an estimator that makes the loss small. However, it is impossible to make the loss small for every possible sample. So we may try to pick an estimator that makes the average loss small.

<u>Def 3.2</u> : Give a loss function $\ell\left(\cdot\,;\cdot\right)$. The risk function, denoted by $R_{\ell}\left(t;\theta\right)$, of an estimator $T = \ell\left(X_1, \ldots, X_n\right)$ is defined by $R_{\ell}\left(\theta\right) = E_{\theta}\left[\ell\left(T;\theta\right)\right]$.

Remark 3.5 : Suppose $X_1,\dots,X_n$ is a random sample from a continuous distribution having pdf $f(x;\theta)$. Then

$$E_\theta\left[\ell\left(T;\theta\right)\right]=E_\theta\left[\ell\left(t\left(X_1,\cdots,X_n\right);\theta\right)\right]$$

$$=\int\cdots\int\ell\left(t\left(X_1,\dots,X_n\right);\theta\right)\prod_{i=1}^{n}f\left(x_i;\theta\right)dx_1\cdots dx_n.$$

If we know the density of T, then $E_\theta\left[\ell\left(T;\theta\right)\right]=\int\ell\left(t;\theta\right)f_T\left(t\right)dt.$

EX 3.14 : The corresponding risks in EX3.13 are :

(i) $E_\theta\left[\left(T-\tau\left(\theta\right)\right)^2\right]\Rightarrow$ mean-squared loss

(ii)$E_\theta\left[\left|T-\tau\left(\theta\right)\right|\right]$ $\Rightarrow$ mean-absolute loss

(iii)$A: p\left(\theta\right)\left[\left|T-\tau\left(\theta\right)\right|>\varepsilon\right]$

(iv)$\rho\left(\theta\right): E_\theta\left[\left|T-\tau\left(\theta\right)\right|^r\right]$

Def 3.8 : Let $T_1 = \ell_1(X_1,\ldots,X_n)$ and $T_2 = \ell_2(X_1,\ldots,X_n)$ be two estimators of θ. Then, T_1 is defined to be a better estimator than T_2 iff $R_{\ell_1}(\theta) \le R_{\ell_2}(\theta)$ for all $\theta \in \Omega$ and $R_{\ell_1}(\theta) < R_{\ell_2}(\theta)$ for at least one $\theta \in \Omega$. An estimator $T = \ell(X_1,\ldots,X_n)$ is defined to be admissible if there is no better estimator.

Def 3.9 : $T^* = \ell^*(X_1,\ldots,X_n)$ is said to be a minimax estimator iff $\sup_\theta R_{\ell^*}(\theta) \le \sup_\theta R_\ell(\theta)$, for every estimator $T = \ell(X_1,\ldots,X_n)$.

EX 3.15 : $X_1,\ldots,X_n \sim U(0,\theta), \theta > 0.$

$Y_1,\ldots,Y_n$: order statistics corresponding to $X_1,\ldots,X_n$.

Let $T_1 = Y_n$, $T_2 = \dfrac{n+1}{n} Y_n$, $T_3 = Y_1 + Y_n$, $T_4 = (n+1) Y_1$, $T_5 = 2\bar{X}$ be estimators of θ. We assume that $n \ge 2$, $R_{\prime}(\theta) = MSE_{\prime}(\theta)$.

(1) $\because T_1 = Y_n < \theta$, $\therefore T_1$ is unbiased.

The p.d.f of T_1 is $f_{T_1}(y) = n\left(\dfrac{y}{\theta}\right)^{n-1} \dfrac{1}{\theta} = \dfrac{ny^{n-1}}{\theta^n}, 0 \le y \le \theta.$

So $E(T_1) = \displaystyle\int_0^\theta y \frac{ny^{n-1}}{\theta^n} dy = \frac{n}{\theta^n} \cdot \left.\frac{y^{n+1}}{n+1}\right|_0^\theta = \frac{n}{n+1}\theta \ne \theta.$

$\therefore$ The bias of T_1 is $b_{T_1}(\theta) = E(T_1) - \theta = \dfrac{-1}{n+1}\theta < 0.$

