Point Estimation

Basic Concepts

Estimators:- Any funⁿ of the random sample which is used to estimate the unknown value of the given parametric funⁿ $g(\theta)$ is called an estimator. If $\underline{X} = (X_1, X_2, \ldots, X_n)$ is a random sample from a popⁿ with the prob distⁿ P_θ, a funⁿ $d(\underline{X})$ used for estimating $g(\theta)$ is known as an estimator.

Let $x_1, x_2, \ldots, x_n$ be a realization of $\underline{X}$. Then the corresponding value $d(\underline{x})$ is called an estimate.

For example, in estimating the average height of adult males in an ~~at~~ ethnic group, we may use the sample mean $\overline{X}$ as an estimator. Now if a random sample of 50 has a sample mean 180 cm, then 180 cm is an estimate of the average height.

Parameter space:- is the set of all possible values of a parameter. Estimators must lie in parameter space.

$\theta \rightarrow \Theta, \Omega$.

Desirable Criteria for Estimators:-

Unbiasedness:- Let $X_1, X_2, \ldots, X_n$ be a random sample from a popⁿ with prob. distⁿ P_θ, $\theta \in \Theta$. An estimator $T(\underline{X})$, $\underline{X} = (X_1, \ldots, X_n)$ is said to be unbiased for estimating $g(\theta)$, if

$$E_\theta[T(\underline{X})] = g(\theta) \quad \forall \theta \in \Theta$$

i.e. Avarage value of the estimator = the original parameter value.

If $E_\theta T(x) = g(\theta) + b(\theta)$
then $b(\theta)$ is called bias of T. If $b(\theta) > 0 \ \forall \theta$
then T is said to over-estimate $g(\theta)$. On the other hand if $b(\theta) < 0 \ \forall \theta$, then T is an underestimator of $g(\theta)$.

Ex 1:- Let $X \sim Bin(n, p)$, n is known; $0 \le p \le 1$.

$E(X) = np$

$\Rightarrow E\left(\frac{X}{n}\right) = p$.

So, $\frac{X}{n}$ (sample proportion) is unbiased for p which is pop^n proportion.

$E\left(X(X-1)\right) = n(n-1)p^2$

$\Rightarrow E\left(\frac{X(X-1)}{n(n-1)}\right) = p^2$

So, $\frac{X(X-1)}{n(n-1)}$ is unbiased for p^2.

$V(X) = np(1-p)$
$= n(p-p^2)$

So, $d(X) = n\left[\frac{X}{n} - \frac{X(X-1)}{n(n-1)}\right] = \frac{X(n-X)}{n-1}$ is unbiased for variance of X

2. Let $X_1, X_2, \ldots, X_n \sim P(\lambda)$, $\lambda > 0$ $E(X_i) = \lambda$ $i = 1(1)n$

$T_1(\underline{X}) = \overline{X}$, $T_2(\underline{X}) = X_i$, $T_3(\underline{X}) = \frac{X_1 + 2X_2}{3}$

are some unbiased estimators for λ.

$V(X) = \lambda$

$T_4(\underline{X}) = \frac{1}{n-1}\Sigma(X_i - \overline{X})^2$

$\sum_{j=1}^{4} c_j T_j = d(\underline{X})$ is also unbiased for λ if $\sum_{j=1}^{4} c_j$ since each T_j are unbiased.

We can generate infinite no. of unbiased estimators.

Scanned by TapScanner

3. If $E(X)$ exists, then any given estimation problem the sample mean is an unbiased estimator of the pop^n mean.

4. let $E(X^2)$ exists, i.e, $V(X) = \sigma^2$ exists. Then $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is unbiased for σ^2.
$\longrightarrow$ sample variance

$E(S^2) = \frac{1}{n-1}\sum_{i=1}^{n} E(X_i - \bar{X})^2$

$= \frac{1}{n-1} E\left(\sum X_i^2 - n\bar{X}^2\right)$

$= \frac{1}{n-1}\left[n(\mu^2+\sigma^2) - n(\mu^2 + \frac{\sigma^2}{n})\right]$

$= \sigma^2$

$E X_i^2 = \mu^2 + \sigma^2$

$E\bar{X}^2 = \mu^2 + \frac{\sigma^2}{n}$

~~$E(\bar{X}^2)$ ds~~

$E\left(\bar{X}^2 - \frac{S^2}{n}\right) = \mu^2 + \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = \mu^2$

So, $d^*(X) = \bar{X}^2 - \frac{S^2}{n}$ is unbiased for μ^2.

5. Let $f_X(x) = \lambda e^{-\lambda x}, x>0, \lambda>0$.

$X_1, X_2, \ldots, X_n$ be an random sample.

