11-04-2022 Unit-4

Monday, April 11, 2022 2:58 PM

[redacted] Unit-4 (Point Estimation)

$E(T_n) = \theta$

Consistent → unbiased

* sampling theory → statistic Estimators → sample's parameter

Estimators → unbiased ✓, consistent ✓, efficient ✓, sufficient.

→ Every unbiased estimator is consistent ??

$\underbrace{E(T_n) = \theta}$ ✓

As $n \to \infty$

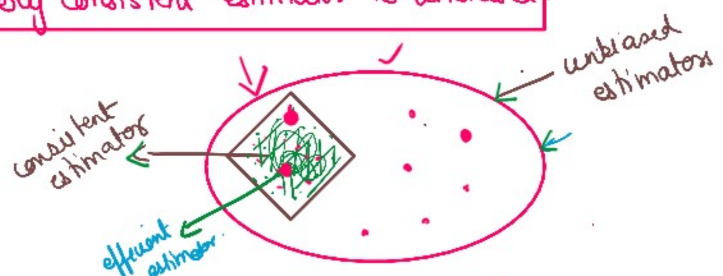
$T_n \to \theta$. ✓

$E(T_n) \to \theta$ & $V(T_n) \to 0$ as $n \to \infty$

NO

T_n is it unbiased?

$\boxed{\text{Every consistent estimator is unbiased}}$



efficient + consistent → unbiased

(3) Efficient Estimator :- In unbiased estimators, sometimes, there exists more than one consistent estimators for a particular parameter. To choose the better estimator, we use the criteria of variance. ✓
That is,

$T_1, T_2, \ldots, T_n$

$Var(T_i) = ?$

T_4 - most efficient estimator.

$\boxed{\begin{array}{c}\text{lesser the variance} \\ \Rightarrow \text{better the efficiency}\end{array}}$

T_1, T_2, T_3
$var(T_1) = 2$
$var(T_2) = 1.9$
$var(T_3) = 2.1$

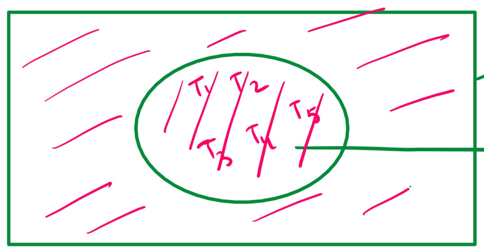
That, if T_1 and T_2 are two consistent estimators used to estimate a parameter θ and $Var(T_1) < Var(T_2)$, then -

$var(T_3) = 2.1$

a parameter θ and $Var(T_1) < Var(T_2)$, then –

T_1 will be more efficient than T_2.

$v(T_i) = ?$
least variance
$\Rightarrow$ efficient estimator



unbiased estimators
$E(T_n) = \theta$

consistent estimators
As $n \to \infty$,
$E(T_n) \to \theta$
$\to$ $V(T_n) \to 0$

§ Efficiency of an estimator – If T_1 is the most efficient estimator with variance V_1 and T_2 is any other estimator with variance V_2, then the efficiency E_2 of T_2 is defined as –

$$E_2 = \frac{V_1}{V_2} \quad \checkmark \qquad \leq 1$$

$\underbrace{T_1, T_2, \ldots, T_n}$ & T_1 is the most efficient. Then,

$$E_1 = \frac{V_1}{V_1} = 1 \quad \checkmark$$

$$E_2 = \frac{V_1}{V_2} < 1$$

$$E_3 = \frac{V_1}{V_3} < 1$$

$$\vdots$$

$$E_n = \frac{V_1}{V_n} < 1$$

i.e., $\boxed{E_i \leq 1}$

$\underbrace{\boxed{0 < E_i \leq 1}}$ $(0, 1]$

* Remark: Efficiency cannot exceed unity.

