Unit 5

Estimation

If $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ is a random sample of size 'n' from a large population $X_1, X_2, X_3, \dots \dots \dots \dots \dots, X_N$

of size 'N' then usually the sample mean and variance are denoted by $\bar{x}$ and s^2 and population mean and

variance are denoted by μ and σ^2, i.e. $\bar{x} = \frac{x_1+x_2+x_3+\cdots\dots\dots\dots+x_n}{n}$, $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$ and

$\mu = \frac{X_1+X_2+X_3+\cdots\dots\dots\dots+X_N}{N}$, $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2 = \frac{1}{N}\sum_{i=1}^{N} X_i^2 - \mu^2$.

Poll Que. If $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ is a random sample of size 'n' from a large population $X_1, X_2, X_3, \dots \dots \dots \dots \dots, X_N$ of size 'N' then which of the following is correct:

(a) $E(x_i) = \mu^2, E(\bar{x}) = \mu$ (b) $E(x_i) = \mu, E(\bar{x}) = \mu$ (c) $E(x_i) = \mu, E(\bar{x}) = \mu^2$ (d) $None\ of\ these.$

Poll Que. If $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ is a random sample of size 'n' from a large population $X_1, X_2, X_3, \dots \dots \dots \dots \dots, X_N$ of size 'N' then which of the following is correct:

(a) $V(x_i) = \frac{\sigma^2}{n}, V(\bar{x}) = \sigma^2$ (b) $V(x_i) = \sigma^2, V(\bar{x}) = \sigma^2$ (c) $V(x_i) = \sigma^2, V(\bar{x}) = \frac{\sigma^2}{n}$ (d) $None\ of\ these.$

Poll Que. If $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ is a random sample of size 'n' from a large population $X_1, X_2, X_3, \dots \dots \dots \dots \dots, X_N$

of size 'N' then which of the following is correct:

(a) $E(s^2) = \left(\frac{n-1}{n}\right)\sigma^2$ (b) $E(s^2) = \left(\frac{n+1}{n}\right)\sigma^2$ (c) $E(s^2) = \sigma^2$

(d) $None\ of\ these.$

Theorem 6·9. *Let $X_1, X_2, \ldots, X_n$ be n random variables then*

$$V\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 V(X_i) + 2\sum_{\substack{i=1 \\ i<j}}^{n}\sum_{j=1}^{n} a_i a_j \, Cov(X_i, X_j)$$

Que. If X_1 and X_2 are two random variables, a_1 and a_2 are two constants then $V(a_1X_1 + a_2X_2) =$?

(a) $a_1^2V(X_1) + a_2^2V(X_2) - 2a_1^2a_2^2Cov(X_1, X_2)$ (b) $a_1^2V(X_1) + a_2^2V(X_2) + 2a_1^2a_2^2Cov(X_1, X_2)$

(c) $a_1^2V(X_1) + a_2^2V(X_2) - 2a_1a_2Cov(X_1, X_2)$ (d) $a_1^2V(X_1) + a_2^2V(X_2) + 2a_1a_2Cov(X_1, X_2)$.

Remark: If $x_1, x_2, x_3, \ldots \ldots \ldots \ldots \ldots, x_n$ is a random sample of size 'n' from a large population $X_1, X_2, X_3, \ldots \ldots \ldots \ldots \ldots, X_N$ of size 'N' then $E(x_i) = \mu, E(\bar{x}) = \mu, V(x_i) = \sigma^2, V(\bar{x}) = \frac{\sigma^2}{n}, E(s^2) = \left(\frac{n-1}{n}\right)\sigma^2, E(S^2) = \sigma^2$ where $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

Remark: If $x_1, x_2, x_3, \ldots \ldots \ldots \ldots \ldots, x_n$ is a random sample of size 'n' from a large population $X_1, X_2, X_3, \ldots \ldots \ldots \ldots \ldots, X_N$ of size 'N' then $E(x_i) = \mu, E(\bar{x}) = \mu, V(x_i) = \sigma^2, V(\bar{x}) = \frac{\sigma^2}{n}, E(s^2) = \left(\frac{n-1}{n}\right)\sigma^2, E(S^2) = \sigma^2$ where $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

12·11. Unbiased Estimate for population Mean (μ) and Variance (σ^2). Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a large population $X_1, X_2, \ldots X_N$ (of size N) with mean μ and variance σ^2. Then the sample mean ($\bar{x}$) and variance (s^2) are given by