Now $E(T_1^2) = \displaystyle\int_0^\theta y^2 \frac{ny^{n-1}}{\theta} dy = \frac{n}{n+2}\theta^2;$

$$Var\left(T_1\right)=E\left(T_1^2\right)-\left[E\left(T_1\right)\right]^2=\frac{n}{n+2}\theta^2-\left(\frac{n\theta}{n+1}\right)^2=\frac{n\theta^2}{\left(n+1\right)^2\left(n+2\right)}$$

$$\therefore R_{T_1}\left(\theta\right)=MSE_{T_1}\left(\theta\right)=Var\left(T_1\right)+b_{T_1}^2\left(\theta\right)=\frac{n\theta^2}{\left(n+1\right)^2\left(n+2\right)}=\left(\frac{-1}{n+1}\theta\right)^2$$

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

(2) $T_2 = \frac{n+1}{n} Y_n$

$\because E(T_2) = E\left(\frac{n+1}{n} Y_n\right) = \frac{n+1}{n} E(Y_n) = \frac{n+1}{n} \cdot \frac{n}{n+1} \theta = \theta.$

$\therefore T_2$ is unbiased and hence $b_{T_2}(\theta) = 0$. Thus, $R_{T_2}(\theta) = Var(T_2)$

$$= Var\left(\frac{n+1}{n} Y_n\right) = \left(\frac{n+1}{n}\right)^2 Var(Y_n) = \left(\frac{n+1}{n}\right)^2 \frac{n\theta^2}{(n+1)^2 (n+2)}$$

$$= \frac{\theta^2}{n(n+2)}. \quad \frac{R_{T_2}(\theta)}{R_{T_1}(\theta)} = \frac{\theta^2}{n(n+2)} \Bigg/ \frac{2\theta^2}{(n+1)(n+2)} = \frac{n+1}{2n} < 1, \; n \geq 2.$$

$\therefore R_{T_2}(\theta) < R_{T_1}(\theta) \quad \forall \theta > 0. \Rightarrow T_2$ is better than T_1.

$\Rightarrow T_1$ is not admissible. Also note that $R_{T_2}(\theta) \approx \frac{1}{2} R_{T_1}(\theta)$ when n is large.

(3) $T_3 = Y_1 + Y_n$

$$E(T_3) = E(Y_1 + Y_n) = E(Y_1) + E(Y_n) = \frac{\theta}{n+1} + \frac{n\theta}{n+1} = \theta$$

$\Rightarrow T_3$ is unbiased and $b_{T_3}(\theta) = 0.$

It's easy to show that $Var(Y_1) = Var(Y_n) = \dfrac{n\theta^2}{(n+1)^2(n+2)}$.

So $R_{T_3}(\theta) = Var(T_3) = Var(Y_1 + Y_n)$

$= Var(Y_1) + Var(Y_n) + 2\text{cov}(Y_1, Y_n) = 2Var(Y_1) + 2\text{cov}(Y_1, Y_n).$

Note that the joint pdf of Y_1 and Y_n is

$$f_{Y_1,Y_n}(y_1, y_n) = \frac{n(n-1)}{\theta^n}(y_n - y_1)^{n-2}, 0 \le y_1 < y_n < \theta.$$

So cov$\left(Y_1, Y_n\right) = E\left(Y_1 Y_n\right) - E\left(Y_1\right) E\left(Y_n\right)$

$$= \int_0^\theta \int_0^{y_n} y_1 y_n f_{Y_1, Y_n}\left(y_1, y_n\right) dy_1 dy_n$$

$$= \frac{\theta^2}{n+2} - \left(\frac{\theta}{n+1}\right)\left(\frac{n\theta}{n+1}\right) = \frac{\theta^2}{\left(n+1\right)^2\left(n+2\right)}.$$

$$\therefore R_{T_3}\left(\theta\right) = 2\frac{n\theta^2}{\left(n+1\right)^2\left(n+2\right)} + \frac{2\theta^2}{\left(n+1\right)^2\left(n+2\right)}$$

$$= \frac{2\theta^2}{\left(n+1\right)\left(n+2\right)} = R_{T_1}\left(\theta\right) > R_{T_2}\left(\theta\right), n \geq 2.$$

$\therefore T_3$ is inadmissible.

(4) $T_4 = (n+1)Y_1$

$E(T_4) = (n+1)E(Y_1) = (n+1)\dfrac{\theta}{n+1} = \theta \Rightarrow T_4$ is unbaised.

