UNIT-III.
ESTIMATION THEORY

3.0 INTRODUCTION

A statistic whose value is used as the point estimate of a parameter θ is called an estimator.

An estimate of the parameter is a particular numerical value of the estimator obtained by sampling.

With respect to estimating a parameter, the following two types of estimates are

(i) Point estimation

(ii) Interval estimation

3.1 POINT ESTIMATION

If an estimate of a population parameter θ is given by a single value, then the estimate is called a point estimate of the parameter.

For example, the sample mean $\bar{x}$ is a point estimate of a population mean μ, sample variance s^2 is a point estimate of a population variance σ^2.

Example of Point Estimate is Given below :

A single value of a statistic is a point estimate of a population parameter. The sample mean x, for example, is a point estimate of the population mean μ. The sample proportion p is a point estimate of the population proportion P, in the same way.

Common Methods of finding point estimates

The method of point estimation entails using the value of a statistic derived from survey data to achieve the best approximation of the population's corresponding unknown parameter. The point estimators can be calculated using a variety of methods, each with its own set of properties.

Given below are two methods of find point estimates :

1. **Method of moments :** It begins with known facts about a population, which are then applied to a sample of the population. The first step is to create equations that connect population moments to unknown parameters.

   The next step is to select a population sample from which to estimate population moments. The sample mean of the population moments is used to solve the equations obtained in step one. This yields the most accurate estimation of unknown population parameters.

If $f(x;\theta_1, \theta_2, ..., \theta_k)$ is the density of the population with k parameters $\theta_1, \theta_2, ..., \theta_k$, our interest is to find the functions of sample moments that will estimate the parameters θ_1, θ_2, ..., θ_k. Let $f(x;\theta_1, \theta_2, ..., \theta_k)$ be the density function of the population, from which the sample $(x_1, x_2, ..., x_n)$ of size n is drawn. If μ_r' denotes the r^{th} order moment about the origin, then it is given by

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x;\theta_1, \theta_2, ..., \theta_k)\, dx, \quad (r = 1, 2, ..., k) \qquad ... (1)$$

In general, $\mu_1', \mu_2', ..., \mu_k'$ will be function of the parameters $\theta_1, \theta_2, ..., \theta_k$

Solving the k-equations in (1), we get θ_1, θ_2, ..., θ_k in terms of $\mu_1', ..., \mu_2', ..., \mu_k'$

Hence,

$$\begin{aligned} \overline{\theta}_i &= \theta_i(\mu_1', \mu_2', ..., \mu_k') \; ; \; i = 1, 2, ..., k \\ &= \theta_i(m_1', m_2', ..., m_k') \qquad ... (2) \end{aligned}$$

Equation (2) is obtained by equating μ_r' by m_r' where m_r' is the r^{th} order moment of the sample about the origin.

Note 1. If we wish to estimate the parameters, we should compute k moments about the origin.

Note 2. Method of moment estimators are usually less efficient than MLEs.

2. **Maximum likelihood estimator :** The maximum likelihood estimator is a point estimation tool that tries to find unknown parameters that increase the likelihood function. It compares data sets and finds the best fit for the data by using the values from a proven model.

For example, a study may be curious about the average weight of premature babies. Because measuring all premature babies in the population will be impractical, the researcher will take a sample from a single area. The researcher will use the maximum likelihood estimator to find the average weight of the entire population of pre-term babies based on the sample data since the weight of pre-term babies follows a normal distribution.

Definition :

A statistic $\hat{\theta}(X_1, ..., X_n)$ is a **maximum likelihood estimator** of θ

if, for each sample $x_1, ..., x_n$, $\overline{\theta}(x_1, ..., x_n)$ is a value for the parameter that maximizes the likelihood function $L(\theta \mid x_1, ..., x_n)$

Method of maximum likelihood

According to this method, a good estimator is the one which maximizes the likelihood function $L = f(x_1, \theta); f(x_2, \theta), ..., f(x_n, \theta)$. It is called the Maximum Likelihood Estimator (MLE).

If t is the MLE for the population parameter θ, then it satisfies

$$\frac{\partial L}{\partial \theta} = 0 \text{ and } \frac{\partial^2 L}{\partial \theta^2} = 0.$$

Since L is positive, $\frac{\partial L}{\partial \theta} = 0$ is equivalent to $\frac{1}{L}\frac{\partial L}{\partial \theta} = 0$ (or) $\frac{\partial}{\partial \theta}(\log L) = 0$.

$\frac{\partial}{\partial \theta}(\log L) = 0$ is more convenient in practice and it is usually referred to as the likelihood equation.

Note : The MLE of a parameter θ is usually denoted as $\overline{\theta}$

Properties of maximum likelihood estimators

Here some of the important properties of maximum likelihood estimators, without proof :

1. MLEs are consistent.
2. MLEs are most efficient.

3. MLEs are sufficients, if sufficient estimators exist.
4. MLEs are not necessarily unbiased.
5. MLEs have the invariance property, viz., if $\overline{\theta}$ is a MLE for θ, then $f(\overline{\theta})$ will be a MLE for $f(\theta)$
6. The distribution of MLE tends to normally for large samples.

CHARACTERISTICS OF ESTIMATION

A good point estimator should satisfy the following criteria

(1) Unbiasedness

(2) Consistency

(3) Efficiency

(4) Sufficiency

(5) Robustness

(1) **Unbiasedness** : Let $\overline{\theta}$ be an estimator of the parameter θ then the estimator $\overline{\theta}$ $(x_1, x_2, x_3, \ldots, x_n)$ is said to be an unbiased estimate of parameter θ.

$$\text{If } E(\overline{\theta}) = \theta \ \forall \ \theta$$

If $E(\overline{\theta}) > \theta$ then $\overline{\theta}$ is said to be positively biased

If $E(\overline{\theta}) < \theta$ then $\overline{\theta}$ is said to be negatively biased

$\{E(\overline{\theta} - \theta\}$ is called the amount of bias is denoted by $b(\theta)$

$$\therefore b(\theta) = \{E(\overline{\theta}) - \theta\}$$

Examples : Since $E(\overline{x}) = \mu$, the sample mean is an unbiased estimate of a population mean μ.

Again $E(S^2) \neq \sigma^2$ where $S^2 = \frac{1}{n} \sum (x_i - \overline{x})^2$ is not an unbiased estimate of σ^2.

But $E(S^2) \neq \sigma^2$ where $S^2 = \frac{1}{n-1} \sum (x_i - \overline{x})^2$ is an unbiased estimate of σ^2. Also observe that $S^2 = \frac{n}{n-1} S^2$

(2) **Consistency** : A statistic $\overline{\theta}$ is obtained from a sample of size 'n' is said to be consistent estimator of the parameter θ, if it comes closer and closer to the parameter when n becomes larger and larger i.e., $\lim_{n\to\infty} p\left[|\overline{\theta}-\theta| \le \varepsilon\right] = 1$ for all $\varepsilon > 0$

i.e., the estimator $\overline{\theta}$ is consistent if it converges to θ, as the sample size n increases.

(3) **Efficiency** : We may get two or more consistent estimators for the same parameter θ. If two consistent estimators exist for the same parameter θ, then the statistic with the smaller variance is called an efficient estimator of θ, while the other statistic is called an inefficient estimator of θ (i.e.,) if for two consistent estimators $\overline{\theta}_1$ and $\overline{\theta}_2$ for θ we have $V(\overline{\theta}_1) < V(\overline{\theta}_2)$, then $\overline{\theta}_1$ is an efficient estimator of θ.

Of all the consistent estimators of a parameter θ, the one with the smallest variance is known as the most efficient estimator or the best estimator of θ.

(i.e.,) If $\overline{\theta}_1, \overline{\theta}_2, \overline{\theta}_3, \ldots$ are consistent estimators for θ and the smallest of $V(\overline{\theta}_1), V(\overline{\theta}_2), V(\overline{\theta}_3), \ldots$ then the corresponding estimate is the most efficient estimator or the best estimator of θ.

Note : If $\overline{\theta}_1$ is the most efficient estimator with variance V_1 and $\overline{\theta}_2$ is any other estimator with variance V_2 then the efficiency E of $\overline{\theta}_2$ is defined $E = \frac{V_1}{V_2}$ which is always less than unity.