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad \text{and} \quad s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Now
$$E(\bar{x}) = E\left(\frac{1}{n} \sum_{i=1}^{n} x_i\right) = \frac{1}{n} \sum_{i=1}^{n} E(x_i)$$

Since x_i is a sample observation from the population X_i, ($i = 1, 2, \ldots, N$) it can take any one of the values $X_1, X_2, \ldots, X_N$ each with equal probability $1/N$.

$$\therefore \quad E(x_i) = \frac{1}{N} X_1 + \frac{1}{N} X_2 + \ldots + \frac{1}{N} X_N$$

$$= \frac{1}{N} (X_1 + X_2 + \ldots + X_N) = \mu \qquad \ldots(1)$$

$$\therefore \quad E(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} (\mu) = \frac{1}{n} n\mu \quad \Rightarrow \quad E(\bar{x}) = \mu \qquad \ldots(12\cdot6)$$

Thus the sample mean ($\bar{x}$) is an unbiased estimate of the population mean (μ).

Now $\quad E(s^2) = E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2\right]$

$$= \frac{1}{n}\sum_{i=1}^{n} E(x_i^2) - E(\bar{x})^2 \qquad \ldots(2)$$

We have $V(x_i) = E[x_i - E(x_i)]^2 = E(x_i - \mu)^2$, $\qquad$ [From (1)]

$$= \frac{1}{N}\left[(X_1 - \mu)^2 + (X_2 - \mu)^2 + \ldots + (X_N - \mu)^2\right] = \sigma^2 \qquad \ldots(3)$$

Also we know that

$$V(x) = E(x^2) - [E(x)]^2 \Rightarrow E(x^2) = V(x) + \{E(x)\}^2 \qquad \ldots(4)$$

In particular

$$E(x_i^2) = V(x_i) + \{E(x_i)\}^2 = \sigma^2 + \mu^2 \qquad \ldots(5)$$

Also from (4), $E(\bar{x}^2) = V(\bar{x}) + \{E(\bar{x})\}^2$

But $\quad V(\bar{x}) = \frac{\sigma^2}{n}$, where σ^2 is the population variance. $\quad$ [*c.f.* § 12·13]

$\therefore \quad E(\bar{x}^2) = \frac{\sigma^2}{n} + \mu^2 \qquad$ [Using (12·6)] $\qquad \ldots(5a)$

$$E(s^2) = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2+\mu^2) - \left(\frac{\sigma^2}{n}+\mu^2\right)$$

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

$$= \frac{n-1}{n}\sigma^2 \qquad \ldots(12{\cdot}7)$$

Since $E(s^2) \neq \sigma^2$, sample variance is not an unbiased estimate of population variance.

From (12·7), we get

$$\frac{n}{n-1}E(s^2) = \sigma^2 \;\Rightarrow\; E\left(\frac{ns^2}{n-1}\right) = \sigma^2$$

$$\Rightarrow \quad E\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right] = \sigma^2 \;\; i.e., \;\; E(S^2) = \sigma^2 \qquad \ldots(12{\cdot}8)$$

where $$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad \ldots(12{\cdot}8a)$$

$\therefore$ S^2 is an unbiased estimate of the population variance σ^2.

Proof. Let x_i, $(i = 1, 2, \ldots, n)$ be a random sample of size n from a population with variance σ^2, then the sample mean $\bar{x}$ is given by

$$\bar{x} = \frac{1}{n}(x_1+x_2+\ldots+x_n)$$

$$\therefore \quad V(\bar{x}) = V\left[\frac{1}{n}(x_1+x_2+\ldots+x_n)\right] = \frac{1}{n^2}V(x_1+x_2+\ldots+x_n)$$

$$= \frac{1}{n^2}\left[V(x_1)+V(x_2)+\ldots+V(x_n)\right],$$

the covariance terms vanish since the sample observations are independent, [*c.f.* Remark (*ii*) § 6·6]

But $V(x_i) = \sigma^2, (i = 1, 2, \ldots, n)$ [From (3) of § 12·11]

$$\therefore \quad V(\bar{x}) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$$

$$\Rightarrow \qquad \text{S.E.}(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} \qquad \ldots(12{\cdot}9)$$

Characteristics of Estimators. The following are some of the criteria that should be satisfied by a good estimator.
(i) Unbiasedness *(ii) Consistency* *(iii) Efficiency* *(iv) Sufficiency.*

Definition: Any function of random sample $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ that are being observed, say, $\mathrm{T_n}(x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n)$ is called a **statistic**. Clearly a statistic is a random variable. If it is used
to estimate an unknown **parameter '$\boldsymbol{\theta}$'** of the distribution then it is called an **estimator**. A particular
value of the estimator, say, is called an **Estimate of '$\boldsymbol{\theta}$'**.