$\therefore R_{T_4}(\theta) = Var(T_4) = Var((n+1)Y_1) = (n+1)^2 Var(Y_1)$

$$= (n+1)^2 \frac{n\theta^2}{(n+1)^2(n+2)} = \frac{n\theta^2}{n+2} . \therefore \frac{R_{T_4}(\theta)}{R_{T_2}(\theta)} = \frac{\dfrac{n\theta^2}{n+2}}{\dfrac{\theta^2}{n(n+2)}}$$

$= n^2 > 1, n \ge 2. \Rightarrow T_4$ is inadmissible.

Note that n^2 grows very fast when n is large.

So T_4 is not a good estimator.

(5) $T_5 = 2\bar{X} \Rightarrow E(T_5) = 2E(\bar{X}) = 2 \cdot \dfrac{\theta}{2} = \theta \Rightarrow T_5$ is unbaised.

$$\therefore R_{T_5}(\theta) = Var(T_5) = Var(2\bar{X}) = 4Var(\bar{X}) = 4 \cdot \frac{\theta^2 / 12}{n} = \frac{\theta^2}{3n}.$$

$$\therefore \frac{R_{T_5}(\theta)}{R_{T_2}(\theta)} = \frac{\dfrac{\theta^2}{3n}}{\dfrac{\theta^2}{n(n+2)}} = \frac{n+2}{3} > 1,\ n \ge 2 \Rightarrow T_5 \text{ is inadmissible.}$$

Summary : $R_{T_2}(\theta) < R_{T_1}(\theta) = R_{T_3}(\theta) < R_{T_5}(\theta) < R_{T_4}(\theta),\ n \ge 2.$

Finally, if $n = 1$, then $Y_n = Y_1 = \bar{X} = X_1$.

So $T_1 = X_1,\ T_2 = T_3 = T_4 = T_5 = 2X_1$. So $R_{T_1}(\theta) = \cdots = R_{T_5}(\theta) = \dfrac{\theta^2}{3}$.

$$T_2 = \frac{n+1}{n} Y_n = \left(1 + \frac{1}{n}\right) Y_n.$$

§3.3 UNBIASED ESTIMATION AND EFFICIENCY

<u>Def 3.10</u> : Let $X_1,\ldots,X_n$ be a random sample from a density $f(x;\theta)$. An estimator $T^* = t^*(X_1,\ldots,X_n)$ of $\tau(\theta)$ is defined to be a uniformly minimum-variance unbiased estimator (UMVUE) of $\tau(\theta)$ iff (i) T^* is unbiased; i.e., $E_\theta(T^*) = \tau(\theta)$, and (ii) $Var_\theta(T^*) \le Var_\theta(T)$ for any other unbiased estimator $T = t(X_1,\ldots,X_n)$ of $\tau(\theta)$.

- Consider a random sample $X_1,\ldots,X_n$ from a pdf $f(x;\theta)$, $\theta \in \Omega$. Assume that Ω is a subset of real line. Let $T = t(X_1,\ldots,X_n)$ be an unbiased estimator of $\tau(\theta)$. We assume that

(i) $f(x;\theta)$ is positive on a set S independent of θ;

(ii) $\dfrac{\partial}{\partial\theta}\log f(x;\theta)$ exists for all x and all θ;

(iii) the integral $\int\cdots\int\prod_{i=1}^{n} f(x_i;\theta)\,dx_1\cdots dx_n$ can be differentiated with respect to θ under the integral sign;

(iv) the integral $\int\cdots\int t(x_1,\ldots,x_n)\prod_{i=1}^{n} f(x_i;\theta)\,dx_1\cdots dx_n$ can be differentiated with respect to θ under the integral sign;

(v) $0 < E_\theta\left(\left(\dfrac{\partial}{\partial\theta}\log f(X;\theta)\right)^2\right) < \infty$ for all $\theta\in\Omega$.

The above assumptions are called regularity conditions.