* Defn: Efficiency is the ability of a consistent estimator to estimate a given parameter.

$$E(T_1) = \frac{V(T_3)}{V(T_1)} \qquad E(T_2) = \frac{V(T_3)}{V(T_2)}$$

$$= \frac{2/n}{3/n} = \frac{2}{3} < 1. \qquad = \frac{2/n}{4/n} = \frac{1}{2}.$$

$= \frac{\ldots}{3/n} = \frac{2}{3} < 1 .$ $\quad \ldots \frac{\ldots}{4/n} = \frac{1}{2} .$

Pb :- Let T_1, T_2, T_3 are any three consistent estimators used to estimate a parameter θ such that - $Var(T_1) = \frac{3}{n}$, $Var(T_2) = \frac{4}{n}$ & $Var(T_3) = \frac{2}{n}$.

(a) T_1
(b) T_2
(c) T_3

(i) which is most efficient?

$\longrightarrow$ Since $Var(T_3)$ is smallest, so T_3 is the most efficient estimator.

(ii) Find the efficiency of T_1 and T_2.

(A) $\frac{2}{3}, \frac{1}{3}$ (B) $\frac{2}{3}, \frac{1}{4}$ (C) $\frac{2}{3}, \frac{1}{2}$ (D) $\frac{2}{3}, \frac{1}{6}$

$\longrightarrow$

$$E_1 = \frac{V_3}{V_1} = \frac{2/n}{3/n} = \frac{2}{3}$$

$$E_2 = \frac{V_3}{V_2} = \frac{2/n}{4/n} = \frac{1}{2}$$

Pb X_1, X_2 and X_3 is a random sample of size 3 from a population with mean value μ and variance σ^2. T_1, T_2, T_3 are the estimators used to estimate mean value μ, where -

$$T_1 = X_1 + X_2 - X_3, \quad T_2 = 2X_1 + 3X_3 - 4X_2, \quad T_3 = \frac{1}{3}(\lambda X_1 + X_2 + X_3).$$

(i) Find the value of λ if T_3 is unbiased. (a) 1 (b) 2 (c) 3 (d) 4

(A) 1 (B) 2 (C) 3

$\rightarrow$ $E(T_3) = \mu$

$$\Rightarrow E\left(\frac{\lambda X_1 + X_2 + X_3}{3}\right) = \mu$$

$$\Rightarrow \frac{\lambda}{3} E(X_1) + \frac{1}{3} E(X_2) + \frac{1}{3} E(X_3) = \mu$$

$$\Rightarrow \frac{\lambda}{3} E(X_1) + \frac{1}{3} E(X_2) + \frac{1}{3} E(X_3) = \mu$$

$$\Rightarrow \frac{\lambda+2}{3} \mu = \mu$$

$$\Rightarrow \lambda + 2 = 3$$

$$\Rightarrow \boxed{\lambda = 1}$$

$Var = \sigma^2$

$Var(X_i) = \sigma^2$

$T_1 = X_1 + X_2 - X_3$

$T_2 = 2X_1 + 3X_2 - 4X_3$

$T_3 = \frac{X_1 + X_2 + X_3}{3}$

$Var(ax+b) = a^2 V(x)$

(b) Which is the best estimator?

Ⓐ T_1 Ⓑ T_2 Ⓒ T_3

$$Var(T_1) = Var(X_1 + X_2 - X_3)$$
$$= 1^2 Var(X_1) + 1^2 Var(X_2) + (-1)^2 Var(X_3)$$
$$= \sigma^2 + \sigma^2 + \sigma^2$$
$$= 3\sigma^2.$$

$$Var(T_2) = Var(2X_1 + 3X_2 - 4X_3)$$
$$= 4\sigma^2 + 9\sigma^2 + 16\sigma^2 = 29\sigma^2$$

$$Var(T_3) = \frac{1}{9}(\sigma^2 + \sigma^2 + \sigma^2) = \frac{\sigma^2}{3}$$

$Var(T_3)$ is the smallest, so T_3 is the best estimator.

$$\boxed{\varepsilon_3 = 1}.$$

(c) Calculate the efficiency of T_1 & T_2.

$$\rightarrow \quad \varepsilon_1 = \frac{V(T_3)}{V(T_1)} = \frac{\sigma^2/3}{3\sigma^2} = \frac{1}{9}$$

$$\varepsilon_2 = \frac{V(T_3)}{V(T_2)} = \frac{\sigma^2/3}{29\sigma^2} = \frac{1}{87}$$

(d) With $\lambda = 1$, is T_3 a consistent estimator?