(4) **Sufficiency** : An estimator is said to be sufficient for a parameter θ, if it contains all the information in the sample regarding the parameter. In other words if $\overline{\theta} = \overline{\theta}(x_1, x_2, x_3, \ldots, x_n)$ is an estimator of a parameter θ, based on a sample $x_1, x_2, x_3, \ldots, x_n$ of size n from the population with density $f(x, \theta)$ such that the conditional distribution of $x_1, x_2, x_3, \ldots, x_n$ given $\overline{\theta}$ is independent of θ then θ is sufficient estimator of θ.

(5) **Robustness** : An estimation technique which is insensitive to small departures from the idealized assumptions which have been used to optimize the algorithm. Classes of such techniques include M-estimates (which follow from maximum likelihood considerations), L-Estimates (which are linear combinations of order statistics), and R-Estimates (based on statistical rank tests).

EXAMPLES :

The **mean** is not a robust measure of **central tendency.** If the dataset is e.g., the values $\{2, 3, 5, 6, 9\}$, then if we add another datapoint with value -1000 or $+1000$ to the data, the resulting mean will be very different to the mean of the original data. Similarly, if we replace one of the values with a datapoint of value -1000 or $+1000$ then the resulting mean will be very different to the mean of the original data.

The **median** is a robust measure of **central tendency**. Taking the same dataset $\{2, 3, 5, 6, 9\}$, if we add another datapoint with value -1000 or $+1000$ then the median will change slightly, but it will still be similar to the median of the original data. If we replace one of the values with a datapoint of value -1000 or $+1000$ then the resulting median will still be similar to the median of the original data.

Described in terms of **breakdown points**, the median has a breakdown point of 50%, meaning that half the points must be outliers before the median can be moved outside the range of the non-outliers, while the mean has a breakdown point of 0, as a single large observation can throw it off.

The **median absolute deviation** and **interquartile range** are robust measures of **statistical dispersion**, while the **standard deviation** and **range** are not.

Trimmed estimators and **Winsorised estimators** are general methods to make statistics more robust. **L-estimators** are a general class of simple statistics, often robust, while **M-estimators** are a general class of robust statistics, and are now the preferred solution, though they can be quite involved to calculate.

VARIETY OF APPLICATIONS

Robust methods also exist for regression problems, generalized linear models, and parameter estimation of various distributions.

Robust **parametric statistics** can proceed in two ways :

* By designing estimators so that a pre-selected behaviour of the influence function is achieved.
* By replacing estimators that are optimal under the assumption of a normal distribution with estimators that are optimal for, or at least derived for, other distributions; for example using the t-distribution with low degrees of freedom (high kurtosis; degrees of freedom between 4 and 6 have often been found to be useful in practice) or with a **mixture** of two or more distributions.

Robust estimates have been studied for the following problems :

- Estimating **location parameters**
- Estimating **scale parameters**
- Estimating **regression coefficients**
- Estimation of model-states in models expressed in **state-space** form, for which the standard method is equivalent to a **Kalman filter.**

WORKED EXAMPLES

Example 1.

Show that the sample mean $\bar{x}$ is an unbiased estimator for the population mean μ.

Solution :

To prove : $E[\bar{x}] = \mu$

Let $x_1, x_2, ..., x_n$ be a random sample of size n

$$\bar{x} = \frac{\Sigma x_i}{n}$$

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

$$= \frac{1}{n}\left[E(x_1) + E(x_2) + ... + E(x_n)\right]$$

$$= \frac{1}{n}[\mu + \mu + ... + \mu] \qquad [\because E[x_i] = \mu]$$

$$= \frac{1}{n}[n\mu] = \mu$$

$\Rightarrow \bar{x}$ is an unbiased estimator of μ.

Example 2.

If 'X' is a binomial variate with parameters n and p, then show that $\frac{X}{n}$, the observed proportion of successes is an unbiased estimator of the parameter p.

Solution :

Given 'X' is a binomial variate with parameters n and p, then the mean of the binomial distribution is np

i.e., $E[X] = np$

To prove : $E\left[\frac{X}{n}\right] = p$

Let $E\left[\frac{X}{n}\right] = \frac{1}{n} E[X] = \frac{1}{n} np = p$

$\therefore \frac{X}{n}$ is an unbiased estimator of the parameter 'p'.

Example 3.

If $\overline{\theta}$ is an unbiased estimate of θ, show that $\overline{\theta}^2$ is a biased estimator of θ^2.

Solution :

Given $\overline{\theta}$ is an unbiased estimate of $\theta \Rightarrow E[\overline{\theta}] = \theta$

To prove : $E[\overline{\theta}^2] \neq \theta^2$

Now $\quad V(\overline{\theta}) = E[\overline{\theta}^2] - [E(\overline{\theta})]^2$

$$\Rightarrow E[\overline{\theta}^2] = V(\overline{\theta})^2 + [E(\overline{\theta})]^2$$

$$= V(\overline{\theta}) + \theta^2 \qquad \because E(\overline{\theta}) = \theta$$

$$\Rightarrow E(\overline{\theta}^2) \neq \theta^2 \qquad \because V(\overline{\theta}^2) > 0$$

$\therefore \overline{\theta}^2$ is a biased estimator of θ^2.

Example 4.

If $x_1, x_2, ..., x_n$ is a random sample from a normal population $N(\mu, 1)$, then show that $t = \frac{1}{n}\sum_{i=1}^{n} x_i$ is an unbiased estimator of $\mu^2 + 1$

Solution :

Given :

$$E(x_i) = \mu \text{ and } V(x_i) = 1$$

$$E(x_i^2) = V(x_i) + [E(x_i)]^2 \qquad \because [V(x_i) = E(x_i^2) - [E(x_i)]^2]$$

$$E(x_i^2) = 1 + \mu^2$$

Let $\quad E(t) = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] = \frac{1}{n}\sum_{i=1}^{n} E(x_i^2)$

$$= \frac{1}{n} \sum_{i=1}^{n} (1 + \mu^2)$$

$$= \frac{1}{n} n (1 + \mu^2) = \mu^2 + 1$$

$$\therefore E(t) = 1 + \mu^2$$

$\therefore$ t is an unbiased estimator of $\mu^2 + 1$

Example 5.

Prove that for a random sample $(x_1, x_2, ..., x_n)$ of size n drawn from a given large population (μ, σ^2), $s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$ is not an unbiased estimator of the parameter σ^2, but $\frac{ns^2}{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ is an unbiased estimator of σ^2.

Solution :

Here, we have assumed that the sample has been drawn from a normal population

Solution :

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} \left[(x_i - \mu) - (\bar{x} - \mu)\right]^2$$

$$= \sum_{i=1}^{n} \left[(x_i - \mu)^2 - 2(x_i - \mu)(\bar{x} - \mu) + (\bar{x} - \mu)^2\right]$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu) \sum_{i=1}^{n} (x_i - \mu) + (\bar{x} - \mu)^2 \sum_{i=1}^{n} 1 \quad \text{... (1)}$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu)\, n\, (\bar{x} - \mu) + (\bar{x} - \mu)^2 (n)$$

$$\left[\because \sum_{i=1}^{n} (x_i - \mu) = n(\bar{x} - \mu) \text{ and } \sum_{i=1}^{n} 1 = n\right]$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 - 2n(\bar{x} - \mu)^2 + n(\bar{x} - \mu)^2$$

i.e., $$\boxed{\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (x_i - \mu)^2 - n(\bar{x} - \mu)^2} \quad \text{... (2)}$$

Given : $s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$

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

$$E[s^2] = \frac{1}{n} \left[\sum_{i=1}^{n} E(x_i - \mu)^2 - n\,E(\bar{x} - \mu)^2 \right] \quad \text{... (3)}$$

We know that, $E(x_i - \mu)^2 = \sigma^2$, $E(\bar{x} - \mu)^2 = \frac{\sigma^2}{n}$

$$(3) \Rightarrow E[s^2] = \frac{1}{n} \left[\sum_{i=1}^{n} \sigma^2 - n\frac{\sigma^2}{n} \right] = \frac{1}{n} \left[\sigma^2 \sum_{i=1}^{n} 1 - \sigma^2 \right]$$

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

$$= \frac{n-1}{n} \sigma^2 < \sigma^2 \quad \text{... (4)}$$

$\therefore s^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$ is not an unbiased estimator of σ^2.