<u>Thm 3.3</u> : (Cramér-Rao inequality)

Under the regularity conditions, if $T = \ell(X_1, \ldots, X_n)$ is an unbiased estimator of $\tau(\theta)$, then we have $Var_\theta(T) \geq \frac{\left[\tau'(\theta)\right]^2}{nI(\theta)}$, where $I(\theta) = E_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(X;\theta)\right)^2\right)$ for all $\theta \in \Omega$. Equality in Equation (3.4) holds iff there exists a function, say $k(\theta;n)$, such that $\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(x_i;\theta)$ $= k(\theta;n)\left[\ell(x_1, \ldots, x_n) - \tau(\theta)\right]$.

<u>Remark 3.6</u> : The regularity conditions can be modified for discrete density functions, leaving the results of Theorem 3.3 unchanged.

- Theorem 3.3 states that if an unbiased estimator whose variance coincides with the Cramėr-Rao lower bound (CRLB), then this estimator is an UMVUE.
- The quantity $I(\theta)$ is called information number or Fisher information of the sample. In continuous case, $I(\theta) = E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(X;\theta)\right)^2\right]$

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

The information number gives a bound on the variance of the best unbiased estimator of $\tau(\theta)$. As the information number gets larger, we have a smaller bound on the variance of the best unbiased estimator of $\tau(\theta)$.

Lemma 3.1 : If $f(x;\theta)$ satisfies the condition

$$\frac{d}{d\theta}E_\theta\left[\frac{\partial}{\partial\theta}\log f(X;\theta)\right]=\int\frac{\partial}{\partial\theta}\left[\left(\frac{\partial}{\partial\theta}\log f(x;\theta)\right)f(x;\theta)\right]dx,$$ then

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

EX 3.16 : $X_1,\ldots,X_n \sim f(x;\theta)=\theta e^{-\theta x},\ x\in(0,\infty)$.

Let $\tau(\theta)=\theta$. Then $\tau'(\theta)=1$.

$$\because \frac{\partial}{\partial\theta}\log f(x;\theta)=\frac{\partial}{\partial\theta}\log\left(\theta e^{-\theta x}\right)=\frac{\partial}{\partial\theta}\left(\log\theta-\theta x\right)=\frac{1}{\theta}-x \quad \therefore I(\theta)$$

$$=E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(X;\theta)\right)^2\right]=E_\theta\left[\left(\frac{1}{\theta}-x\right)^2\right]=E_\theta\left[\left(x-\frac{1}{\theta}\right)^2\right]=Var(x)=\frac{1}{\theta^2}.$$

So the CRLB for the variance of an unbiased

estimator T of θ is $Var_\theta(T)\geq 1\Big/ n\cdot\frac{1}{\theta^2}=\frac{\theta^2}{n}$. Similarly, if $\tau(\theta)=\frac{1}{\theta}$,

then $\tau'(\theta)=\frac{-1}{\theta^2}$,and $Var_\theta(T)\geq \left(\frac{-1}{\theta^2}\right)^2\Big/ n\cdot\frac{1}{\theta^2}=\frac{1}{n\theta^2}$.

Note that $\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(x;\theta)=\sum_{i=1}^{n}\left(\frac{1}{\theta}-x_i\right)=\frac{n}{\theta}-\sum_{i=1}^{n}x_i=-n\left(\bar{x}-\frac{1}{\theta}\right)$.

We see that $\bar{X}$ is an UMVUE of $\frac{1}{\theta}$. Note : $Var(\bar{X})=\frac{1}{\theta^2}\Big/n$

$=\frac{1}{n\theta^2}$ coincides with the CRLB.

EX 3.17 : $X_1,\ldots,X_n \sim f(x;\theta)=f(x;\lambda)=\frac{e^{-\lambda}\lambda^x}{x!}, x=0,1,\ldots$

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

$$=-1+\frac{x}{\lambda}\quad\therefore I(\lambda)=E_\theta\left[\left(\frac{\partial}{\partial\lambda}\log f(X;\theta)\right)^2\right]$$

$$=E\left[\left(\frac{x}{\lambda}-1\right)^2\right]=\frac{1}{\lambda^2}E\left[(x-\lambda)^2\right]=\frac{1}{\lambda^2}Var(x)=\frac{1}{\lambda^2}\cdot\lambda=\frac{1}{\lambda}.$$

Take $\tau(\lambda) = e^{-\lambda} = P(X = 0)$. Then the CRLB for the variance of an unbiased estimator T of $e^{-\lambda}$ is $Var(T) \geq \left(-e^{-\lambda}\right)^2 \Big/ n \cdot \frac{1}{\lambda}$