Definition: A statistic $\mathrm{T_n}(x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n)$ is said to be an Unbiased estimator of parameter $\gamma(\theta)$
if $E(T_n) = \gamma(\theta)$ for all $\theta \in \Theta$.

Remark: If $E(T_n) > \gamma(\theta)$ then is said to be **positively biased** and if $E(T_n) < \gamma(\theta)$ then is said to be **negatively biased.**
The amount of bias $b(\theta) = E(T_n) - \gamma(\theta)$.

Poll Que. If $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ is a random sample of size 'n' from a No
then $t = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ is an unbiased estimator of:
(a) $1 + \mu$ (b) $1 + \mu^2$ (c) $n + \mu^2$ (d) $None\ of\ these.$

Poll Que. Which of the following is/are correct?
(a) Sample mean is an Unbiased Estimator of population mean.

(b) Sample variance is an Unbiased Estimator of population variance.

(c) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ is an Unbiased Estimator of population variance.

(d) $None\ of\ these.$

Poll Que. If $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ is a random sample of size 'n' from a Normal population $N(\mu, 1)$
then $t = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ is an unbiased estimator of:
(a) $1 + \mu$ (b) $1 + \mu^2$ (c) $n + \mu^2$ (d) $None\ of\ these.$

Remark: If 'T' is an unbiased estimator of 'θ' then 'T^2' is a biased estimator of 'θ^2'.

7·1. Bernoulli Distribution. A random variable X which takes two values 0 and 1, with probabilities q and p respectively, *i.e.*, $P(X = 1) = p$, $P(X = 0) = q$, $q = 1 - p$ is called a *Bernoulli variate* and is said to have a Bernoulli distribution.

Remark: (i) If in binomial distribution, number of trials 'n' $= 1$ then the binomial distribution becomes
Bernoulli distribution.
(ii) Sum of 'n' independent Bernoulli variates is a binomial variate i.e. if $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$
is a random sample of size 'n' drawn from Bernoulli population with parameter 'θ' then $T = \sum_{i=1}^{n} x_i \sim B(n, \theta)$.

Poll Que. If $x_1, x_2, x_3, \dots \dots \dots \dots \dots, x_n$ is a random sample of size 'n' drawn on 'X' which takes the values
'1' or '0' with respective probabilities 'θ' and '$(1 - \theta)$' then $t = \frac{[\sum x_i(\sum x_i - 1)]}{n(n-1)}$ is an unbiased estimator of:
(a) $1 + \theta$ (b) $1 + \theta^2$ (c) $n + \theta^2$ (d) $None\ of\ these.$

Poll Que. Let 'X' be distributed in the Poisson form with parameter 'θ'.Then the only unbiased estimate of $\exp\{-(k+1)\theta\}, k > 0$ is:

(a) $T(X) = (-k)^{-X}$ (b) $T(X) = (k)^{-X}$ (c) $T(X) = (-k)^{X}$ (d) *None of these.*

Example 15·4. *Let X be distributed in the Poisson form with parameter θ. Show that the only unbiased estimator of exp $[-(k+1)\theta]$, $k > 0$, is $T(X) = (-k)^X$ so that*

$$T(x) > 0 \text{ if } x \text{ is even}$$

and

$$T(x) < 0 \text{ if } x \text{ is odd.}$$

15·3. Consistency. An estimator $T_n = T(x_1, x_2, \ldots, x_n)$, based on a random sample of size n, is said to be consistent estimator of $\gamma(\theta)$, $\theta \in \Theta$, the parameter space, if T_n converges to $\gamma(\theta)$ in probability.

i.e., if
$$T_n \xrightarrow{p} \gamma(\theta) \text{ as } n \to \infty \qquad \ldots(15{\cdot}1)$$

In other words, T_n is a consistent estimator of $\gamma(\theta)$ if for every $\varepsilon > 0$, $\eta > 0$, there exists a positive integer $n \geq m(\varepsilon, \eta)$ such that

$$P\left[|T_n - \gamma(\theta)| < \varepsilon\right] \to 1 \text{ as } n \to \infty \qquad \ldots(15{\cdot}2)$$

$$\Rightarrow \quad P\left[|T_n - \gamma(\theta)| < \varepsilon\right] > 1 - \eta\ ;\ \forall\ n \geq m \qquad \ldots(15{\cdot}2a)$$

where m is some very large value of n.