Let us consider $E\left[\frac{n}{n-1} s^2\right] = \frac{n}{n-1} E[s^2]$

$$= \left(\frac{n}{n-1}\right)\left(\frac{n-1}{n}\right)\sigma^2 \text{ by (4)}$$

$$= \sigma^2$$

$\therefore \frac{n}{n-1} s^2$ is an unbiased estimator of σ^2

Note : $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad \Rightarrow \bar{x} - \mu = \frac{1}{n} \sum_{i=1}^{n} x_i - \mu$

$$\Rightarrow \bar{x} - \mu = \frac{1}{n} \sum_{i=1}^{n} x_i - \frac{1}{n} \sum_{i=1}^{n} \mu \qquad \left[\because \sum_{i=1}^{n} \mu = \mu \sum_{i=1}^{n} 1 = \mu(n) = n\mu\right]$$

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

$$\Rightarrow \sum_{i=1}^{n} (x_i - \mu) = n\,\overline{(x - \mu)} \text{ and } \sum_{i=1}^{n} 1 = n$$

Example 6.

Show that $s^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$ **is an unbiased estimator of the parameter** σ^2.

Solution :

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} \left[(x_i - \mu) - (\bar{x} - \mu)\right]^2$$

$$= \sum_{i=1}^{n} \left[(x_i - \mu)^2 - 2(x_i - \mu)(\bar{x} - \mu) + (\bar{x} - \mu)^2\right]$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu) \sum_{i=1}^{n} (x_i - \mu) + (\bar{x} - \mu)^2 \sum_{i=1}^{n} 1 \quad \text{... (1)}$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu)\, n\, (\bar{x} - \mu) + (\bar{x} - \mu)^2 (n)$$

$$\left[\because \sum_{i=1}^{n} (x_i - \mu) = n(\bar{x} - \mu) \text{ and } \sum_{i=1}^{n} 1 = n\right]$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 - 2n(\bar{x} - \mu)^2 + n(\bar{x} - \mu)^2$$

i.e., $$\boxed{\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (x_i - \mu)^2 - n(\bar{x} - \mu)^2} \quad \text{... (2)}$$

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

$$E[s^2] = \frac{1}{n-1} \left[\sum_{i=1}^{n} E(x_i - \mu)^2 - n\,E(\bar{x} - \mu)^2\right] \quad \text{... (3)}$$

We know that $E[(x_i - \mu)^2] = \sigma^2, \quad E[(\bar{x} - \mu)^2] = \frac{\sigma^2}{n}$

$$\therefore (3) \Rightarrow \quad E[s^2] = \frac{1}{n-1} \left[\sum_{i=1}^{n} \sigma^2 - n \frac{\sigma^2}{n}\right]$$

$$= \frac{1}{n-1} \left[\sigma^2 \sum_{i=1}^{n} 1 - \sigma^2\right]$$

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

$$= \frac{1}{n-1}(n-1)\sigma^2 = \sigma^2$$

Hence, $s^2 = \frac{1}{n-1}\Sigma (x_i - \bar{x})^2$ is an unbiased estimator of the parameter σ^2

Note : $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \Rightarrow \bar{x} - \mu = \frac{1}{n}\sum_{i=1}^{n} x_i - \mu$

$$\Rightarrow \bar{x} - \mu = \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n} \mu \qquad [\because \sum_{i=1}^{n} \mu = \mu \sum_{i=1}^{n} 1 = \mu(n) = n\mu]$$

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

$$\Rightarrow \sum_{i=1}^{n} (x_i - \mu) = n(\bar{x} - \mu) \text{ and } \sum_{i=1}^{n} 1 = n$$

Example 7.

If $s_1^2, s_2^2, \ldots, s_r^2$ are r sample variables based on random samples of sizes $n_1, n_2, \ldots, n_r$ respectively drawn from a large population with variance σ^2 and $T = \frac{1}{k}\sum_{i=1}^{r} n_i s_i^2 = \sigma^2$ is an unbiased estimate of σ^2, find the value of k.

Solution : T is an unbiased estimate of σ^2

$$\Rightarrow \quad E(T) = \sigma^2$$

$$E\left[\frac{1}{k}\sum_{i=1}^{r} n_i s_i^2\right] = \sigma^2$$

$$\frac{1}{k}\sum_{i=1}^{r} n_i E[s_i^2] = \sigma^2$$

$$\frac{1}{k}\sum_{i=1}^{r} n_i \left(\frac{n_i - 1}{n_i}\right)\sigma^2 = \sigma^2$$

$$\Rightarrow \sum_{i=1}^{r} (n_i - 1) = k$$

$$\Rightarrow \quad k = \sum_{i=1}^{r} (n_i - 1) = \sum_{i=1}^{r} n_i - \sum_{i=1}^{r} 1$$

$$= (n_1 + n_2 + \ldots + n_r) - r$$

Example 8.

If $x_1, x_2, \ldots, x_n$ is a random sample from $N(\mu, \sigma^2)$ and if $T = \frac{1}{n}\sum_{i=1}^{n} |x_i - \mu|$, then $T\sqrt{\frac{\pi}{2}}$ is an unbiased estimate of σ.

Solution :

$$T = \frac{1}{n}\sum_{i=1}^{n} |x_i - \mu|$$

$$\Rightarrow E(T) = \frac{1}{n}\sum_{i=1}^{n} E\,|x_i - \mu|$$

$E\,|x_i - \mu|$ is the mean deviation about the mean μ of the given normal distribution, which is given by $\sqrt{\frac{2}{\pi}}\,\sigma$

$$\Rightarrow E(T) = \frac{1}{n}.n\sqrt{\frac{2}{\pi}}\,\sigma$$

$$\therefore E\left(\sqrt{\frac{\pi}{2}}\,T\right) \text{ is an unbiased estimator of } \sigma.$$

Example 9.

If $x_1, x_2, \ldots, x_n$ is a random sample from $N(0, \sigma^2)$ and if $T = \frac{1}{n}\sum_{i=1}^{n} x_i^2$, then T is an unbiased estimate of σ^2.

Solution :

$$T = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\Rightarrow E(T) = \frac{1}{n}\sum_{i=1}^{n} E(x_i^2)$$

Since x, follows $N(0, \sigma^2)$, $\text{Var}(x_i) = E(x_i^2) - [E(x_i)]^2 = \sigma^2$

$$\Rightarrow E(x_i^2) = \sigma^2 \text{ since } E(x_i) = 0$$

$$\therefore E(T) = \frac{1}{n}.n\,\sigma^2 = \sigma^2$$

$\Rightarrow T$ is an unbiased estimate of σ^2.

Example 10.

Below you are given the values obtained from a random sample of observations taken from an infinite population

32 34 35 39

(a) Find a point estimator for μ. Is this an unbiased estimate of μ? Explain.

(b) Find a point estimator for σ^2. Is this an unbiased estimate of σ^2. Explain.

(c) Find a point estimator for σ

(d) What can be said about the sampling distribution of $\bar{x}$? Be sure to discuss the expected value, the standard deviation, and the shape of the sampling distribution of $\bar{x}$? **[A.U Jan. 2014 (MBA)]**

Solution :

(a) Point estimator of $\mu = \dfrac{32 + 34 + 35 + 39}{4} = 35$

It is an unbiased estimator of μ.

(b) Point estimator of $\sigma^2 = \dfrac{\Sigma (x - \bar{x})^2}{n - 1}$

$$= \frac{(32 - 35)^2 + (34 - 35)^2 + (35 - 35)^2 + (39 - 35)^2}{4 - 1}$$

$$= \frac{9 + 1 + 0 + 16}{3} = 8.6667$$

It is an unbiased estimator of σ^2.

(c) Point estimator of $\sigma = \sqrt{\sigma^2} = \sqrt{8.6667} = 2.9439$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \qquad \therefore \mu = \bar{x} \pm Z \frac{\sigma}{\sqrt{n}}$$

(d) The sampling distribution has a mean equal to the population mean.

The sampling distribution has a standard deviation equal to the population standard deviation of divided by the square root of the sample size.

The shape of the sampling distribution of $\bar{x}$ will be a normal curve.

Example 11.

A random sample (x_1, x_2, x_3, x_4) of size 4 is drawn from a normal population unknown mean μ. Consider the following estimators to estimate μ.

(i) $t_1 = \dfrac{x_1 + x_2 + x_3 + x_4}{4}$

(ii) $t_2 = \dfrac{x_1 + x_2 + x_3}{3} + x_4$

(iii) $t_3 = \dfrac{x_1 + 2x_2 + \lambda x_3}{3}$ where

λ is such that t_3 is an unbiased estimator. 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.