YES NO

$$T_3 = \frac{X_1 + X_2 + X_3}{3} = \bar{X}$$

* Weak law of large no-s - sample mean is always a consistent estimator of population mean.

Lecture 30

[redacted] Unit - 4 (Point Estimation)

Unbiasedness
Consistency
Efficiency

* Methods of Estimation -

→ Method of maximum likelihood estimation :-

1st time C.F. Gauss

Prof. R.A. Fisher

Imp § Likelihood function :- Let, $x_1, x_2, \ldots, x_n$ be a random sample of size n from a population with density function $f(x, \theta)$. Then, the likelihood function of the sample values $x_1, x_2, \ldots, x_n$, usually denoted by $L = L(\theta)$ is their joint density function -

$$L = f(x_1, \theta) f(x_2, \theta) f(x_3, \theta) \cdots f(x_n, \theta)$$

$$\Rightarrow \boxed{L = \prod_{i=1}^{n} f(x_i, \theta)}$$

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

→ $\theta = \hat{\theta}$ for which L is maximum.

To find the maximum value of likelihood function $L(\theta)$, we follow the steps below :-

(i) $\frac{\partial}{\partial \theta} L(\theta) = 0 \Rightarrow$ points of extrema.

(ii) $\frac{\partial^2}{\partial \theta^2} L(\theta)$. If $\frac{\partial^2}{\partial \theta^2} L(\theta) < 0$, θ is a point of maxima.

Let $\theta = \hat{\theta}$.

$L(\theta)$

Then, the maximum likelihood estimator is -

$L(\hat{\theta})$.

§ Properties of Maximum likelihood Estimators :-

(i) The first & second order derivatives i.e. $\frac{\partial}{\partial \theta} \log L$ and $\frac{\partial^2}{\partial \theta^2} \log L$ exist and are continuous functions of θ in a range R for almost all x.

(ii) The 3rd order derivative $\frac{\partial^3}{\partial \theta^3} \log L$ exists such that

$\left| \frac{\partial^3}{\partial \theta^3} \log L \right| < M(x)$, where $E[M(x)] < K$, a +ve quantity.

(iii) For every θ in R,

$$E\left(-\frac{\partial^2}{\partial \theta^2} \log L\right) = \int_{-\infty}^{\infty} \left(-\frac{\partial^2}{\partial \theta^2} \log L\right) L \, dx = I(\theta), \text{ is}$$

finite & non-zero.

(iv) The range of integration is independent of θ. But, if the range of integration depends on θ, then $f(x, \theta)$ vanishes at the extremes depending on θ.

$$\boxed{E(x) = \int_{-\infty}^{\infty} x f(x) \, dx}$$

* Cramer-Rao Theorem — With probability approaching unity as $n \to \infty$, the likelihood equation $\frac{\partial}{\partial \theta} \log L = 0$, has a solution which converges in probability to the true value θ_0. In other words, M.L.E's are consistent.

$$\boxed{\frac{\partial}{\partial \theta} \log L = 0}$$ likelihood equation.

As $n \to \infty$
$E(T_n) \to \theta$
$V(T_n) \to 0$

$|T_n - \theta| < \epsilon$
$T_n \to \theta$

As $n \to \infty$, $P\{|E(T_n) - \theta| < \epsilon\} \to 1$, as $n \to \infty$

$P\{|E(\theta) - \theta_0| < \epsilon\} \to 1$.