$= \frac{\lambda e^{-2\lambda}}{n}$. Note that $T = \frac{1}{n}\sum_{i=1}^{n} I_{\{0\}}(x_i)$ is an unbiased estimator of $e^{-\lambda}$. Since $E(T) = E\left(\frac{1}{n}\sum_{i=1}^{n} I_{\{0\}}(x_i)\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(I_{\{0\}}(x_i)\right)$

$= \frac{1}{n}\sum_{i=1}^{n}\left[1 \cdot P(x_i = 0) + 0 \cdot P(x_i \neq 0)\right] = \frac{1}{n}\sum_{i=1}^{n} e^{-\lambda} = e^{-\lambda}$. T is the proportion of observations in the sample that are equal to 0.

Now, $Var\left(T\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n} I_{\{0\}}\left(x_i\right)\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var\left(I_{\{0\}}\left(x_i\right)\right)$

$= \frac{1}{n^2}\sum_{i=1}^{n}\left(e^{-\lambda} - e^{-2\lambda}\right) = \frac{1}{n}e^{-\lambda}\left(1-e^{-\lambda}\right)$ and $\frac{1}{n}e^{-\lambda}\left(1-e^{-\lambda}\right) \geq \frac{\lambda e^{-2\lambda}}{n}$

$\Rightarrow T$ may not be an UMVUE!! Note that $\sum_{i=1}^{n}\frac{\partial}{\partial\lambda}\log f\left(x_i;\lambda\right)$

$= \sum_{i=1}^{n}\left(-1+\frac{x_i}{\lambda}\right) = -n + \frac{\sum x_i}{\lambda} = \frac{n}{\lambda}\left(\bar{X}-\lambda\right). \therefore \bar{X}$ is an UMVUE of λ.

<u>EX 3.18</u> : $X_1,\ldots,X_n \sim U\left(0,\theta\right). f\left(x;\theta\right) = \frac{1}{\theta}, 0 < x < \infty.$

Since $\frac{\partial}{\partial\theta}\log f\left(x;\theta\right) = \frac{\partial}{\partial\theta}\log\left(\frac{1}{\theta}\right) = \frac{\partial}{\partial\theta}\left(-\log\theta\right) = \frac{-1}{\theta},$

we have $I\left(\theta\right) = E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f\left(X;\theta\right)\right)^2\right] = E_\theta\left(\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}.$

If T is an unbiased estimator of θ, then $Var_\theta(T) \geq 1 \Big/ \frac{n}{\theta^2}$

$= \frac{\theta^2}{n} =$ CRLB.Consider the sufficient statistic $Y_n = \max\left[X_1, \ldots, X_n\right]$.

The pdf of Y_n is $n\left[F(y)\right]^n f(y)$. $f_{Y_n}(y) = n \cdot \left(\frac{y}{\theta}\right)^{n-1} \frac{1}{\theta}$ and $E_\theta(Y_n)$

$= \int_0^\theta y \cdot \frac{y^{n-1} n}{\theta^n} dy = \frac{n\theta}{n+1}$. So, $E_\theta\left(\frac{n+1}{n} Y_n\right) = \theta$, i.e. $\frac{n+1}{n} Y_n$ is an

unbiased estimator of θ. Now, $Var_\theta\left(\frac{n+1}{n} Y_n\right) = \left(\frac{n+1}{n}\right)^2 Var(Y_n)$

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

Clearly, $Var_\theta\left(\frac{n+1}{n}Y_n\right) < $ CRLB. Thus,the Cramėr-Rao inequality is not applicable to this pdf. To see that this is so,

$$\frac{\partial}{\partial\theta}\int_0^\theta \ell(x) f(x;\theta)\,dx = \frac{\partial}{\partial\theta}\int_0^\theta \ell(x)\frac{1}{\theta}\,dx = \frac{\partial}{\partial\theta}\left[\frac{1}{\theta}\int_0^\theta \ell(x)\,dx\right]$$

$$= \left(\frac{-1}{\theta^2}\right)\int_0^\theta \ell(x)\,dx + \frac{1}{\theta}\frac{\partial}{\partial\theta}\int_0^\theta \ell(x)\,dx = \left(\frac{-1}{\theta^2}\right)\int_0^\theta \ell(x)\,dx + \frac{\ell(x)}{\theta}$$

$\neq \int_0^\theta \ell(x)\frac{\partial}{\partial\theta}f(x;\theta)\,dx$. In general, if the range of the pdf depend on θ,the Cramėr-Rao inequality will not be applicable.

