STATISTICAL INFERENCE

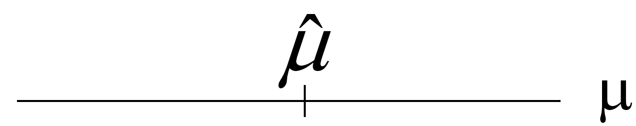
PART I

POINT ESTIMATION

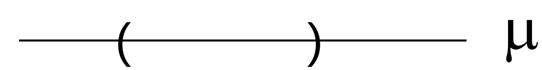
STATISTICAL INFERENCE

- Determining certain unknown properties of a probability distribution on the basis of a sample (usually, a r.s.) obtained from that distribution

Point Estimation: $\hat{\mu} = 5$



Interval Estimation: $3 \leq \mu \leq 8$



Hypothesis Testing: $H_0 : \mu = 5$

$H_1 : \mu \neq 5$

STATISTICAL INFERENCE

- ***Parameter Space (Ω):*** The set of all possible values of an unknown parameter, θ; $\theta \in \Omega$.
- A pdf with unknown parameter: $f(x;\ \theta),\ \theta \in \Omega$.
- **Estimation:** Where in Ω, θ is likely to be?

$\{\ f(x;\ \theta),\ \theta \in \Omega\ \} \longrightarrow$ $\boxed{\text{The family of pdfs}}$

STATISTICAL INFERENCE

- **Statistic:** A function of rvs (usually a sample rvs in an estimation) which does not contain any unknown parameters.

$$\bar{X}, S^2, etc$$

- **Estimator** of an unknown parameter θ: $\hat{\theta}$

A statistic used for estimating θ.

$$\hat{\theta}: estimator = U(X_1, X_2, \cdots, X_n)$$

$$\bar{X}: Estimator$$

$$\downarrow$$

An observed value

$\bar{x}$: *Estimate* : A particular value of an estimator

METHODS OF ESTIMATION

Method of Moments Estimation,
Maximum Likelihood Estimation

METHOD OF MOMENTS ESTIMATION (MME)

- Let $X_1, X_2, \ldots, X_n$ be a r.s. from a population with pmf or pdf $f(x;\theta_1, \theta_2, \ldots, \theta_k)$. The MMEs are found by equating the first k population moments to corresponding sample moments and solving the resulting system of equations.

Population Moments	Sample Moments
$\mu_k = E\left[X^k\right]$	$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

METHOD OF MOMENTS ESTIMATION (MME)

$$\mu_1 = M_1 \qquad \mu_2 = M_2 \qquad \mu_3 = M_3$$

so on…

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad E(X^2) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \qquad E(X^3) = \frac{1}{n}\sum_{i=1}^{n} X_i^3$$

Continue this until there are enough equations to solve for the unknown parameters.

EXAMPLES

- $X \sim Exp(\theta)$. For a r.s of size n, find the MME of θ.

EXAMPLES

- $X \sim Gamma(\alpha, \beta)$. For a r.s of size n, find the MMEs of α and β.

DRAWBACKS OF MMES

- Although sometimes parameters are positive valued, MMEs can be negative.

- If moments does not exist, we cannot find MMEs.

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

- Let $X_1, X_2, \ldots, X_n$ be a r.s. from a population with pmf or pdf $f(x;\theta_1, \theta_2, \ldots, \theta_k)$, the likelihood function is defined by

$$L(\theta_1,\theta_2,\cdots,\theta_k \mid x_1,x_2,\cdots,x_n) = f(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)$$

$$= \prod_{i=1}^{n} f(x_i \mid \theta_1,\theta_2,\cdots,\theta_k)$$

$$= \prod_{i=1}^{n} f(x_i;\theta_1,\theta_2,\cdots,\theta_k)$$

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

- For each sample point $(x_1, \ldots, x_n)$, let

$$\hat{\theta}_1(x_1,\ldots,x_n),\cdots,\hat{\theta}_k(x_1,\ldots,x_n)$$

be a parameter value at which

$L(\theta_1, \ldots, \theta_k | x_1, \ldots, x_n)$ attains its maximum as a function of $(\theta_1, \ldots, \theta_k)$, with $(x_1, \ldots, x_n)$ held fixed. A maximum likelihood estimator (MLE) of parameters $(\theta_1, \ldots, \theta_k)$ based on a sample $(X_1, \ldots, X_n)$ *is*

$$\hat{\theta}_1(x_1,\ldots,x_n),\cdots,\hat{\theta}_k(x_1,\ldots,x_n)$$

- The MLE is the parameter point for which the observed sample is most likely.