Solution :

We have $E(x_i) = \mu$, Var $(x_i) = \sigma^2$ ... (1)

Cov $(x_i, x_j) = 0$, $i \neq j = 1, 2, 3, 4, \dots n$

$$E[t_1] = E\left[\frac{x_1 + x_2 + x_3 + x_4}{4}\right]$$

$$= \frac{1}{4}\left[E(x_1) + \dots + E(x_4)\right]$$

$$= \frac{1}{4}[\mu + \mu + \mu + \mu]$$

$$= \frac{1}{4}(4\mu) = \mu$$

$$E[t_2] = E\left[\frac{x_1 + x_2 + x_3}{3} + x_4\right]$$

$$= \frac{1}{3}[\mu + \mu + \mu] + \mu$$

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

$E[t_3] = \mu$ (Given) [$\because$ t_3 is an unbiased estimator]

$$E\left[\frac{x_1 + 2x_2 + \lambda x_3}{3}\right] = \mu$$

$$\Rightarrow \quad \frac{1}{3}[\mu + 2\mu + \lambda\mu] = \mu$$

$$\Rightarrow \quad \frac{1}{3}[3\mu + \lambda\mu] = \mu$$

$$\Rightarrow \quad \frac{1}{3}[3 + \lambda]\mu = \mu$$

$$\Rightarrow \quad \frac{1}{3}[3 + \lambda] = 1$$

$$\Rightarrow \quad 3 + \lambda = 3$$

$$\Rightarrow \quad \lambda = 0$$

$$V(t_1) = \frac{1}{4^2}\left[V(x_1) + V(x_2) + V(x_3) + V(x_4)\right]$$

$$= \frac{1}{16}[\sigma^2 + \sigma^2 + \sigma^2 + \sigma^2]$$

$$= \frac{1}{16}[4\sigma^2] = \frac{1}{4}\sigma^2$$

$$V(t_2) = \frac{1}{3^2}[V(x_1) + V(x_2) + V(x_3)] + V(x_4)$$

$$= \frac{1}{9}[\sigma^2 + \sigma^2 + \sigma^2] + \sigma^2$$

$$= \frac{3}{9}\sigma^2 + \sigma^2 = \frac{1}{3}\sigma^2 + \sigma^2 = \frac{4}{3}\sigma^2$$

$$V(t_3) = \frac{1}{3^2}V(x_1) + \left(\frac{2}{3}\right)^2 V(x_2) + 0$$

$$= \frac{1}{9}\sigma^2 + \frac{4}{9}\sigma^2 = \frac{5}{9}\sigma^2$$

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

Example 12.

Let x_1, x_2, x_3, and x_4 be a random sample of size 4 from a population with man value μ and variance σ^2. Let T_1, T_2, T_3, T_4 be the estimators used to estimate mean value μ where

$$T_1 = x_1 + x_2 + x_3 - 2x_4$$

$$T_2 = x_1 + 2x_2 + 3x_3 - 5x_4$$

$$T_3 = \frac{kx_1 + x_2 + x_3 + x_4}{4}$$

(i) **Find whether T_1 and T_2 are unbiased estimators?**

(ii) **Find the value of k such that T_2 is unbiased estimator for m.**

(iii) **With this values of k is T_3 a consistent estimator?**

(iv) **Which is the best estimator ?**

Solution :

Since x_1, x_2, x_3, x_4 is an random sample from a population with mean m and variance σ^2,

$$E(x_i) = \mu,\ Var(x_i) = \sigma^2,\ Cov(x_i, x_j) = 0,\ (i \neq j = 1, 2, 3, ..., n) \quad ... (1)$$

(i) $E(T_1) = 3\mu - 2\mu = \mu$ which implies that T_1 is a unbiased estimator of μ,

$$E(T_2) = \mu + 2\mu + 3\mu - 5\mu = \mu$$

which implies that T_2 is also an unbiased estimator of μ.

(ii) We are given $E(T_3) = \mu$

$$\Rightarrow \quad \frac{k\mu + \mu + \mu + \mu}{4} = \mu$$

$$\Rightarrow \quad \frac{(3+k)\mu}{4} = \mu$$

$$\frac{3+k}{4} = 1$$

$$\Rightarrow \quad 3 + k = 4$$

$$\Rightarrow \quad k = 1$$

(iii) With $k = 1$, $T_3 = \frac{1}{4}(x_1 + x_2 + x_3 + x_4) = \bar{x}$

Since sample mean is a consistent estimator of population μ, by the Week Law of large numbers, T_3 is a consistent estimator of μ.

(iv) $\text{Var}(T_1) = V(x_1) + V(x_2) + V(x_3) + 4\,Var(x_4) = 3\sigma^2 + 4\sigma^2 = 7\sigma^2$

$$\text{Var}(T_2) = \sigma^2 + 4\sigma^2 + 9\sigma^2 + 25\sigma^2 = 39\sigma^2$$

$$\text{Var}(T_3) = \frac{1}{16}[\sigma^2 + \sigma^2 + \sigma^2 + \sigma^2] = \frac{\sigma^2}{4}$$

Since Var (T_3) is minimum, T_3 is the best estimator of μ in the sense minimum variance.

Example 13.

If $x_1, x_2, \ldots x_n$ are random observations of a Bernoulli's variate x which assumes values 1 and 0 with probabilities θ and $(1-\theta)$ respectively. Show that $\dfrac{T(n-T)}{n(n-1)}$ is an unbiased estimator of $\theta(1-\theta)$, where $T = x_1 + x_2 + \ldots + x_n$ **[A.U A/M 2017 (MBA)]**

Solution :

$P(x=1) = p$ and $P(x=0) = 1-p = q$, replacing θ by the conventional p.

$$\therefore \quad T = (x_1 + x_2 + \ldots + x_n)$$

follows a binomial distribution with mean np and variance npq.

$$\therefore \quad E(T) = np \text{ and } Var(T) = E(T^2) - [E(T)]^2 = npq$$

$$E(T^2) = npq + n^2p^2$$

$$\text{Now} \quad E\left[\frac{T(n-T)}{n(n-1)}\right] = \frac{1}{n(n-1)}\left[nE(T) - E(T^2)\right]$$

$$= \frac{1}{n(n-1)}\left[n^2p - npq - n^2p^2\right]$$

$$= \frac{1}{n(n-1)}\left[n^2pq - npq\right]$$

$$= \frac{1}{n(n-1)}n(n-1)pq = pq = \theta(1-\theta)$$

Hence the result follows.

Example 14.

If $\{x_1, x_2, \ldots, x_n\}$ is a random sample of size n, drawn from a geometric distribution, the probability mass function of which is given by $P(x=r) = pq^{r-1}$; $r = 1, 2, 3, \ldots, \infty$. Prove that the mean of the sample is a consistent estimator of the population mean.

Solution :

Mean and variance of the given geometric population are $\dfrac{1}{p}$ and $\dfrac{q}{p^2}$.

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

$$\text{Var}(\bar{x}) = \text{Var}\left(\frac{1}{n}\sum x_i\right)$$

$$= \frac{1}{n^2} \sum_{n=1}^{n} \text{Var}(x_i), \text{ since } x_1, x_2, ..., x_n \text{ are independent.}$$

$$= \left(\frac{1}{n^2}\right)\left(n \frac{q}{p^2}\right) = \frac{q}{n p^2}$$

$$\text{as } n \to \infty,\ E(\bar{x}) \to \frac{1}{p} \text{ and Var } (\bar{x}) = 0$$

$\therefore$ $\bar{x}$ is a consistent estimator of the population mean $\frac{1}{p}$.

Example 15.

If t_1 is a most efficient estimator and t_2 is an unbiased estimator (of some population parameter) with efficiency e, and if the correlation coefficient ρ between t_1 and t_2 is ρ show that $\rho = \sqrt{e}$.