Thm3.4 : If $\hat{\theta} = \hat{\vartheta}(x_1,\ldots,x_n)$ is the maximum likelihood estimate of θ, and if $T^* = \ell^*(X_1,\ldots,X_n)$ is an unbiased estimator of $\tau^*(\theta)$ whose variance coincides with the CRLB, then $\ell^*(x_1,\ldots,x_n) = \tau^*\left(\hat{\vartheta}(x_1,\ldots,x_n)\right)$.

EX 3.19 : $X_1,\ldots,X_n \sim N(\mu,\sigma^2)$, $\mu\in\mathbb{R}$, $\sigma>0$, $n\geq 2$.

(1) σ^2 is known : $\tau(\theta)=\mu$.

$$f(x;\theta)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-1}{2\sigma^2}(x-\mu)^2}, x\in\mathbb{R}.$$

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

$$\therefore I(\theta)=E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(X;\theta)\right)^2\right]=E_\theta\left[\left(\frac{x-\mu}{\sigma^2}\right)^2\right]=\frac{1}{\sigma^4}E_\theta\left[(x-\mu)^2\right]$$

$$=\frac{1}{\sigma^4}Var_\theta(X)=\frac{1}{\sigma^4}\sigma^2=\frac{1}{\sigma^2}.$$

So, CRLB $=\dfrac{\left[\tau'(\theta)\right]^2}{n\cdot I(\theta)}=\dfrac{1}{n\left(1/\sigma^2\right)}=\dfrac{\sigma^2}{n}=Var(\bar{X})\Rightarrow \bar{X}$ is an UMVUE of μ.

(2) μ is known : $\tau(\theta) = \sigma^2$.

$$L(\sigma^2; x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-1}{2\sigma^2}(x_i - \mu)^2} = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\frac{\partial}{\partial\sigma^2} \log L(\sigma^2; x_1, \ldots, x_n) = \frac{n}{2\sigma^4}\left(\sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n} - \sigma^2\right) = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \text{ is the MLE of } \sigma^2.$$

$\Rightarrow \hat{\sigma}^2$ is an UMVUE of σ^2 (Thm 3.4).

(3) μ, σ^2 : unknown : Consider estimation of σ^2.

$$\frac{\partial}{\partial(\sigma^2)^2}\left[\log(2\pi\sigma^2)^{-1/2} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x - \mu)^2\right] = \frac{1}{2\sigma^4} - \frac{(x-\mu)^2}{\sigma^6}$$

$$\therefore I(\theta) = -E_\theta\left[\left(\frac{\partial}{\partial(\sigma^2)^2}\log f(X;\theta)\right)^2\right] = -E_\theta\left[\frac{1}{2\sigma^4} - \frac{(x-\mu)^2}{\sigma^6}\right]$$

$$= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}E_\theta\left[(x-\mu)^2\right] = -\frac{1}{2\sigma^4} + \frac{1}{\sigma^6}\cdot\sigma^2 = \frac{1}{2\sigma^4}.$$

$$\therefore \text{CRLB} = \frac{1}{n\left(1/2\sigma^4\right)} = \frac{2\sigma^4}{n}\left(\tau(\theta) = \sigma^2, \therefore \tau'(\theta) = 1\right)$$

Thus, any unbiased estimator T of σ^2 must satisfy $Var(T) \geq \frac{2\sigma^4}{n}$

$$\because \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \quad \therefore Var(S^2) = Var\left(\frac{\sigma^2}{n-1}\frac{n-1}{\sigma^2}S^2\right)$$

$$= \left(\frac{\sigma^2}{n-1}\right)^2 Var\left(\frac{(n-1)S^2}{\sigma^2}\right) = \frac{\sigma^4}{(n-1)^2}\cdot 2(n-1) = \frac{2\sigma^4}{n-1} > \frac{2\sigma^4}{n} = \text{CRLB}.$$

Thm3.5 : If $T^* = \ell^*(X_1,\ldots,X_n)$ is an unbiased estimator of some $\tau^*(\theta)$ whose variance coincides with the CRLB, then $f(\cdot\,;\,\theta)$ is a member of exponential class. Conversely, if $f(\cdot\,;\,\theta)$ is a member of exponential class, then there exists an unbiased estimator, say T^*, of some function of $\tau^*(\theta)$, whose variance coincides with the CRLB (i.e., T^* is an UMVUE of $\tau^*(\theta)$) .