$E(X_i) = \frac{1}{\lambda}$

then $\bar{X}$ is an unbiased estimator for $\frac{1}{\lambda}$.

$E(X_i^k) = \frac{k!}{\lambda^k}$

$d_1(X) = \frac{1}{n k!}\sum_{i=1}^{n} X_i^k$ is unbiased for $\frac{1}{\lambda^k}$.

$Y = \sum X_i \sim \text{Gamma}(n, \lambda)$

$f_Y(y) = \frac{\lambda^n}{\Gamma(n)} e^{-\lambda y} y^{n-1}, y>0$

$E\left(\frac{1}{Y}\right) = \int_0^\infty \frac{1}{y} f(y)\,dy$

$= \int_0^\infty \frac{\lambda^n}{\Gamma(n)} e^{-\lambda y} y^{n-2}\,dy$

$= \frac{\lambda^n}{\Gamma(n)} \frac{\Gamma(n-1)}{\lambda^{n-1}} = \frac{\lambda}{n-1}$

Scanned by TapScanner

$E\left(\frac{n-1}{Y}\right) = \lambda \quad , n > 1$

~~Ie~~ for more that one observation we can estimate unbiasedly the reciprocal of the mean (rate).

(*) $E\left(\frac{1}{Y^k}\right) = \int_0^\infty \frac{\lambda^n}{\Gamma(n)} e^{-\lambda y} y^{n-k-1} \, dy \quad , n > k$

$$= \frac{\lambda^n}{\Gamma(n)} \frac{\Gamma(n-k)}{\lambda^{n-k}} = \frac{\Gamma(n-k)}{\Gamma(n)} \lambda^k$$

$\frac{\Gamma n}{\Gamma n-k} \cdot \frac{1}{Y^k}$ is unbiased for λ^k if $n > k$

6. Let $X \sim P(\lambda)$. We want to estimate the probability of no occurance, i.e. $P(X=0) = e^{-\lambda}$

$\Rightarrow$ Let $I(X) = \begin{cases} 1 & \text{if } X = 0 \\ 0 & \text{if } X \neq 0. \end{cases}$

$$E(I(X)) = 1 \cdot P(X=0) + 0 \cdot P(X \neq 0)$$
$$= P(X=0) = e^{-\lambda}$$

So, $I(X)$ is an unbiased estimator of $e^{-\lambda}$.

A Method to find unbiased estimator is to solve directly the eqn $E[T(X)] = g(\theta) \; \forall \theta$

Ex:- Let X be a truncated poisson r.v. with zero missing

$$P_X(x) = \frac{1}{e^{\lambda}-1} \frac{\lambda^x}{x!} \; ; x = 1, 2, \ldots$$

Suppose we want to estimate λ here.

$E(T(X)) = \lambda \quad \forall \lambda > 0$

$$\sum_{x=1}^{\infty} T(x) \frac{\lambda^x}{x!} \cdot \frac{1}{e^{\lambda}-1} = \lambda \quad \forall \lambda > 0.$$

or, $\sum_{n=1}^{\infty} T(n) \frac{\lambda^n}{n!} = \lambda(e^{\lambda}-1)$

or, $T(1)\frac{\lambda}{1!} + T(2)\frac{\lambda^2}{2!} + T(3)\frac{\lambda^3}{3!} + \cdots = \lambda(e^{\lambda}-1)$

$= \lambda\left[\lambda + \frac{\lambda^2}{2!} + \cdots\right]$

So, ~~on an open~~
Since the two power series can be identical on an open interval iff their all coefficients match, we get $T(1)=0$

$T(2) = 2! = 2$

$T(3) = 3$

$T(n) = n \ldots$

So, the unbiased estimator is

$T(n) = \begin{cases} 0 & \text{if } x=1 \\ x & \text{if } x = 2,3,\ldots \end{cases}$

↳ may not be useful in cont. distn.

<u>Remarks</u>:- 1. Unbiased estimators may not exist.

$X \sim Bin(n,p)$., $g(p) = p^{n+1}$

$E(T(X)) = p^{n+1}$; $0 \le p \le 1$

$\Rightarrow \sum T(x)\binom{n}{x} p^x (1-p)^{n-x} = p^{n+1}$ ——(1), $\forall\, 0 \le p \le 1$

Since the LHS is a polynomial of degree atmost 'n' it cannot equal the RHS which has a power $n+1$. So, eqn (1) has no soln.

So, unbiased estimator of p^{n+1} does not exist.

$\sin p$, $\ln p$, e^p, $\frac{1}{p}$ → unbiased estimators not exists.