$|T_n - \theta| < \epsilon$

$T_n \to \theta$

$$\boxed{\begin{array}{l} \text{As } n \to \infty, \quad P\{|E(T_n) - \theta| < \epsilon\} \to 1, \quad \text{as } n \to \infty \\ P\{|E(\theta) - \theta| < \epsilon\} \to 1. \end{array}}$$

Remark:- MLE's are always consistent estimators but need not be unbiased.

For example- in sampling from $N(\mu, \sigma^2)$ population-

$MLE(\mu) = \bar{x}$ (sample mean), which is both unbiased and consistent.

but $MLE(\sigma^2) = s^2$ (sample ~~mean~~ variance), which is consistent but not unbiased.

$$\boxed{E\left(\frac{n}{n-1} s^2\right) = \sigma^2} \qquad E(s^2) \neq \sigma^2$$

Theorem (Hazoor Bazar's theorem):- Any consistent solution of the likelihood equation provides a maximum of the likelihood with probability tending to unity as the sample size$(n) \to \infty$.

$$\frac{\partial}{\partial \theta} \log L = 0 \leftarrow \text{maximum of the likelihood as } n \to \infty.$$

Theorem:- If M.L.E. exists, it is the most efficient in the class of such estimators.

$Var(\theta)$ minimum

Maximum likelihood estimator.

$\theta = \hat{\theta} \longrightarrow$ most efficient

$T \to \theta$

$\psi(T) \to \psi(\theta)$

$\theta = \hat{\theta} \longrightarrow$ most efficient $(\varepsilon = 1)$

§ Invariance property of M.L.E :-

If T is the M.L.E of θ, and $\psi(\theta)$ is one-to-one function of θ, then $\psi(T)$ is the M.L.E of $\psi(\theta)$.

Ex- In random sampling from normal population $N(\mu, \sigma^2)$, find the maximum likelihood estimator for μ when σ^2 is known.

Solution: Here, $X \sim N(\mu, \sigma^2)$, then -

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

so,

$$f(x_i, \theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}.$$

Now, Likelihood function is -

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \cdots \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdots e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}$$

$\sum^{n} (x_i-\mu)^2$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right) e^{\dots} \cdots e^{\dots}$$

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

$$\Rightarrow \log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \quad \rightarrow (1)$$

Now, when σ^2 is known, the likelihood equation is –

$$\frac{\partial}{\partial\mu}\log L = 0$$

$$\Rightarrow -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)(-1) = 0 \quad [\text{from } (1)]$$

$$\Rightarrow \sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\Rightarrow (x_1-\mu)+(x_2-\mu)+\cdots(x_n-\mu) = 0$$

$$\Rightarrow x_1+x_2+\cdots+x_n = n\mu$$

$$\Rightarrow \mu = \frac{x_1+x_2+\cdots+x_n}{n}$$

$$\Rightarrow \mu = \bar{x} \text{ (sample mean).}$$

$\therefore$ M.L.E for μ is the sample mean $\bar{x}$.

(b) M.L.E for σ^2, when μ is known.

Diff^n (1) partially w.r.t σ^2, we have –

$$\frac{\partial}{\partial\sigma^2}\log L = 0$$

$$\Rightarrow -\frac{n}{2}\times\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

$$\Rightarrow n - \frac{1}{}\sum(x_i-\mu)^2 = 0$$

$$\Rightarrow \quad n - \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

$$\Rightarrow \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\Rightarrow \quad \sigma^2 = s^2$$

$\therefore$ M.L.E for σ^2 is the sample variance s^2.

(c) Find the M. L. E for the simultaneous estimation of μ & σ^2.

$\rightarrow$ The likelihood equations for the simultaneous estimation of μ & σ^2 are –

$$\frac{\partial}{\partial \mu} \log L = 0 \quad \& \quad \frac{\partial}{\partial \sigma^2} \log L = 0$$

$$\Rightarrow \quad \hat{\mu} = \bar{x} \qquad \Rightarrow \quad \hat{\sigma}^2 = s^2 \quad //$$