Def 3.11 : The retative efficiency of an unbiased estimator T of $\tau(\theta)$ to another unbiased estimator $T^*$of $\tau(\theta)$ is given by $re(T,T^*) = \dfrac{Var(T^*)}{Var(T)}$.

An unbiased estimator $T^*$of $\tau(\theta)$ is said to be efficient if $re(T,T^*) \le 1$ for all unbiased estimators T of $\tau(\theta)$ and for all $\theta \in \Omega$. The efficiency of an unbiased estimator T of $\tau(\theta)$ is $e(T) = re(T,T^*)$, if T^*an unbiased estimator of $\tau(\theta)$.

EX 3.20 : $X_1,\ldots,X_n \sim f(x;\theta)=\theta e^{-\theta x},\ x\in(0,\infty)$.

In EX3.16, we have shown that $\bar{X}$ is an UMVUE of $\dfrac{1}{\theta}$.

Let $Y_1=\min[X_1,\ldots,X_n]$. Then $f_{Y_1}(y)$

$=n\left[1-\left(1-e^{-\theta y}\right)\right]^{n-1}\left(\theta e^{-\theta y}\right)=(n\theta)e^{-(n\theta)y}$.So, $Y_1 \sim Exp(n\theta)$,

and have $E(Y_1)=\dfrac{1}{n\theta}\left(\text{or } E(nY_1)=\dfrac{1}{\theta}\right)$.Thus, nY_1 is

unbiased for $\dfrac{1}{\theta}$ and $Var(nY_1)=n^2Var(Y_1)=n^2\dfrac{1}{(n\theta)^2}=\dfrac{1}{\theta^2}$.

Recall that $Var(\bar{X})=\dfrac{1}{\theta^2}\Big/n=\dfrac{1}{n\theta^2}$.

We have $e\left(nY_1\right) = re\left(nY_1, \bar{X}\right) = \dfrac{Var\left(\bar{X}\right)}{Var\left(nY_1\right)} = \dfrac{1/n\theta^2}{1/\theta^2} = \dfrac{1}{n}$

(efficiency of nY_1). $\therefore \bar{X}$ is efficient and nY_1 is a very poor estimator of $\dfrac{1}{\theta}$ since its efficiency is small for large n.

§3.4 SUFFICIENCY AND UNBIASEDNESS

In this section, the concept of sufficiency will be used in our search for UMVUEs. Recall that if X and Y are any two random variables,then we have $E(X) = E\left[E(X|Y)\right]$, $Var(X) = Var\left[E(X|Y)\right] + E\left[Var(X|Y)\right]$

These toois are used to prove the following theorem.

<u>**Thm3.6**</u> : (Rao-Blackwell)

Let $X_1,\ldots,X_n$ be a random sample from a density $f(\cdot\,;\,\theta)$, and let $T = t(X_1,\ldots,X_n)$ be an unbiased estimator of $\tau(\theta)$. Let $S = s(X_1,\ldots,X_n)$ be a sufficient statistic. Define $T' = E(T|S)$. Then

(a) T' is a statistic, and it is a function of the sufficient statistic S;

(b) T' is an unbiased estimator of $\tau(\theta)$; that is, $E_\theta(T') = \tau(\theta)$;

(c) $Var_\theta(T') \le Var_\theta(T)$ for every θ.

• The Rao-Blackwell theorem says that, given an unbiased estimator, another unbiased estimator that is a function of sufficient statistic is a uniformly better unbiased estimator of $\tau(\theta)$.

<u>Thm 3.7</u> : (Lemann-Scheffė)

Let $X_1,\ldots,X_n$ be a random sample from a density $f(\cdot\,;\,\theta)$, and let $S = s(X_1,\ldots,X_n)$ be a complete sufficient statistic. If $T^* = t^*(S)$, a function of S, is an unbiased estimator of $\tau(\theta)$, then T^* is an UMVUE of $\tau(\theta)$.