2. Unbiased estimators are not reasonable

$\Rightarrow e^{-3\lambda} = E[(-2)^X] = \sum (-2)^x \frac{e^{-\lambda}\lambda^x}{x!}$

$0 < e^{-3\lambda} < 1$

$(-2)^X \to 1, -2, 4, -8, \ldots$

↳ not in $[0,1]$

Consistency :- (Large sample property)

An estimator $T_n = t(x_1, \dots, x_n)$ is said to be consistent for estimating $g(\theta)$ if for each $\varepsilon > 0$,

$$P(|T_n - g(\theta)| > \varepsilon) \to 0$$

as $n \to \infty$ $\forall \theta \in \Theta$

$$T_n \xrightarrow{P} g(\theta)$$

Example :- 1. Let $x_1, x_2, \dots, x_n$ be a random sample from a popn with mean μ and variance σ^2.

$$P(|\bar{x} - \mu| > \varepsilon) \leq \frac{var(\bar{x})}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0 \text{ as } n \to \infty$$

so, sample mean $\bar{x}$ is consistent for popn mean.

2. Let $x_1, x_2, \dots, x_n$ be a sequence of i.i.d. r.vs with mean μ, then by weak law of large numbers the, $\bar{x} \xrightarrow{P} \mu$

so, whenever popn mean exists, the sample mean is consistent for it.

$$f_X(x) = \frac{1}{\pi} \frac{1}{1+(x-\mu)^2}, \quad x \in \mathbb{R}, \mu \in \mathbb{R}$$

$E(X)$ does not exist

$\bar{x} \stackrel{d}{=} X_1$, $\quad P(|\bar{x} - \mu| < \varepsilon) = \frac{2}{\pi} \tan^{-1}(\varepsilon) \not\to 0$

~~So~~ Criteria of consistency fails.

Theorem :- If T_n is consistent for θ and h is a continuous funn, then $h(T_n)$ is consistent for $h(\theta)$.

(Unbiased case it is not applicable)

Scanned by TapScanner

Theorem:- If $E(T_n) = \theta_n \to \theta$, $V(T_n) = \sigma_n^2 \to 0$, then T_n is consistency for θ.

Theorem:- If T_n is consistent for θ and $a_n \to 1$, $b_n \to 0$, then $a_n T_n + b_n$ is also consistent for θ.

T_n, $\frac{n}{n+1} T_n + \frac{1}{n+1}$, $\frac{n+2}{n+4} T_n + \frac{1}{n}$ are all consistent for parameter θ.

T_n for θ

Lect-03

Examples:-

1. Let $x_1, x_2, \ldots, x_n$ be a random sample from an exponential dist^n with a location parameter μ.

$$f_{x_1}(x) = \begin{cases} e^{\mu - x}, & x > \mu \\ 0 & \text{otherwise.} \end{cases}$$

$$F_{x_1}(x) = \int_\mu^x e^{\mu - t}\, dt = 1 - e^{\mu - x}$$

$E(x_i) = \mu + 1$.

$E(\bar{x}) = \mu + 1$

$\bar{x}$ is consistent estimator for $\mu + 1$

$T_1 = \bar{X} - 1$ is unbiased & consistent for estimating μ.

Consider $Y = x_{(1)} = \min\{x_1, x_2, \ldots, x_n\}$

random sample so they are iid.

$$\begin{aligned} F_{x_{(1)}}(x) &= P(x_{(1)} \le x) \\ &= 1 - P(x_{(1)} > x) \\ &= 1 - P(x_1 > x)\, P(x_2 > x) \cdots P(x_n > x) \\ &= 1 - \left[1 - F_{x_1}(x)\right]^n \\ &= 1 - e^{n(\mu - x)}, \quad x > \mu \end{aligned}$$

so, the probability density of $x_{(1)}$ is

$f_{X_{(1)}}(x) = n e^{n(\mu - x)}, \quad x > \mu$

$E(X_{(1)}) = \mu + \frac{1}{n}$.

$T_2 = X_{(1)} - 1/n$. Then T_2 is also unbiased for μ

$V(X_{(1)}) = \frac{1}{n^2}$

$V(T_2) = \frac{1}{n^2} \to 0$ as $n \to \infty$

So, T_2 is also consistent for μ.

Here we have two ~~es~~ unbiased & consistent estimators.

2. Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform distⁿ on $[0, \theta]$.

$$f_{X_i}(x) = \begin{cases} \frac{1}{\theta}, & 0 \le x_i \le \theta \\ 0, & \text{otherwise} \end{cases}$$

$E(X_{(i)}) = \frac{\theta}{2}$

$$F_{X_i}(x) = \begin{cases} 0, & x_i < 0 \\ \frac{x}{\theta}, & 0 \le x \le \theta \\ 1, & x > \theta. \end{cases}$$

$E(\bar{X}) = \frac{\theta}{2}$

$\Rightarrow E(2\bar{X}) = \theta$

$T_1 = 2\bar{X}$, then T_1 is unbiased & consistent for θ.

$X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$

$$F_{X_{(n)}}(x) = P(X_{(n)} \le x) = P(X_1 \le x, X_2 \le x, \ldots, X_n \le x)$$

$$= [F_{X_i}(x)]^n$$

$$= \begin{cases} 0, & \text{if } x < 0 \\ (x/\theta)^n, & 0 \le x \le \theta \\ 1, & x > \theta. \end{cases}$$

$$f_{X_{(n)}}(x) = \begin{cases} \frac{n x^{n-1}}{\theta^n}, & 0 \le x \le \theta. \\ 0, & ~~x > \theta~~ \text{ otherwise} \end{cases}$$

Scanned by TapScanner

$$E(X_{(n)}) = \int_0^\theta x \cdot f_{X_{(n)}}(x)\,dx$$

$$= \int_0^\theta \frac{n x^n}{\theta^n}\,dx = \frac{n\, x^{n+1}}{(n+1)\,\theta^n} = \left(\frac{n}{n+1}\right)\theta.$$

$$E\left(\frac{n+1}{n} X_{(n)}\right) = \theta$$

$T_2 = \frac{n+1}{n} X_{(n)}$. Then T_2 is unbiased for θ.

$$\lim_{n\to\infty} F_{X_{(n)}}(x) = \begin{cases} 0, & \text{for } x<\theta \\ 1, & x \geq \theta. \end{cases}$$

This is the cdf of a RV which takes value θ with probability 1. So, $X_{(n)} \xrightarrow{P} \theta$

This fact can also be proved by considering

$$P(|X_{(n)}-\theta|>\varepsilon) = P(\theta - X_{(n)} > \varepsilon)$$

$$= P(X_{(n)} < \theta - \varepsilon).$$

$$= \left(\frac{\theta-\varepsilon}{\theta}\right)^n \to 0 \text{ as } n\to\infty.$$

So, $T_3 = X_{(n)}$ is consistent for θ.

$\Rightarrow$ T_2 is also consistent for θ as $\frac{n+1}{n} \to 1$ as $n\to\infty$

So, we have two different unbiased & consistent estimators.

3. Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous popn with cdf $F(x)$ and the range of variables is interval $[a,b]$.

Let $U = X_{(1)}$, $V = X_{(n)}$.

Then U is consistent for a & V is consistent for b.

$$\Rightarrow F_U(u) = \begin{cases} 0, & x<a \\ 1-[1-F(u)]^n, & a\leq x<b \\ 1, & x\geq b \end{cases}$$

$$F_V(v) = \begin{cases} 0, & x < a \\ [F(v)]^n, & a \le x < b \\ 1, & x \ge b \end{cases}$$

$\lim_{n \to \infty} F_V(v) = 0, \quad x < a$
$= 1, \quad x \ge a$ } This is cdf of a r.v. degenerate at a.

so, $V \xrightarrow{d} a \Rightarrow U \xrightarrow{P} a$.

$\lim_{n \to \infty} F_V(v) = 0, \quad x < b$
$= 1, \quad x \ge b$ } degenerate at b

$V \xrightarrow{d} b \Rightarrow V \xrightarrow{P} b$

sample range $x_{(n)} - x_{(1)}$

$\therefore V - U$ is consistent ~~esi test~~ estimator for the popn range $(b-a)$.

If lower limit $-\infty$
upper " $+\infty$ $\rightarrow$ ~~that~~ In this case
$V \not\rightarrow -\infty$
$U \not\rightarrow +\infty$

<u>Efficiency of Estimators</u>

$\rightarrow$ no easy to evaluate

$E|T - g(\theta)| \rightarrow$ Mean absolute error.

$MSE(T) = E(T - g(\theta))^2 \rightarrow$ Mean squared error.

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

$= E\{T - E(T)\}^2 + \{E(T) - g(\theta)\}^2 + 2E\{T - E(T)\}\{E(T) - g(\theta)\}$

$= V(T) + B^2(T) + 0$

We ~~si~~ can say that estimator T_1 is better (more efficient) if $MSE(T_1) \le MSE(T_2) \; \forall \theta \in \Theta$

If T is unbiased for $g(\theta)$, then $MSE(T) = var(T)$.