15·4·2. Sufficient Conditions for Consistency.

Theorem 15·2. Let $\{T_n\}$ *be a sequence of estimators such that for all* $\theta \in \Theta$,

(i) $E_\theta(T_n) \to \gamma(\theta),\ n \to \infty$

and *(ii)* $Var_\theta(T_n) \to 0$, *as* $n \to \infty$.

Then T_n is a consistent estimator of $\gamma(\theta)$.

Poll Que. Which of the following is/are correct?

(a) Sample mean is only an Unbiased Estimator of population mean.

(b) Sample mean is only a consistent Estimator of population mean.

(c) Sample mean is both Unbiased and consistent Estimator of population mean.

(d) $None\ of\ these.$

Example 15·5. *(a) Prove that in sampling from a $N(\mu, \sigma^2)$ population, the sample mean is a consistent estimator of μ.*

Unbiasedness is a property associated with finite n. A statistic $T_n = T(x_1, x_2, \ldots, x_n)$, is said to be an unbiased estimator of $\gamma(\theta)$ if

$$E(T_n) = \gamma(\theta), \text{ for all } \theta \in \Theta$$

15·4·1. Invariance Property of Consistent Estimators.

Theorem 15·1. *If T_n is a consistent estimator of $\gamma(\theta)$ and $\psi(\gamma(\theta))$ is a continuous function of $\gamma(\theta)$, then $\psi(T_n)$ is a consistent estimator of $\psi(\gamma(\theta))$.*

Poll Que. If $x_1, x_2, x_3, \ldots \ldots \ldots \ldots \ldots, x_n$ are random observations on a Bernoulli variate 'X' taking the value '1' with probability 'p' and the value '0' with probability '$(1 - p)$', then $\frac{\sum x_i}{n}\left(1 - \frac{\sum x_i}{n}\right)$ is a consistent estimator of:

(a) $p(1-p)$ (b) $1-p$ (c) p (d) $None\ of\ these.$

Example 15·6. *If $X_1, X_2, \ldots, X_n$ are random observations on a Bernoulli variate X taking the value 1 with probability p and the value 0 with probability (1 − p), show that :*

$$\frac{\sum x_i}{n}\left(1 - \frac{\sum x_i}{n}\right) \text{ is a consistent estimator of } p(1-p).$$

Solution. Since $X_1, X_2, \ldots, X_n$ are *i.i.d* Bernoulli variates with parameter 'p',

$$T = \sum_{i=1}^{n} x_i \sim B(n, p)$$

$$\Rightarrow \qquad E(T) = np \quad \text{and} \quad \text{Var}(T) = npq$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{T}{n}$$

$$\therefore \qquad E(\bar{X}) = \frac{1}{n} E(T) = \frac{1}{n} \cdot np = p$$

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{T}{n}\right) = \frac{1}{n^2} \cdot \text{Var}(T) = \frac{pq}{n} \to 0 \text{ as } n \to \infty.$$

Since $E(\bar{X}) \to p$ and $\text{Var}(\bar{X}) \to 0$, as $n \to \infty$; $\bar{X}$ is a consistent estimator of p.

Also $\dfrac{\Sigma x_i}{n}\left(1 - \dfrac{\Sigma x_i}{n}\right) = \bar{X}\ (1 - \bar{X})$, being a polynomial in $\bar{X}$, is a continuous function of $\bar{X}$.

Since $\bar{X}$ is consistent estimator of p, by the invariance property of consistent estimators (Theorem 15·1), $\bar{X}\ (1 - \bar{X})$ is a consistent estimator of $p\,(1 - p)$.