• The Lemann-Scheffė theorem says that if a complete sufficient statistic S exists, then there is an UMVUE of $\tau(\theta)$, and the UMVUE is the unique unbiased estimator of $\tau(\theta)$ which is a function of S.

§3.5 LARGE SAMPLE PROPERTIES

§3.5.1 CONSISTENCY

Def 3.12 : A sequence of estimators $T_n = \ell_n(X_1, \ldots, X_n)$ is defined to be a consistent sequence of estimators of $\tau(\theta)$ if for every $\varepsilon > 0$ and every $\theta \in \Omega$, $\lim_{n\to\infty} P_\theta\left[\left|T_n - \tau(\theta)\right| < \varepsilon\right] = 1$. It is equivalent to $\lim_{n\to\infty} P_\theta\left[\left|T_n - \tau(\theta)\right| \geq \varepsilon\right] = 0$.

Thm 3.8 : If $T_n = \ell_n(X_1, \ldots, X_n)$ is a sequence of estimators of $\tau(\theta)$ satisfying $\lim_{n\to\infty} E_\theta\left[\left(T_n - \tau(\theta)\right)^2\right] = 0$, for every $\theta \in \Omega$, then T_n is a consistent sequence of estimators of $\tau(\theta)$.

- Recall that the MSE of T_n is given by

$$MSE_{T_n}(\theta) = E_\theta\left[\left(T_n - \tau(\theta)\right)^2\right]$$

$$= Var(T_n) + \left[E_\theta(T_n) - \tau(\theta)\right]^2$$

$$= Var(T_n) + b^2_{T_n}(\theta).$$

An equivalent statement is this: For every $\theta \in \Omega$, if T_n satisfies

(i) $\lim_{n\to\infty} Var(T_n) = 0$, and

(ii) $\lim_{n\to\infty} b^2_{T_n}(\theta) = 0$,

then T_n is a consistent sequence of estimators of $\tau(\theta)$.

Thm 3.9 : Let $T_n = \ell_n(X_1, \ldots, X_n)$ be a consistent sequence of estimators of $\tau(\theta)$. Let $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ be sequences of constants satisfying

(i) $\lim_{n\to\infty} a_n = 1$,

(ii) $\lim_{n\to\infty} b_n = 0$,

then the sequence $U_n = a_n T_n + b_n$ is a consistent sequence of estimators of $\tau(\theta)$.

Def 3.13 : A sequence of estimators $T_n = \ell_n(X_1, \ldots, X_n)$ is said to be asymptotically unbiased for $\tau(\theta)$ if $\lim_{n\to\infty} E(T_n) = \tau(\theta)$, for each $\theta \in \Omega$.

§3.5.1 EFFICIENCY

Def 3.14 : Let T_n and T_n^* be asymptotically unbiased sequences of estimators for $\tau(\theta)$. The asymptotically relative efficiency of T_n relative to T_n^* is defined to be $are(T_n, T_n^*) = \lim_{n\to\infty} \frac{Var(T_n^*)}{Var(T_n)}$. The sequence T_n^* is said to be asymptotically efficient if $are(T_n, T_n^*) \leq 1$, for all other asymptotically unbiased sequences T_n, and for all $\theta \in \Omega$. The asymptotically efficiency of T_n is defined to be $ae(T_n) = are(T_n, T_n^*)$.

Def 3.15 : A sequence of estimators T_n of $\tau(\theta)$ is said to be asymptotically normal if $\frac{\sqrt{n}\left[T_n - \tau(\theta)\right]}{\sigma(\theta)} \xrightarrow{d} N[0,1]$, as $n \to \infty$, where $\sigma(\theta)$ is a function of θ.

Def 3.16 : A sequence of estimators T_n of $\tau(\theta)$ is said to be best asymptotically normal (BAN) if it is asymptotically normal and $\sigma^2(\theta) = \frac{\left[\tau'(\theta)\right]}{I(\theta)}$, for all $\theta \in \Omega$, where $I(\theta)$ is the Fisher information.

Remark 3.7 : Under certain regularity conditions, the MLE of $\tau(\theta)$ is BAN.