Solution :

Let V_1 and V_2 be the variance of t_1 and t_2 respectively. Then

$$e = \frac{V_1}{V_2} \Rightarrow V_2 = \frac{V_1}{e} \qquad ... (1)$$

Let $t_3 = p\,t_1 + q\,t_2$, where $p + q = 1$ and let V_3 be the variance of t_3.

$$\text{Then } V_3 = Var\,(pt_1 + qt_2)$$

$$= p^2 V_1 + q^2 V_2 + 2pq \,.\, \text{Cov}\,(t_1, t_2)$$

$$= p^2 V_1 + q^2 . \frac{V_1}{e} + 2pq \,.\, \rho \sqrt{V_1 . \frac{V_1}{e}}$$

$$\rho = \frac{\text{Cov}\,(t_1, t_2)}{\sqrt{V_1 . V_2}}$$

$$\Rightarrow Cov\,(t_1, t_2) = \rho \sqrt{V_1 V_2}$$

$$V_3 = \left(p^2 + \frac{q^2}{e} + \frac{2pq\rho}{\sqrt{e}}\right) V_1 \qquad ... (2)$$

Since V_3 is a less efficient estimator, $V_3 \geq V_1$. i.e.,

$$\left(p^2 + \frac{q^2}{e} + \frac{2pq\rho}{\sqrt{e}}\right) V_1 \geq (p+q)^2 . V_1 \text{ since } P + q = 1 \text{ and from (2)}$$

i.e., $$p^2 + \frac{q^2}{e} + \frac{2pq\rho}{\sqrt{e}} \geq p^2 + q^2 + 2pq$$

i.e., $$q^2\left(\frac{1}{e}-1\right)+2pq\left(\frac{\rho}{\sqrt{e}}-1\right)\geq 0 \quad \cdots (3)$$

In (3), the equality holds good, when $q = 0$ and $p = 1$

$$(3) \Rightarrow 2\left(\frac{\rho}{\sqrt{e}}-1\right) = 0$$

$$\Rightarrow \frac{\rho}{\sqrt{e}}-1 = 0$$

$$\Rightarrow \rho = \sqrt{e}$$

Example 16.

If t_1 and t_2 are both most efficient estimators with equal variance V and if t_3 is the average of t_1 and t_2, prove that Var $(t_3) = \frac{1}{2}V(1+\rho)$, where ρ is the coefficient of correlation between t_1 and t_2.

Solution :

Let $t_3 = \frac{1}{2}(t_1+t_2)$

$$\text{Var}\,(at_1+bt_2) = a^2V(t_1)+b^2V(t_2)+2ab\,\text{Cov}\,(t_1,t_2)$$

$$\therefore \quad \text{Var}\,(t_3) = \text{Var}\left(\frac{1}{2}t_1+\frac{1}{2}t_2\right)$$

$$= \frac{1}{4}V+\frac{1}{4}V+2\times\frac{1}{2}\times\frac{1}{2}\text{Cov}\,(t_1,t_2)$$

$$= \frac{1}{2}V+\frac{1}{2}\rho\sqrt{V(t_1).V(t_2)} = \frac{1}{2}V(1+\rho)$$

Method of Moments

Example 17.

Find the estimator of θ in the population with density function $f(x,\theta) = \theta x^{\theta-1}$; $0 < x < 1$; $\theta > 0$, by method of moments.

Solution :

The first order moment (about the origin) of the population is given by

$$\mu_1' = \int_0^1 x\,.\,\theta x^{\theta-1}\,dx$$

$$= \theta \int_0^1 x\, x^{\theta - 1}\, dx$$

$$= \theta \int_0^1 x^{\theta}\, dx$$

$$= \theta \left(\frac{x^{\theta+1}}{\theta+1}\right)_0^1 = \frac{\theta}{\theta+1}$$

The first order moment of the sample $(x_1, x_2, ..., x_n)$ about the origin is given by

$$m_1{}' = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

By the method of moments,

$$\bar{x} = \frac{\theta}{\theta+1}$$

$$(\theta+1)\bar{x} = \theta$$

$$\theta\bar{x} + \bar{x} = \theta$$

$$\bar{x} = \theta - \theta\bar{x}$$

$$\bar{x} = \theta(1-\bar{x})$$

$$\text{i.e., } \theta = \frac{\bar{x}}{1-\bar{x}}$$

Therefore, $\bar{\theta} = \dfrac{\bar{x}}{1-\bar{x}}$

Example 18.

If $(x_1, x_2, ..., x_n)$ a random sample from the uniform population with the density function $f(x, a, b) = \dfrac{1}{b-a}$; $a < x < b$. Find the estimators of a and b by the method of moments.

Solution :

$$\mu_1{}' = \int_a^b \left(\frac{1}{b-a}\right) x\, dx = \frac{1}{b-a}\int_a^b x\, dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right) = \frac{1}{2}(a+b)$$

$$\mu_2' = \int_a^b \left(\frac{1}{b-a}\right) x^2\, dx = \frac{1}{b-a}\int_a^b x^2\, dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

$$= \frac{1}{b-a}\left(\frac{b^3-a^3}{3}\right) = \frac{1}{3}(a^2+ab+b^2)$$

If $\quad m_1' = \bar{x} = \frac{1}{n}\sum x_i$

and $\quad m_2' = s^2 = \frac{1}{n}\sum x_i^2$

are the first and second order moments of the sample about the origin, then the method of moments gives

$$a+b = 2\bar{x} \quad \Rightarrow b = 2\bar{x} - a \qquad \text{... (1)}$$

and $$a^2+ab+b^2 = 3s^2 \qquad \text{... (2)}$$

(2) $\Rightarrow \quad a^2 + a(2\bar{x}-a) + (2\bar{x}-a)^2 = 3s^2 \qquad$ by (1)

i.e., $\quad a^2 - 2\bar{x}a + (4\bar{x}^2 - 3s^2) = 0$

$$\therefore \quad a = \frac{2\bar{x} \pm \sqrt{4\bar{x}^2 - 4(4\bar{x}^2) - 3s^2)}}{2}$$

$$a = \bar{x} \pm \sqrt{3(s^2 - (\bar{x})^2}$$

Again from (1) $\quad b = \bar{x} \mp \sqrt{3(s^2 - (\bar{x})^2)}$

Since $a < b$, we have $\quad \bar{a} = \bar{x} - \sqrt{3(s^2 - (\bar{x})^2)}$

and $\quad \bar{b} = \bar{x} + \sqrt{3(s^2 - (\bar{x})^2)}$

Example 19.

For the probability mass function $f(x,p) = 3\,C_x \cdot \dfrac{p^x(1-p)^{3-x}}{1-(1-p)^3}$; $x = 1, 2, 3$ obtain the estimator of p by the method of moments, if the frequencies at $x = 1$, 2 and 3 are respectively 22, 20 and 18.

Solution :

$$f(x,p) = \frac{1}{1-(1-p)^3} B(3;p)$$

Therefore, the first order moment about the origin, viz., the mean of the given distribution is given by

$$\mu_1' = \frac{1}{1-(1-p)^3} \cdot 3p$$

The mean of the observed sample is given by

$$\bar{x} = \frac{(1)(22) + (2)(20) + (3)(18)}{22 + 20 + 18}$$

$$= \frac{116}{60} \text{ (or) } \frac{29}{15}$$

By the method of moments, $\mu_1' = \bar{x}$

i.e., $$\frac{3p}{3p - 3p^2 + p^3} = \frac{29}{15}$$

i.e., $$29p^2 - 87p + 42 = 0$$

$$\Rightarrow \quad p = \frac{87 \pm 51.93}{58}$$

$$= 2.395 \text{ (or) } 0.605$$

Since 2.395 is inadmissible, $\bar{p} = 0.605$

Example 20.

Let $(X_1, X_2, ..., X_n)$ be a random sample from the exponential distribution with probability density functions of X as

$$f(x, \theta) = \theta e^{-\theta x}, \; 0 < x < \infty, \; \theta > 0$$
$$= 0, \text{ elsewhere}$$

Estimate θ using the method of moments.

Solution :

Here there is only one parameter θ to estimate so $E[X] = \bar{X}$

For the exponential distribution, $E(X) = \frac{1}{\theta}$

$\therefore E(X) = \bar{X}$ results in $\frac{1}{\theta} = \bar{X}$,

So $\bar{\theta} = \frac{1}{\bar{X}}$ is the moment estimation of θ.

Example 21.

The time to failure of an electronic component follows an exponential distribution with parameter λ. Eight units are randomly selected and tested resulting in the following failure time (in hours) :

13.03, 6.07, 68.44, 17.11, 32.54, 8.77, 12.14, 23.42

Find the moment estimate of λ.

Solution :

$$\text{Here } \bar{x} = \frac{1}{8}(13.03 + 6.07 + 68.44 + 77.11 + 32.54 + 8.77 + 12.14 + 23.42)$$

$$= \frac{1}{8}(181.52) = 22.69$$

$\therefore$ the moment estimate of λ is

$$\lambda = \frac{1}{\bar{x}} = \frac{1}{22.69} = 0.04407$$

Example 22.