15·5. Efficient Estimators. *Efficiency*. Even if we confine ourselves to unbiased estimates, there will, in general, exist more than one consistent estimator of a parameter. For example, in sampling from a normal population $N(\mu, \sigma^2)$, when σ^2 is known, sample mean $\bar{x}$ is an unbiased and consistent estimator of μ [*c.f.* Example 15·5*a*].

$$\therefore \qquad V(Md) = \frac{1}{4n} \cdot 2\pi\sigma^2 = \frac{\pi\sigma^2}{2n}$$

$$\left.\begin{array}{ll}\text{Since} & E(Md) = \mu \\ \text{and} & V(Md) \to 0\end{array}\right\} \text{, as } n \to \infty$$

median is also an unbiased and consistent estimator of μ.

If, of the two consistent estimators T_1, T_2 of a certain parameter θ, we have

$$V(T_1) < V(T_2), \text{ for all } n \qquad \ldots(15{\cdot}11)$$

then T_1 is more efficient than T_2 for all samples sizes.

We have seen above :

For all n, $\quad V(\bar{x}) = \dfrac{\sigma^2}{n}$

and for large n, $\quad V(Md) = \dfrac{\pi\sigma^2}{2n} = 1{\cdot}57\,\dfrac{\sigma^2}{n}$

Since $V(\bar{x}) < V(Md)$, we conclude that for normal distribution, sample mean is more efficient estimator for μ than the sample median, for large samples at least.

15·5·1. Most Efficient Estimator. *If in a class of consistent estimators for a parameter, there exists one whose sampling variance is less than that of any such estimator, it is called the most efficient estimator. Whenever such an estimator exists, it provides a criterion for measurement of efficiency of the other estimators.*

Efficiency *(Def.) 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 of T_2 is defined as :*

$$E = \frac{V_1}{V_2} \qquad \ldots(15\cdot12)$$

Obviously, E cannot exceed unity.

If $T, T_1, T_2, \ldots, T_n$ are all estimators of $\gamma(\theta)$ and $\text{Var}(T)$ is minimum, then the efficiency E_i of T_i, $(i = 1, 2, \ldots, n)$ is defined as :

$$E_i = \frac{\text{Var } T}{\text{Var } T_i}\,;\ i = 1, 2, \ldots, n \qquad \ldots(15\cdot12a)$$

Obviously $E_i \leq 1, i = 1, 2, \ldots n$.

For example, in the normal samples, since sample mean $\bar{x}$ is the most efficient estimator of μ [c.f. Remark to Example 15·31], the efficiency E of Md for such samples, (for large n), is :

$$E = \frac{V(\bar{x})}{V(Md)} = \frac{\sigma^2/n}{\pi\sigma^2/(2n)} = \frac{2}{\pi} = 0\cdot637$$

Example 15·7. *A random sample $(X_1, X_2, X_3, X_4, X_5)$ of size 5 is drawn from a normal population with unknown mean μ. Consider the following estimators to estimate μ :*

(i) $t_1 = \dfrac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$

(ii) $t_2 = \dfrac{X_1 + X_2}{2} + X_3$, (iii) $t_3 = \dfrac{2X_1 + X_2 + \lambda X_3}{3}$

where λ is such that t_3 is an unbiased estimator of μ.

Find λ. Are t_1 and t_2 unbiased ? State giving reasons, the estimator which is best among t_1, t_2 and t_3.

(a) $\lambda = -1$ (b) $\lambda = 0$ (c) $\lambda = 1$ (d) *None of these.*

(a) only 't_1' is unbiased (b) only 't_2' is unbiased (c) both 't_1' and 't_2' are unbiased (d) *None of these.*

(a) 't_1' is most efficient (b) 't_2' is most efficient (c) 't_3' is most efficient (d) *None of these.*

$$V(t_1) = \frac{1}{25}[V(X_1) + V(X_2) + V(X_3) + V(X_4) + V(X_5)] = \frac{1}{5}\sigma^2$$

$$V(t_2) = \frac{1}{4}[V(X_1) + V(X_2)] + V(X_3) = \frac{1}{2}\sigma^2 + \sigma^2 = \frac{3}{2}\sigma^2$$

$$V(t_3) = \frac{1}{9}[4V(X_1) + V(X_2)] = \frac{1}{9}(4\sigma^2 + \sigma^2) = \frac{5}{9}\sigma^2 \qquad (\because \lambda = 0)$$

Since $V(t_1)$ is the least, t_1 is the best estimator (in the sense of least variance) of μ.