EXAMPLES

- Let $X \sim Bin(n,p)$. One observation on X is available, and it is known that n is either 2 or 3 and $p=1/2$ or $1/3$. Our objective is to estimate the pair (n,p).

x	$(2,1/2)$	$(2,1/3)$	$(3,1/2)$	$(3,1/3)$	*Max. Prob.*
0	$1/4$	$4/9$	$1/8$	$8/27$	$4/9$
1	$1/2$	$4/9$	$3/8$	$12/27$	$1/2$
2	$1/4$	$1/9$	$3/8$	$6/27$	$3/8$
3	0	0	$1/8$	$1/27$	$1/8$

$$(\hat{n},\hat{p})(x)=\begin{cases}(2,1/3) \text{ if } x=0\\(2,1/2) \text{ if } x=1\\(3,1/2) \text{ if } x=2\\(3,1/2) \text{ if } x=3\end{cases}$$

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

- It is convenient to work with the logarithm of the likelihood function.
- Suppose that $f(x;\theta_1, \theta_2, \ldots, \theta_k)$ is a positive, differentiable function of $\theta_1, \theta_2, \ldots, \theta_k$. If a supremum $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k$ exists, it must satisfy the likelihood equations

$$\frac{\partial \ln L\left(\theta_1, \theta_2, \cdots, \theta_k; x_1, x_2, \cdots, x_k\right)}{\partial \theta_j} = 0, \; j = 1, 2, \ldots, k$$

- MLE occurring at boundary of Ω cannot be obtained by differentiation. So, use inspection.

EXAMPLES

1. $X \sim Exp(\theta)$, $\theta > 0$. For a r.s of size n, find the MLE of θ.

EXAMPLES

2. $X \sim N(\mu, \sigma^2)$. For a r.s. of size n, find the MLEs of μ and σ^2.

EXAMPLES

3. $X \sim Uniform(0,\theta)$, $\theta>0$. For a r.s of size n, find the MLE of θ.

EXAMPLES

4. $X \sim Uniform(\theta, \theta+1)$, $\theta > 0$. For a r.s of size n, find the MLE of θ.

INVARIANCE PROPERTY OF THE MLE

- If $\hat{\theta}$ is the MLE of θ, then for any function $\tau(\theta)$, the MLE of $\tau(\theta)$ is $\tau\left(\hat{\theta}\right)$.

Example: $X \sim N(\mu,\sigma^2)$. For a r.s. of size n, the MLE of μ is $\bar{X}$. By the invariance property of MLE, the MLE of μ^2 is $\bar{X}^2$.

ADVANTAGES OF MLE

- Often yields good estimates, especially for large sample size.
- Usually they are consistent estimators.
- Invariance property of MLEs
- Asymptotic distribution of MLE is Normal.
- Most widely used estimation technique.

DISADVANTAGES OF MLE

- Requires that the pdf or pmf is known except the value of parameters.
- MLE may not exist or may not be unique.
- MLE may not be obtained explicitly (numerical or search methods may be required.). It is sensitive to the choice of starting values when using numerical estimation.
- MLEs can be heavily biased for small samples.
- The optimality properties may not apply for small samples.

SOME PROPERTIES OF ESTIMATORS

- **UNBIASED ESTIMATOR (UE):** An estimator $\hat{\theta}$ is an UE of the unknown parameter θ, if

$$E\left[\hat{\theta}\right] = \theta \ for\ all\ \theta \in \Omega$$

Otherwise, it is a **Biased Estimator** of θ.

$$\underset{\theta}{Bias}\left(\hat{\theta}\right) = E\left[\hat{\theta}\right] - \theta \longrightarrow \boxed{\text{Bias of } \hat{\theta} \text{ for estimating } \theta}$$

If $\hat{\theta}$ is UE of θ, $\underset{\theta}{Bias}\left(\hat{\theta}\right) = 0.$