Suppose that $X_1, X_2, ..., X_n$ is a random sample from a normal distribution $N(\mu, \sigma^2)$. Find the moment estimators of μ and σ^2.

Solution :

For the normal distribution

$E(X) = \mu$ and $E(X^2) = \mu^2 + \sigma^2$.

Equating $E(X)$ to $\overline{X}$ and $E(X^2)$ to $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ gives

$$\mu = \overline{X}, \ \mu^2 + \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Solving these equations gives the moment estimators

$$\therefore \bar{\mu} = \overline{X} \text{ and}$$

$$\bar{\sigma}^2 = \frac{\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_n\right)^2}{n}$$

$$= \frac{\sum (X_i - \overline{X})^2}{n}$$

Example 23.

A random variable X takes the values 0, 1, 2 with respective probabilities $\frac{1}{2}-\theta$, $\frac{\alpha}{2}+2(1-\alpha)\theta$ and $\left(\frac{1-\alpha}{2}\right)+(2\alpha-1)\theta$, where α and θ are the parameters. If a sample of size 75 drawn from the population yielded the values 0, 1, 2 with respective frequencies 27, 38, 10 respectively, find the estimators of α and θ by the method of moments.

Solution :

$$\mu_1' = E(X) = (0)\left(\frac{1}{2}-\theta\right)+(1)\left\{\frac{\alpha}{2}+2(1-\alpha)\theta\right\}+(2)\left\{\frac{1-\alpha}{2}+(2\alpha-1)\theta\right\}$$

$$= 1-\frac{\alpha}{2}+2\alpha\theta$$

$$\mu_2' = E(X^2) = (1^2)\left\{\frac{\alpha}{2}+2(1-\alpha)\theta\right\}+(2^2)\left\{\frac{1-\alpha}{2}+(2\alpha-1)\theta\right\}$$

$$= 2-\frac{3}{2}\alpha+(6\alpha-2)\theta$$

The mean of the observed sample

$$m_1' = \frac{(38)(1)+(10)(2)}{75} = \frac{58}{75}$$

The second order moment about the origin is given by

$$m_2' = s^2 = \frac{1}{75}(38\times 1^2+10\times 2^2) = \frac{78}{75}$$

By the method of moments, $\mu_1' = \bar{x}$ and $\mu_2' = s^2$

$$1-\frac{\alpha}{2}+2\alpha\theta = \frac{58}{75} \quad \text{... (1)}$$

and

$$2-\frac{3}{2}\alpha+(6\alpha-2)\theta = \frac{78}{75} \quad \text{... (2)}$$

Solving the equations (1) and (2), we get

$$\bar{\alpha} = \frac{34}{33} \text{ and } \bar{\theta} = \frac{7}{50}$$

METHOD OF MAXIMUM LIKELIHOOD ESTIMATOR

Example 24.

Consider a characteristic that occurs in proportion p of a population. Let $X_1, , X_2, ..., X_n$ be a random sample of size n so

$P[X_i = 0] = 1 - p$ and $P[X_i = 1] = p$ for $i = 1, ..., n$

where $0 \le p \le 1$. Obtain the maximum likelihood estimator of p.

(OR)

Find the maximum likelihood estimator for the parameter P of the binomial distribution $B(N, P)$, where N is very large but finite, on the basis of sample of size n. Also find its variance.

Solution :

The probability mass function of the binomial distribution is

$$P(X = x) = p(x; N, P) = N\,C_x\,P^x (1 - P)^{N-x};\ x = 0, 1, 2, ..., N.$$

Therefore, the likelihood function of the random sample $(x_1, x_2, ..., x_n)$ is given by

$$L(x_1, x_2, ..., x_n, P) = \prod_{i=1}^{n} n\,C_{x_i}\,P^{\Sigma x_i} (1 - P)^{n^2 - \Sigma x_i}$$

$$\therefore \quad = \prod_{i=1}^{n} n\,C_{x_i}\,P^{n\bar{x}} (1 - P)^{n^2 - n\bar{x}}$$

$$\log L = \sum_{i=1}^{n} \log(n\,C_{x_i}) + n\bar{x} \log P + n(n - \bar{x}) \log(1 - P)$$

The likelihood equation is

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

i.e.,
$$\frac{n\bar{x}}{P} - \frac{n(n - \bar{x})}{1 - P} = 0$$

i.e.,
$$(1 - P)\bar{x} - P(n - \bar{x}) = 0$$

$$\Rightarrow \quad \bar{x} - P\bar{x} - Pn + P\bar{x} = 0$$

$$\bar{x} - nP = 0 \Rightarrow P = \frac{\bar{x}}{n}$$

By Rao-Cramer's formula,

$$\left\{Var\,(\bar{P})\right\}^{-1} = -E\left[\frac{\partial^2}{\partial P^2}\log L\right]$$

$$= -E\left[\frac{\partial}{\partial P}\left\{\frac{n\,\bar{x}}{P} - \frac{n\,(n-\bar{x})}{1-P}\right\}\right]$$

$$= -E\left[-\frac{n\,\bar{x}}{P^2} - \frac{n\,(n-\bar{x})}{(1-P)^2}\right]$$

$$= n\left[\frac{1}{P^2}E\,(\bar{x}) + \frac{1}{(1-P)^2}\left[(n - E\,(\bar{x}))\right]\right]$$

$$= n\left[\frac{1}{P^2}\,n\,P + \frac{1}{(1-P)^2}\,(n-nP)\right]$$

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

$$= \frac{n^2}{PQ}$$

$$Var\,(\bar{P}) = \frac{PQ}{n^2}$$

Note : For the binomial distribution

$$E\,(\bar{x}) = E\left\{\frac{1}{n}\sum x\right\} = np]$$

Example 25.

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the Poisson distribution

$$f\,(x \mid \lambda) = \frac{\lambda^x\,e^{-\lambda}}{x\,!}$$

where $0 \le \lambda < \infty$. Obtain the maximum likelihood estimator of λ.

Solution :

The p.d.f is $P\,[X = x] = \dfrac{\lambda^x\,e^{-\lambda}}{x\,!}$

$$L\,[\lambda : x_1, x_2, \ldots, x_n] = \prod_{i=1}^{n}\frac{\lambda^{x_i}\,e^{-\lambda}}{x_i\,!}$$

$$\log L = \sum_{i=1}^{n} \log \left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i !} \right)$$

$$= \sum_{i=1}^{n} \left[\log (\lambda^{x_i} e^{-\lambda}) - \log (x_i !) \right]$$

$$= \sum_{i=1}^{n} \left[\log (\lambda^{x_i} + \log (e^{-\lambda}) - \log (x_i !) \right]$$

$$= \sum_{i=1}^{n} \left[x_i \log \lambda - \lambda - \log (x_i !) \right]$$

$$= \log \lambda \sum_{i=1}^{n} x_i - n \lambda - \sum_{i=1}^{n} \log (x_i !)$$

$$\frac{d}{d\lambda} (\log 2) = \text{(The likelihood equation)}$$

$$\frac{1}{\lambda} \sum_{i=1}^{n} x_i - n = 0$$

$$\frac{1}{\lambda} \sum_{i=1}^{n} x_i = n$$

$$\frac{1}{n} \sum_{i=1}^{n} x_i = \lambda$$

$$\lambda = \frac{1}{n} \sum_{j=1}^{n} x_i = \bar{x}$$

Example 26.

The number of defective hard drives produced daily by a production line can be modeled as a Poisson distribution. The counts for ten days are

7 3 1 2 4 1 2 3 1 2

Obtain the maximum likelihood estimate of the probability of 0 to 1 defectives on one day.

Solution :

We know that, the maximum likelihood estimate of λ is

$$\bar{\lambda} = \bar{x} = \frac{7+3+1+2+4+1+2+3+1+2}{10} = \frac{26}{10} = 2.6.$$

Consequently, by the invariance property, the maximum likelihood estimate of

$$P(X = 0 \text{ or } 1) = e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!}$$

is

$$e^{-\bar{\lambda}} + \frac{\bar{\lambda} e^{-\bar{\lambda}}}{1!} = e^{-2.6} + \frac{2.6 \,.\, e^{-2.6}}{1!} = 0.267$$

There will 1 or fewer defectives on just over one-quarter of the days.

Example 27.