Example 15·8. *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*

$T_1 = X_1 + X_2 - X_3$, $T_2 = 2X_1 + 3X_3 - 4X_2$, and $T_3 = (\lambda X_1 + X_2 + X_3)/3$

(i) *Are T_1 and T_2 unbiased estimators ?*

(ii) *Find the value of λ such that T_3 is unbiased estimator for μ.*

(iii) *With this value of λ is T_3 a consistent estimator ?*

(iv) *Which is the best estimator ?*

Poll Que. Which of the following is the probability density function of the random variable $X \sim N(\mu, \sigma^2)$?

(a) $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right)$ (b) $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right)$ (c) $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ (d) $\frac{1}{2\pi\sqrt{\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.

15·10. Methods of Estimation. So far we have been discussing the requisites of a good estimator. Now we shall briefly outline some of the important methods for obtaining such estimators. Commonly used methods are

(i) *Method of Maximum Likelihood Estimation.*

(ii) *Method of Minimum Variance.*

(iii) *Method of Moments.*

(iv) *Method of Least Squares.*

(v) *Method of Minimum Chi-square*

(vi) *Method of Inverse Probability.*

Likelihood Function. *Definition.* 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, given by

$$L = f(x_1, \theta) f(x_2, \theta) \ldots f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta). \quad \ldots(15{\cdot}53)$$

L gives the relative likelihood that the random variables assume a particular set of values $x_1, x_2, \ldots, x_n$. For a given sample $x_1, x_2, \ldots, x_n$, L becomes a function of the variable θ, the parameter.

The principle of maximum likelihood consists in finding an estimator for the unknown parameter $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$, say, which maximises the likelihood function $L(\theta)$ for variations in parameter *i.e.,* we wish to find $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$ so that

$$L(\hat{\theta}) > L(\theta) \qquad \forall\, \theta \in \Theta$$

i.e., $$L(\hat{\theta}) = \operatorname{Sup} L(\theta) \;\forall\, \theta \in \Theta.$$

Thus if there exists a function $\hat{\theta} = \hat{\theta}(x_1, x_2, \ldots, x_n)$ of the sample values which maximises L for variations in θ, then $\hat{\theta}$ is to be taken as an estimator of θ. $\hat{\theta}$ is usually called *Maximum Likelihood Estimator (M.L.E.).* Thus $\hat{\theta}$ is the solution, if any, of

$$\frac{\partial L}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial^2 L}{\partial \theta^2} < 0 \qquad \ldots(15{\cdot}54)$$

Since $L > 0$, and $\log L$ is a non-decreasing function of L; L and $\log L$ attain their extreme values (maxima or minima) at the same value of $\hat{\theta}$. The first of the two equations in (15·54) can be rewritten as

$$\frac{1}{L} \cdot \frac{\partial L}{\partial \theta} = 0 \quad \Rightarrow \quad \frac{\partial \log L}{\partial \theta} = 0, \qquad \ldots(15{\cdot}54a)$$

a form which is much more convenient from practical point of view.

If θ is vector valued parameter, then $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$, is given by the solution of simultaneous equations :

$$\frac{\partial}{\partial \theta_i} \log L = \frac{\partial}{\partial \theta_i} \log L (\theta_1, \theta_2, \ldots, \theta_k) = 0 \; ; \; i = 1, 2, \ldots, k \qquad \ldots(15{\cdot}54b)$$

Equations (15·54*a*) and (15·54*b*) are usually referred to as the *Likelihood Equations* for estimating the parameters.

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

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

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

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

Poll Que. Which of the following is the probability density function of the random variable $X \sim N(\mu, \sigma^2)$?

(a) $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right)$ (b) $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right)$ (c) $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ (d) $\frac{1}{2\pi\sqrt{\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.

Poll Que. If $x_1, x_2, x_3, \ldots \ldots \ldots \ldots \ldots, x_n$ is a random sample of size 'n' from a Normal population $N(\mu, \sigma^2)$ then $\log L = ?$

(a) $-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$ (b) $-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$

(c) $-\frac{n}{2}\log(\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$ (d) $-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$.

Example 15·31. *In random sampling from normal population $N(\mu; \sigma^2)$, find the maximum likelihood estimators for*

(i) μ when σ^2 is known,

(ii) σ^2 when μ is known, and

(iii) the simultaneous estimation of μ and σ^2.