SOME PROPERTIES OF ESTIMATORS

- **ASYMPTOTICALLY UNBIASED ESTIMATOR (AUE):** An estimator $\hat{\theta}$ is an AUE of the unknown parameter θ, if

$$\underset{\theta}{Bias}\left(\hat{\theta}\right) \neq 0 \; but \; \lim_{n\to\infty} \underset{\theta}{Bias}\left(\hat{\theta}\right) = 0$$

SOME PROPERTIES OF ESTIMATORS

- **CONSISTENT ESTIMATOR (CE):** An estimator $\hat{\theta}$ which converges in probability to an unknown parameter θ for all $\theta \in \Omega$ is called a CE of θ.

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

  - MLEs are generally CEs.

EXAMPLES

1. For a r.s. of size n,

$$E(\bar{X}) = \mu \Rightarrow \bar{X} \text{ is an UE of } \mu.$$

By WLLN,

$$\bar{X} \xrightarrow{p} \mu$$

$$\Rightarrow \bar{X} \text{ is a CE of } \mu.$$

EXAMPLES

2. $X \sim Uniform(0,\theta)$, $\theta > 0$. For a r.s of size n, is the MLE of θ an UE and a CE of θ?

EXAMPLES

3. Let $X_1, X_2, \ldots, X_n$ be a r.s. from $NB(2,\theta)$ distribution

$$f(x;\theta)=(x+1)\theta^2(1-\theta)^x, x=0,1,\ldots,0<\theta<1$$

where $\mu=\dfrac{2(1-\theta)}{\theta}$ *and* $\sigma^2=\dfrac{2(1-\theta)}{\theta^2}.$

a) Find the MLE of θ.

b) Find the MLE of μ.

c) Find an UE of $1/\theta$.

MEAN SQUARED ERROR (MSE)

- The **Mean Square Error (MSE)** of an estimator $\hat{\theta}$ for estimating θ is

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

If $\underset{\theta}{MSE}\left(\hat{\theta}\right)$ is smaller, $\hat{\theta}$ is the better estimator of θ.

$$\textit{For two estimators, } \hat{\theta}_1 \textit{ and } \hat{\theta}_2 \textit{ of } \theta, \textit{ if}$$

$$\underset{\theta}{MSE}\left(\hat{\theta}_1\right) < \underset{\theta}{MSE}\left(\hat{\theta}_2\right), \theta \in \Omega$$

$\hat{\theta}_1$ *is better estimator of* θ.

EXAMPLE

Let $X_1, X_2, \ldots, X_n$ be a r.s. from $Uniform(0,\theta)$ distribution

a) Find the MLE of θ.

b) Find the MME of θ.

c) Compare MSEs of MLE and MME of θ and comment on which one is better estimator of θ.

MEAN SQUARED ERROR CONSISTENCY

- T_n is called **mean squared error consistent** (or consistent in quadratic mean) if $E\{T_n-\theta\}^2\to 0$ as $n\to\infty$.

Theorem: T_n is consistent in MSE iff

i) $Var(T_n)\to 0$ as $n\to\infty$.

ii) $\lim_{n\to\infty} E[T_n]=\theta.$

- If $E\{T_n-\theta\}^2\to 0$ as $n\to\infty$, T_n is also a CE of θ.

SUFFICIENT STATISTICS

- X, $f(x;\theta)$, $\theta \in \Omega$
- $X_1, X_2, \ldots, X_n$ be a sample rvs
- $Y = U(X_1, X_2, \ldots, X_n)$ is a statistic.
- A **sufficient statistic**, Y is a statistic which contains all the information for the estimation of θ.

SUFFICIENT STATISTICS

- Given the value of Y, the sample contains no further information for the estimation of θ.

- Y is a **sufficient statistic (ss)** for θ if the conditional distribution of sample rvs given the value of y of Y, *i.e.* $h(x_1,x_2,\ldots,x_n|y)$ does not depend on θ for every given $Y=y$.

- A ss for θ is not unique.