In one area along the interstate, the number of dropped wireless phone connections per call follows a Poisson distribution. From four calls, the number of dropped connections is

$$2 \quad 0 \quad 3 \quad 1$$

Find the maximum likelihood estimate of λ.

Solution :

We know that, the maximum

likelihood estimate of $\lambda = \bar{\lambda} = \bar{x} = \frac{2+0+3+1}{4} = \frac{6}{4} = \frac{3}{2} = 1.5$

Example 28.

For random sampling from a 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.

Solution :

For the normal distribution $N(\mu, \sigma^2)$, the likelihood function is given by

$$L = f(x_1, \theta) \,.\, f(x_2, \theta) \,\ldots\, f(x_n, \theta)$$

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

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

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

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

The MLE of μ is given by

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

i.e., $\bar{x}-\mu = 0$

$\therefore \bar{\mu}$ = MLE of $\mu = \bar{x}$, since it is seen that

$$\frac{\partial^2}{\partial\mu^2}(\log L) = \frac{-n}{\sigma^2} < 0$$

(ii) Differentiating both sides of (1) w.r.t. σ^2, we get

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

The MLE of σ^2 is given by

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

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

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

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

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

i.e., $\bar{\mu} = \bar{x}$ ... (2)

and $\bar{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ ... (3)

Using (2) in (3), we get $(\bar{\sigma})^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{\mu})^2$

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

Note : $E(\bar{\mu}) = E(\bar{x}) = \mu$ and $E(\bar{\sigma})^2 = E(s^2) \neq \sigma^2$. This means that the MLEs need not necessarily be unbiased.

Example 29.

One process of making green gasoline takes biomass in the form of sucrose and converts it into gasoline using catalytic reactions. At one step in a pilot plant process, the output includes carbon chains of length 3. Fifteen runs with same catalyst produced the yields (gal).

5.57	5.76	4.18	4.64	7.02	6.62	6.33	7.24
5.57	7.89	4.67	7.24	6.43	5.59	5.39	

Treating the yields as a random sample from a normal population.

(a) Obtain the maximum likelihood estimates of the mean yield and the variance.

(b) Obtain the maximum likelihood estimate of the coefficient of variation σ/μ

ion :

(a) We calculate

$$\bar{\mu} = \bar{x} = \frac{5.57 + 5.76 + \ldots + 5.39}{15} = \frac{90.14}{15} = 6.009 \text{ gal}$$

Recall that the maximum likelihood estimate of variance uses divisor n, not $n - 1$.

$$(\bar{\sigma})^2 = \frac{1}{n} \sum_{i=1}^{15} (x_i - \bar{x})^2 = \frac{1}{15}(16.263) = 1.084$$

(b) The coefficient of variation is a function of μ and σ^2, so its maximum likelihood estimate is that same function of $\bar{\mu}$ and $\bar{\sigma}^2$.

$$\left(\frac{\bar{\sigma}}{\mu}\right) = \frac{\bar{\sigma}}{\bar{\mu}} = \frac{\sqrt{1.084}}{6.009} = 0.173$$

Example 30.

Obtain the maximum likelihood estimators of a and b in terms of the sample observations $x_1, x_2, \ldots, x_n$ taken from the exponential population with density function

$$f(x, a, b) = k\, e^{-b(x-a)} \; ; \; x \geq a, \; b > 0$$

Solution :

To find the unknown constant k, we have

$$\int_a^{\infty} f(x, a, b)\, dx = 1$$

i.e.,
$$k \int_a^{\infty} e^{-b(x-a)}\, dx = 1$$

i.e.,
$$k\, e^{ab} \left(\frac{e^{-bx}}{-b}\right)_a^{\alpha} = 1$$

i.e.,
$$\frac{k}{b} = 1 \text{ or } k = b$$

$$\therefore f(x, a, b) = be^{-b(x-a)}; \; x \geq a, \; b > 0$$

Now
$$L(x_1, x_2, \ldots, x_n, a, b) = b^n \cdot e^{-b \sum_{i=1}^{n} (x_i - a)}$$

$$= b^n\, e^{-b(n\bar{x} - na)} \qquad \ldots (1)$$

The likelihood equations for the simultaneous estimation of a and b are

$$\frac{\partial}{\partial a}(\log L) = 0 \qquad \ldots (2)$$

and
$$\frac{\partial}{\partial b}(\log L) = 0 \qquad \ldots (3)$$

From (1), we get

$$\log L = n \log b - n\, b\, (\bar{x} - a)$$

$\therefore$ (2) $\Rightarrow nb = 0$, which means $b = 0$, which is absard as b is given to be positive.

$$(3) \Rightarrow \frac{n}{b} - n\,(\bar{x} - a) = 0$$

$$\therefore\ b = \frac{1}{\bar{x} - a}$$

viz., b is not unique as a is not definitely known.

Again from (1), we note that L is maximum, for a given positive value of b, when $e^{-n\,b\,(\bar{x} - a)}$ is maximum

i.e., when $(\bar{x} - a)$ is minimum.

i.e., when $\frac{1}{n}\left[(x_1 - a) + (x_2 - a) + \ldots + (x_n - a)\right]$ is minimum.

i.e., when a = smallest among $x_1, x_2, \ldots, x_n$, say x_s.

For this value of

$$\bar{a} = x_s\ ;\ \bar{b} = \frac{1}{\bar{x} - x_s}$$

Example 31.

Find the maximum likelihood estimator of θ in the population with density function $f(x, \theta) = (1 + \theta)x^{\theta}$; $0 \le x \le 1$; $\theta > 0$; based on a random sample of size n. Test whether this estimator is a sufficient estimator of θ.

Solution :

$$L(x_1, x_2, \ldots, x_n, \theta) = f(x, \theta) \ldots f(x_n, \theta)$$

$$= (1 + \theta)^n \cdot (x_1, x_2, \ldots, x_n)^{\theta}$$

$$\therefore\ \log L = n \log(1 + \theta) + \theta \sum_{i=1}^{n} \log x_i$$

The likelihood equation is

$$\frac{\partial}{\partial \theta}(\log L) = 0$$

$$\Rightarrow \quad \frac{n}{1 + \theta} + \sum_{i=1}^{n} \log x_i = 0$$

$$\Rightarrow \quad \frac{n}{1 + \theta} + n \log G = 0$$

where G is the geometric mean of the sample.

$$\therefore\ \bar{\theta} = -\left\{1 + \frac{1}{\log G}\right\}$$

Now $L(x_1, x_2, ..., x_n, \theta) \left\{(1+\theta)^n \prod_{i=1}^{n} x_i^{\theta-1}\right\} \left\{\prod_{i=1}^{n} x_i\right\}$

$$= \left((1+\theta)^n . G^{n(\theta-1)}\right) \left\{\prod_{i=1}^{n} x_i\right\}$$

$$= L_1(G, \theta) . L_2(x_1, x_2, ..., x_n)$$

$\therefore \bar{\theta}$ which is given by $-\left\{1 + \frac{1}{\log G}\right\}$ is a sufficient estimator of θ.

3.2 INTERVAL ESTIMATION

In an estimate of a population parameter θ given by two distinct numbers between which the parameter may be considered to lie, then the estimate is called an interval estimate of the parameter θ.

The interval estimate or a confidence interval consists of an upper confidence limit and lower confidence limit and we assign a probability (usually stated as $\alpha\%$ of 95% or 99% etc.) that this interval contains the unknown population parameter.

Table values of Z_α

Confidence level $(1-\alpha)$	Z_α (or) t_α
90%	1.64
95%	1.96
98%	2.33
99%	2.575
Without any reference to the confidence level	3.00

Interval estimates of the population parameter

1. **Mean μ :**

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

i.e., $\mu \in \left(\bar{x} \pm z_{\alpha/2} S.E(\bar{x})\right)$ (or) $\mu \in \left(\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$

If 95% confidence then the population parameter lies between

$$\mu \in \left(\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

If 99% confidence then the population parameter lies between

$$\mu \in \left(\bar{x} \pm 2.58 \frac{\sigma}{\sqrt{n}}\right)$$

2. **Proportion P :**

$$P \in \left[p \pm z_{\alpha/2} S.E\,(p)\right]$$

3. **Difference of means $(\mu_1 - \mu_2)$:**

$$(\mu_1 - \mu_2) \in [\,(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} S.E\,(\bar{x}_1 - \bar{x}_2)]$$

4. **Difference of proportions $(P_1 - P_2)$:**

$$(P_1 - P_2) \in [\,(p_1 - p_2) \pm z_{\alpha/2} S.E\,(p_1 - p_2)]$$

Maximum Error of Estimate E

3.2(a) Maximum Error of Estimate (For large samples)

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \qquad (\sigma \text{ known})$$

Example 1.