For part (i) (a) MLE for 'μ' when 'σ^2' is known is $\frac{\bar{x}}{2}$ (b) MLE for 'μ' when 'σ^2' is known is $\bar{x}$

(c) MLE for 'μ' when 'σ^2' is known is $2\bar{x}$ (d) $None\ of\ these.$

For part (ii) (a) MLE for 'σ^2' when 'μ' is known is $\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$

(b) MLE for 'σ^2' when 'μ' is known is $\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2$

(c) MLE for 'σ^2' when 'μ' is known is $\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$

(d) $None\ of\ these.$

For part (iii) (a) MLE for 'μ' and 'σ^2' are $\bar{x}$ and s^2 respectively

(b) MLE for 'μ' and 'σ^2' are $\bar{x}$ and $\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$ respectively

(c) MLE for 'μ' and 'σ^2' are $\bar{x}$ and $\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ respectively (d) $None\ of\ these.$

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

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

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

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

Case (i). When σ^2 is known, the likelihood equation for estimating μ is

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

or

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

$$\Rightarrow \qquad \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x} \qquad \ldots(*)$$

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

Case (ii). When μ is known, the likelihood equation for estimating σ^2 is

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

$$\Rightarrow \qquad n - \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 = 0, \quad i.e., \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \qquad \ldots(**)$$

Case (iii). The likelihood equations for simultaneous estimation of μ and σ^2 are

$$\frac{\partial}{\partial \mu} \log L = 0 \quad \text{and} \quad \frac{\partial}{\partial \sigma^2} \log L = 0, \text{ thus giving}$$

$$\hat{\mu} = \bar{x} \qquad \text{[From (*)]}$$

and

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2 \qquad \text{[From (**)]}$$

$$= \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = s^2, \text{ the sample variance.}$$

Poll Que. (i) Which of the following is the maximum likelihood estimate of the parameter 'α' of

a population having density function: $\frac{2}{\alpha^2}(\alpha - x), 0 < x < \alpha$ when a sample of unit size is drawn

from the population? (Here 'x' is the sample value).

(a) $2x$ (b) x (c) $\frac{x}{2}$ (d) $None\ of\ these.$

Poll Que. (ii) Is the above MLE is biased?

(a) Yes (b) No.

Example 15·32. *Prove that the maximum likelihood estimate of the parameter α of a population having density function :*

$$\frac{2}{\alpha^2}(\alpha - x),\ 0 < x < \alpha$$

for a sample of unit size is 2x, x being the sample value. Show also that the estimate is biased. [*Burdwan Univ. B.Sc. (Maths. Hons.), 1991*]

Solution. For a random sample of unit size ($n = 1$), the likelihood function is :

$$L(\alpha) = f(x, \alpha) = \frac{2}{\alpha^2}(\alpha - x)\ ;\ 0 < x < \alpha$$

Likelihood equation gives :

$$\frac{d}{d\alpha}\log L = \frac{d}{d\alpha}\left[\log 2 - 2\log\alpha + \log(\alpha - x)\right] = 0$$

$$\Rightarrow \quad -\frac{2}{\alpha} + \frac{1}{\alpha - x} = 0 \quad \Rightarrow \quad 2(\alpha - x) - \alpha = 0 \quad \Rightarrow \quad \alpha = 2x$$

Hence MLE of α is given by $\hat{\alpha} = 2x$.

$$E(\hat{\alpha}) = E(2X) = 2\int_0^{\alpha} x . f(x, \alpha)\, dx$$

$$= \frac{4}{\alpha^2}\int_0^{\alpha} x(\alpha - x)\, dx = \frac{4}{\alpha^2}\left|\frac{\alpha x^2}{2} - \frac{x^3}{3}\right|_0^{\alpha} = \frac{2}{3}\alpha$$

Since $E(\hat{\alpha}) \neq \alpha$, $\hat{\alpha} = 2x$ is not an unbiased estimate of α.

Springer Ebooks.pdf

$P(-3 \le Z \le 3) = 0.9973$, *i.e.*, $P(|Z| \le 3) = 0.9973$ $P(|Z| > 3) = 1 - P(|Z| \le 3) = 0.0027$

$P(-1.96 \le Z \le 1.96) = 0.95$, *i.e.*, $P(|Z| \le 1.96) = 0.95$

$P(|Z| > 1.96) = 1 - P(|Z| \le 1.96) = 0.05$

$P(|Z| \le 2.58) = 0.99$ $P(|Z| > 2.58) = 0.01$