SUFFICIENT STATISTICS

- The conditional distribution of sample rvs given the value of y of Y, is defined as

$$h(x_1, x_2, \cdots, x_n | y) = \frac{f(x_1, x_2, \cdots, x_n, y; \theta)}{g(y; \theta)}$$

$$h(x_1, x_2, \cdots, x_n | y) = \frac{L(\theta; x_1, x_2, \cdots, x_n)}{g(y; \theta)}$$

- If Y is a ss for θ, then

Not depend on θ for every given y.

$$h(x_1, x_2, \cdots, x_n | y) = \frac{L(\theta; x_1, x_2, \cdots, x_n)}{g(y; \theta)} = H(x_1, x_2, \cdots, x_n)$$

ss for θ

may include y or constant.

- Also, the conditional range of X_i given y not depend on θ.

SUFFICIENT STATISTICS

EXAMPLE: $X \sim Ber(p)$. For a r.s. of size n, find a ss for p is exists.

SUFFICIENT STATISTICS

- If Y is a ss for θ, then a 1-1 transformation of Y, say $Y_1=fn(Y)$ is also a ss for θ.
- **Neyman's Factorization Theorem:** Y is a ss for θ iff

$$L(\theta)=k_1(y;\theta)k_2(x_1,x_2,\cdots,x_n)$$

The likelihood function

Does not contain any other x_i

Not depend on θ for every given y (also in the conditional range of x_i.)

where k_1 and k_2 are non-negative functions and k_2 does not depend on θ for every given y.

EXAMPLES

1. $X \sim Ber(p)$. For a r.s. of size n, find a ss for p is exists.

EXAMPLES

2. $X \sim N(\mu,\sigma^2)$ where μ is known. For a r.s. of size n, find a ss for σ^2.

EXAMPLES

3. Let $X_1, X_2, \ldots, X_n$ be a r.s. from Uniform(0,θ) distribution. Find a ss for θ, if exists.

EXAMPLES

4. Let $X_1, X_2, \ldots, X_n$ be a r.s. from Cauchy(θ) distribution. Find a ss for θ, if exists.

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

SUFFICIENT STATISTICS

- A ss may not exist. Jointly ss $Y_1, Y_2, \ldots, Y_k$ may needed.
- A ss for θ is not unique, if exists.
- If the MLE of θ exists and unique and if a ss for θ exists, then MLE is a function of a ss for θ.

EXAMPLES

5. $X \sim N(\mu, \sigma^2)$. For a r.s. of size n, find jss for μ and σ^2.

EXAMPLES

6. $X \sim Uniform(\theta_1, \theta_2)$. For a r.s. of size n, find jss for θ_1 and θ_2.

MINIMAL SUFFICIENT STATISTICS

- A ss $T(X)$ is called minimal ss if, for any other ss $T'(X)$, $T(x)$ is a function of $T'(x)$.
- **THEOREM:** Let $f(x;\theta)$ be the pmf or pdf of a sample $X_1, X_2, \ldots, X_n$. Suppose there exist a function $T(x)$ such that, for two sample points $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$, the ratio

$$\frac{f(x_1, x_2, \cdots, x_n; \theta)}{f(y_1, y_2, \cdots, y_n; \theta)}$$

is constant as a function of θ iff $T(x)=T(y)$. Then, $T(X)$ is a minimal sufficient statistic for θ.

EXAMPLE

- $X \sim N(\mu, \sigma^2)$. For a r.s. of size n, find minimal jss for μ and σ^2.

RAO-BLACKWELL THEOREM

- Let $X_1, X_2, \ldots, X_n$ have joint pdf or pmf $f(x_1, x_2, \ldots, x_n; \theta)$ and let $S=(S_1, S_2, \ldots, S_k)$ be a vector of jss for θ. If T is an UE of $\tau(\theta)$ and $\varphi(T)=E(T|S)$, then

i) $\varphi(T)$ is an UE of $\tau(\theta)$.

ii) $\varphi(T)$ is a fn of S.

iii) $Var(\varphi(T)) \le Var(T)$ for all $\theta \in \Omega$.

- $\varphi(T)$ *is a uniformly better unbiased estimator of* $\tau(\theta)$.

SUFFICIENT STATISTICS

- If an UE of θ exists, than an UE of θ which is a function of a ss for θ also exists.

- The minimum variance unbiased estimator (MVUE) of θ, if exists, should be a function of a ss for θ.

ANCILLARY STATISTIC

- A statistic $S(X)$ whose distribution does not depend on the parameter θ is called an **ancillary statistic**.
- An ancillary statistic contains no information about θ.
- An ancillary statistic and ss for θ are dependent.

EXAMPLE

- $X \sim N(\theta,1)$. $X_1, X_2, \ldots, X_n$ be a r.s.

$$S^2 \sim Gamma[(n-1)/2, 2/n]$$

$$\downarrow$$

The distribution is free from θ.

$$\downarrow$$

S^2 is ancillary for θ.

COMPLETENESS AND UNIQUENESS

- Let $\{f(x;\ \theta),\ \theta\in\Omega\}$ be a family of pdfs (or pmfs) and $U(x)$ be an arbitrary function of x not depending on θ. If

$$E_\theta\left[U\left(X\right)\right]=0 \text{ for all } \theta\in\Omega$$

requires that the function itself equal to 0 for all possible values of x; then we say that this family is a complete family of pdfs (or pmfs).

$$E_\theta\left[U\left(X\right)\right]=0 \text{ for all } \theta\in\Omega \Rightarrow U\left(x\right)=0 \text{ for all } x.$$

EXAMPLES

1. Show that the family $\{Bin(n=2,\theta);\ 0<\theta<1\}$ is complete.

EXAMPLES

- $X \sim Uniform(-\theta,\theta)$. Show that the family $\{f(x;\theta),\ \theta>0\}$ is not complete.

EXAMPLES

- Consider the family $\{f(x;\theta),\ \theta\in\Omega\}$ where

$$f(x;\theta)=\frac{1}{\theta},0<x<\theta$$

a) Show that the family is complete in $\Omega=\{\theta:0<\theta<1\}$.

b) Show that the family is **not** complete in $\Omega=\{\theta:1<\theta<\infty\}$.

BASU THEOREM

- If $T(X)$ is a *complete and minimal sufficient statistic*, then $T(X)$ is independent of every ancillary statistic.

- $X \sim N(\mu,\sigma^2)$.

$$\bar{X} : \text{the mss for } \mu$$

$$S^2 \sim \chi^2_{n-1} \longrightarrow \text{Ancillary statistic for } \mu$$

By Basu theorem, $\bar{X}$ and S^2 are independent.

COMPLETE AND SUFFICIENT STATISTICS (css)

- Y is a complete and sufficient statistic (css) for θ if Y is a ss for θ and the family

$$\left\{ g\left(y;\theta\right);\theta\in\Omega \right\}$$

is complete.

The pdf of Y.

1) Y is a ss for θ.

2) $u(Y)$ is an arbitrary function of Y. $E(u(Y))=0$ for all $\theta\in\Omega$ implies that $u(y)=0$ for all possible $Y=y$.

LEHMANN-SCHEFFE THEOREM

- Let Y be a **css** for θ. If there is a function Y which is an **UE** of θ, then the function is the **unique** Minimum Variance Unbiased Estimator (MVUE) of θ.
  - Y css for θ.
  - $T(y)=fn(y)$ and $E[T(Y)]=\theta$.

$\rightarrow T(Y)$ is the MVUE of θ.
$\rightarrow$So, it is the best estimator of θ.

THE MINIMUM VARIANCE UNBIASED ESTIMATOR

- Let Y be a css for θ. Since Y is complete, there could be only a unique function of Y which is an UE of θ.
- Let $U_1(Y)$ and $U_2(Y)$ be two function of Y. $W(Y)=U_1(Y)-U_2(Y)=0$ for all possible values of Y. Therefore, $U_1(Y)=U_2(Y)$ for all Y.
- **Rao-Blackwell Theorem:** If T is am unbiased estimator of θ, and S is a css for θ, then $\mathrm{E}[\varphi(T)]=E[E(T|S)]=\theta$ and $\varphi(T)$ is the unique MVUE of θ.

EXAMPLES

1. $X \sim Ber(\theta)$. Let $X_1, X_2, \ldots, X_n$ be a r.s.

   Show that $Y = \sum_{i=1}^{n} X_i$ is a css for θ and find the MVUE of θ.

EXAMPLES

2. Let X have the pdf $f(x;\theta)$ with $\mu=2\theta$ and $\sigma^2=2\theta^2$. Let $Y=X_1+X_2+...+X_n$ be a css for θ for a r.s. of size n.

a) Find the MVUE of θ.

b) Find the MVUE of θ^2.

c) Find $E[X_1+X_2|Y]$.

EXAMPLES

3. Let X have the pdf

$$f\left(x;\theta\right)=3\theta^{2}x^{4},x\geq\theta,\theta>0$$

where $\mu=3\theta/2$ and $\sigma^2= 3\theta^2/4$.

- Consider a r.s. $X_1, X_2, \ldots, X_n$.
- *Let* $X_{(1)}=min(X_1, X_2, \ldots, X_n)$ *with pdf*

$$g_{X_{(1)}}\left(y\right)=3n\theta^{3n}y^{-\left(3n+1\right)},y\geq\theta$$

and $E(X_{(1)})=3n\theta/(3n-1)$.

Find the MVUE of $\sigma^2= 3\theta^2/4$.

EXPONENTIAL CLASS OF PDFS

- X is a continuous (discrete) rv with pdf $f(x;\theta)$, $\theta \in \Omega$. If the pdf can be written in the following form

$$f(x;\theta) = e^{P(\theta)K(x)+S(x)+Q(\theta)}, a < x < b$$

the pdf is a member of exponential class of pdfs of the continuous (discrete) type.

REGULAR CASE OF THE EXPONENTIAL CLASS OF PDFS

- We have a regular case of the exponential class of pdfs of the continuous type if

a) Range of X does not depend on θ.

b) $P(\theta)$ is a non-trivial continuous function of θ for $\theta \in \Omega$.

c) $dK(x)/dx \not\equiv 0$ and a continuous function of x for $a<x<b$.

d) $S(x)$ is a continuous function of x for $a<x<b$.

REGULAR CASE OF THE EXPONENTIAL CLASS OF PDFS

- $Exponential\ Class + Regular\ Case + Random\ Sample$

$$Y = \sum_{i=1}^{n} K\left(X_i\right) \text{ is a css for } \theta.$$

$\boxed{\text{If } Y \text{ is an UE of } \theta,\ Y \text{ is the MVUE of } \theta.}$

EXAMPLES

1. $X \sim N(\mu,\sigma^2)$ where μ is known. Find a css for σ^2 and find the MVUE of σ^2.

EXAMPLES

2. $X \sim N(\mu, \sigma^2)$ where σ^2 is known. Find a css for μ and find the MVUE of μ.

REGULAR CASE OF THE DISCERETE EXPONENTIAL CLASS OF PDFS

- We have a regular case of the exponential class of pdfs of the continuous type if

a) Range of X does not depend on θ.

b) $P(\theta)$ is a non-trivial continuous function of θ for $\theta \in \Omega$.

c) $K(x) \not\equiv c$ for $x = a_1, a_2, \ldots$.

EXAMPLES

3. $X \sim Ber(\theta)$.

a) Find a css for θ.

b) Find the MVUE of θ.

c) Find the MVUE of θ^2.

EXAMPLES

4. $X \sim Poisson(\theta)$. Find a css for θ and find the MVUE of $P(X \leq 1) = (1+\theta)e^{-\theta}$.

EXAMPLES

5. $X\sim Uniform(\theta)$. Is this pdf a member of exponential class of pdfs? Why?

FISHER INFORMATION AND INFORMATION CRITERIA

- *X, f(x;θ), θ∈Ω, x∈A (not depend on θ).*

$$\lambda(x;\theta) = \ln f(x;\theta)$$

$$\lambda'(x;\theta) = \frac{\partial \ln f(x;\theta)}{\partial \theta}$$

$$\lambda''(x;\theta) = \frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}$$

$$f'(x;\theta) = \frac{\partial f(x;\theta)}{\partial \theta}$$

$$f''(x;\theta) = \frac{\partial^2 f(x;\theta)}{\partial \theta^2}$$

69

FISHER INFORMATION AND INFORMATION CRITERIA

$$\int_A f(x;\theta)dx = 1 \Rightarrow \frac{d}{d\theta}\int_A f(x;\theta)dx = 0$$

$$\Rightarrow \int_A f'(x;\theta)dx = 0$$

$$\Rightarrow \int_A f''(x;\theta)dx = 0$$

$$\lambda'(x;\theta) = \frac{\partial \ln f(x;\theta)}{\partial\theta} = \frac{f'(x;\theta)}{f(x;\theta)}$$

$$\lambda''(x;\theta) = \frac{f''(x;\theta)}{f(x;\theta)} - [\lambda'(x;\theta)]^2$$

70

FISHER INFORMATION AND INFORMATION CRITERIA

$$E[\lambda'(X;\theta)] = \int_A \lambda'(x;\theta) f(x;\theta) dx = 0$$

$$E[\lambda''(X;\theta)] = \int_A \lambda''(x;\theta) f(x;\theta) dx$$

$$= \int_A \left\{ \frac{f''(x;\theta)}{f(x;\theta)} - [\lambda'(x;\theta)]^2 \right\} f(x;\theta) dx$$

$$= \int_A f''(x;\theta) dx - \int_A [\lambda'(x;\theta)]^2 f(x;\theta) dx$$

$$= 0 - E[\lambda'(x;\theta)]^2$$

71

FISHER INFORMATION AND INFORMATION CRITERIA

$$\Rightarrow E\left[\lambda''(x;\theta)\right] = -E\left[\lambda'(x;\theta)\right]^2 = -V\left[\lambda'(x;\theta)\right]$$

The Fisher Information in a random variable X:

$$I[\theta] = E\left[\lambda'(x;\theta)\right]^2 = V\left[\lambda'(x;\theta)\right] = -E\left[\lambda''(x;\theta)\right] \geq 0$$

The Fisher Information in the random sample:

$$I_n[\theta] = nI(\theta)$$

72

CRAMER-RAO LOWER BOUND (CRLB)

- Let $X_1, X_2, \ldots, X_n$ be sample random variables.
- $Y=U(X_1, X_2, \ldots, X_n)$: a statistic not containing θ.
- $E(Y)=m(\theta)$.
- $Z=\lambda'(x_1, x_2, \ldots, x_n; \theta)$ *is a r.v.*
- $E(Z)=0$ *and* $V(Z)=I_n(\theta)$.

73

CRAMER-RAO LOWER BOUND (CRLB)

- $Cov(Y,Z)=E(YZ)-E(Y)E(Z)=E(YZ)$

$$E(Y.Z)=\int\cdots\int u(x_1,x_2,\cdots,x_n)\lambda_n'(x_1,x_2,\cdots,x_n;\theta)dx_1dx_2\cdots dx_n$$

$$=\int\cdots\int u(x_1,\cdots,x_n)\frac{f'(x_1,\cdots,x_n;\theta)}{f(x_1,\cdots,x_n;\theta)}f(x_1,\cdots,x_n;\theta)dx_1\cdots dx_n$$

$$=\int\cdots\int u(x_1,\cdots,x_n)f'(x_1,\cdots,x_n;\theta)dx_1\cdots dx_n=m'(\theta)$$

74

CRAMER-RAO LOWER BOUND (CRLB)

- $E(Y.Z)=m'(\theta)$
- $-1 \le Corr(Y,Z) \le 1 \Rightarrow -1 \le \dfrac{Cov(Y,Z)}{\sqrt{V(Y)}\sqrt{V(Z)}} \le 1$
- $0 \le Corr(Y,Z)^2 \le 1 \Rightarrow 0 \le \dfrac{[Cov(Y,Z)]^2}{V(Y)V(Z)} \le 1$

$$\Downarrow$$

$$0 \le \frac{[m'(\theta)]^2}{V(Y)I_n(\theta)} \le 1$$

The Cramer-Rao Inequality
(Information Inequality)

$$V(Y) \ge \frac{[m'(\theta)]^2}{I_n(\theta)} \Rightarrow \text{The Cramer - Rao Lower Bound}$$

75

CRAMER-RAO LOWER BOUND (CRLB)

- $CRLB$ is the lower bound for the variance of the unbiased estimator of $m(\theta)$.
- When $V(Y)=CRLB$, Y is the MVUE of $m(\theta)$.
- *For a r.s.,*

$$V(Y) \geq \frac{[m'(\theta)]^2}{nI(\theta)} \Rightarrow \text{The Cramer - Rao Lower Bound}$$

76

EFFICIENT ESTIMATOR

- Y is an efficient estimator (EE) of its expectation, $m(\theta)$, if its variance reaches the CRLB.
- An EE of $m(\theta)$ may not exist.
- The EE of $m(\theta)$, *if exists, is unique.*
- *The EE of $m(\theta)$ is the unique MVUE of $m(\theta)$.*
- *If the MVUE of $m(\theta)$ is not EE of $m(\theta)$, then an EE of $m(\theta)$ does not exist.*

77

ASYMPTOTIC EFFICIENT ESTIMATOR

- Y is an asymptotic EE of $m(\theta)$ *if*

$$\lim_{n\to\infty} E(Y) = m(\theta)$$

$$and$$

$$\lim_{n\to\infty} V(Y) = CRLB$$

78

EXAMPLES

1. $X \sim Poi(\mu)$.

a) Find CRLB for μ.

b) Find CRLB for $e^{-\mu}$.

c) Find MLE of μ.

d) Show that $\left(\dfrac{n-1}{n}\right)^{\sum_{i=1}^{n} X_i}$ is an UE of $e^{-\mu}$.

79

EXAMPLES

2. $X \sim Uniform(0,\theta)$

80

EXAMPLES

3. $X \sim Exp(1/\theta)$. Find an EE of θ, if exists.

81

STRUCTURAL FORM

- If

$$\frac{d \ln L(\theta)}{d\theta} = k(\theta)\left(\hat{\theta} - \theta\right), k(\theta) \neq 0$$

where $k(\theta)$ is free from $x_1, x_2, \ldots, x_n$, then automatically $\hat{\theta}$ is an UE of θ and $\hat{\theta}$ is the MVUE of θ.

Remark: $Var\left(\hat{\theta}\right) = \frac{1}{k(\theta)}.$

82

EXAMPLE

$X \sim Exp(\theta)$. Find an EE of θ, if exists.

83

LIMITING DISTRIBUTION OF MLEs

- $\hat{\theta}$: MLE of θ (obtained by differentiation)
- $X_1, X_2, \ldots, X_n$ is a random sample.

$$m\left(\hat{\theta}\right) \overset{asymptotically}{\sim} N\left(m(\theta), \underset{m(\theta)}{RCLB} = \frac{\left[m'(\theta)\right]^2}{nI(\theta)} \right)$$

$$\hat{\theta} \underset{\text{large } n}{\overset{asympt.}{\sim}} N\left(\theta, \frac{1}{nI(\theta)} \right)$$

$$\Rightarrow \frac{\hat{\theta} - \theta}{\sqrt{\dfrac{1}{nI(\theta)}}} = \sqrt{nI(\theta)}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N(0,1)$$

LIMITING DISTRIBUTION OF MLEs

- Let $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_m$ be MLEs of $\theta_1, \theta_2, \ldots, \theta_m$.

$$\hat{\theta}_i \overset{asympt.}{\sim} N\left(\theta_i, RCLB_{\theta_i}\right)$$

$$m\left(\hat{\theta}_i\right) \overset{asympt.}{\sim} N\left(m\left(\theta_i\right), RCLB_{m\left(\theta_i\right)}\right), i = 1, 2, \ldots, m$$

- EE of $m(\theta) = m\left(\hat{\theta}\right) = fn\left(ss \text{ for } \theta\right)$

- If Y is an EE of θ, then $Z = a + bY$ is an EE of $a + bm(\theta)$ where a and b are constants.

EXAMPLE

- $X \sim N(\mu,\sigma^2)$ where σ^2 is known. Assume that we have a r.s of size n.

a) Find a Fisher Information.

b) Find the EE of μ, if it exists.

c) Find the EE of μ^2, if it exists.

EXAMPLE

- $X \sim Exp(\theta^2)$. Assume that we have a r.s of size n.

a) Find a Fisher Information.

b) Using CRLB, show that $\bar{X}$ is the EE of θ^2.

c) Find the EE of θ^4, if it exists. If it does not exist, find an asymptotic EE of θ^4 and specify its asymptotic distribution.

d) Find the MVUE of θ^4, if it exists.