A random sample of size 100 has a standard deviation of 5. What can you say about the maximum error with 95% confidence?

Solution :

Given : $n = 100$; $\sigma = 5$

The maximum error $= z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ where $z_{\alpha/2}$ at 95% $= 1.96$

$\therefore$ Maximum Error $= (1.96)\left(\frac{5}{\sqrt{100}}\right) = (1.96)\left(\frac{1}{2}\right) = 0.98$

Example 2.

An industrial engineer intends to use the mean of a random sample of size $n = 150$ to estimate the average mechanical aptitude (as measured by a certain test) of assembly line workers in a large industry. If, on the basis of experience, the engineer can assume that $\sigma = 6.2$ for such data, what can he assert with probability 0.99 about the maximum size of his error?

Solution :

Given : $n = 150, \sigma = 6.2$

The maximum error $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$, $Z_{\alpha/2}$ at 99% $= 2.575$

$$\Rightarrow \quad E = 2.575 \cdot \frac{6.2}{\sqrt{150}}$$

$$= 1.30$$

Hence, the engineer can assert with probability 0.99 that his error will be at most 1.30

Example 3.

Assuming that $\sigma = 20.0$, how large a random sample be taken to assert with probability 0.95 that the sample mean will not differ from the true mean by more than 3.0 points ?

Solution :

Given : $\sigma = 20, \quad z_{\alpha/2} = 1.96 \qquad E = 3, \qquad$ To find n

We know that,

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \quad 3 = (1.96) \frac{20}{\sqrt{n}}$$

$$\sqrt{n} = \frac{(1.96)\ (20)}{3}$$

i.e., $\sqrt{n} = 13.067$

$$\therefore n = 170.729 \approx 171$$

Example 4.

If the standard deviation of a sample is 20 and the maximum error with 99% confidence is 1.72, how large a sample might be ?

Solution :

Given : $s = 20$; $E = 1.72$; $z_{\alpha/2} = 2.58$

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

(i.e.,) $$1.72 = 2.58 \times \frac{20}{\sqrt{n}}$$

$$\Rightarrow \quad \sqrt{n} = \frac{2.58 \times 20}{1.72} = 30$$

$$\therefore \quad n = 900$$

The required sample size is 900.

Example 5.

What is the maximum error one can expect to make with probability 0.9, when using the mean of a random sample of size $n = 64$ to estimate the mean of a population with $\sigma^2 = 2.567$?

Solution :

Given : $n = 64$

Maximum error $$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$z_{\alpha/2} = 1.645 \text{ (90\% confidence)}$$

Standard deviation $$\sigma = \sqrt{2.56} = 1.6$$

$$\therefore E = (1.645)\left(\frac{1.6}{8}\right) = 0.329$$

$\therefore$ Maximum error $= 0.329$

Example 6.

A research worker wants to determine the average time it takes a mechanic to rotate the tyres of a car and he wants to be able to assert with 95% confidence that the mean of his sample is of by atmost 0.5 minutes. If he can presume from past experience that $\sigma = 1.6$ minutes, how large a sample will have to take ?

Solution :

Given : $E = 0.5$, $\sigma = 1.6$ minutes, $z_{\alpha/2} = 1.96$

n = sample size = ?

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$0.5 = (1.96) \frac{1.6}{\sqrt{n}}$$

$$0.5 = \frac{3.136}{\sqrt{n}}$$

$$\sqrt{n} = \frac{3.136}{0.5} = 6.272$$

$$n = 39.33 \approx 39$$

$\therefore$ The required size of the sample is 39.

Example 7.

The dean of a college wants to use the mean of a random sample to estimate the average amount of time students take to get from one class to the next and she wants to be able to assert with 99% confidence that the error is atmost 0.25 minute. If it can be presumed from experience that σ = 1.4 min, how large a sample will she have to take ?

Solution :

Given : $E = 0.25$, $\sigma = 1.4$, $z_{\alpha/2} = 2.575$

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$0.25 = (2.575) \left(\frac{1.4}{\sqrt{n}}\right)$$

$$0.25 = \frac{3.605}{\sqrt{n}}$$

$$\sqrt{n} = \frac{3.605}{0.25} = 14.42$$

$$\Rightarrow n = 207.93 \approx 208$$

$\therefore$ The required sample size = 208

Example 8.

It is desired to estimate the mean number of hours of continuous use until a certain computer will first require repairs. If it can be assumed that $\sigma = 48$ hours how large a sample be needed so that one will be able to assert with 90% confidence that the sample mean is off by at most 10 hours.

Solution :

Given : $E = 10, \quad \sigma = 48, \quad z_{\alpha/2} = 1.645$

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$10 = (1.645)\left(\frac{48}{\sqrt{n}}\right)$$

$$10 = \frac{78.96}{\sqrt{n}}$$

$$\sqrt{n} = \frac{78.96}{10} = 7.896$$

$$\Rightarrow n = 62.34 \approx 62$$

$\therefore$ The required sample size = 62.

3.2.(b) Maximum Error of estimate (For small samples)

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}} \quad (\sigma \text{ unknown})$$

Example 9.

In six determinations of the melting point of an aluminium alloy, a chemist obtained a mean of 532.26 degrees Celsius with a standard deviation of 1.14 degree. If he uses this mean to estimate the actual melting point of the alloy, what can the chemist assert with 98% confidence about the maximum error?

Solution :

Given $n = 6$, $s = 1.14$ and $t_{0.01} = 3.365$ (for $n - 1 = 5$ degrees of freedom)

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$E = 3.365 \cdot \frac{1.14}{\sqrt{6}} = 4.24$$

Hence, the chemist can assert with 98% confidence that his figure for the melting point of the aluminium alloy is off by at most 4.24 degrees.

3.2(c) Confidence interval for the population mean for large samples (σ is known) is

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \qquad \text{(or)}$$

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Example 10.

A random sample of size $n = 100$ is taken from a population with $\sigma = 5.1$. Given that the sample mean is $\bar{x} = 21.6$, construct a 95% confidence interval for the population mean μ.

Solution :

Given : $n = 100, \bar{x} = 21.6, \sigma = 5.1, z_{\alpha/2} = 1.96$

$$\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$21.6 - 1.96 \cdot \frac{5.1}{\sqrt{100}} < \mu < 21.6 + 1.96 \cdot \frac{5.1}{\sqrt{100}}$$

(or) $20.6 < \mu < 22.6$. Of course, either the interval from 20.6 to 22.6 contains the population mean μ, or it does not, but we are 95% confident means that the method by which the interval was obtained "works" 95% of the time.

Example 11.

The average monthly electricity consumption for a sample of 100 families is 1250 units. Assuming the standard deviation of electric consumption of all families is 150 units, construct a 95% confidence interval estimate of the actual mean electric consumption.

Solution :

Given : $\bar{x} = 1250, \quad \sigma = 150, \quad n = 100, \quad z_{\alpha/2} = 1.96$

$$\bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1250 \pm 1.96 \frac{150}{\sqrt{100}} = 1250 \pm 29.40 \text{ units.}$$

Thus, for 95% level of confidence, the population mean μ is likely to fall between 1220.60 units and 1274.40 units, i.e., $1220.60 \le \mu \le 1274.40$

Example 12.

In order to introduce some incentive for higher balance in savings accounts, a random sample size of 64 savings accounts at a bank's branch was studied to estimate the average monthly balance in saving bank accounts. The mean and standard deviation were found to be Rs. 8,500 and Rs. 2,000, respectively. Find

(1) 90%

(2) 95% and

(3) 99% confidence intervals for the population mean.

Solution :

Confidence limits with confidence level $(100 - \alpha)$ % for average monthly balance in savings accounts are given as :

(1) 90% confidence limits	(2) 95% confidence limits	(3) 99% confidence limits
$\bar{x} \pm 1.645 \dfrac{\sigma}{\sqrt{n}}$	$\bar{x} \pm 1.96 \dfrac{\sigma}{\sqrt{n}}$	$\bar{x} \pm 2.575 \dfrac{\sigma}{\sqrt{n}}$
$8500 \pm 1.645 \dfrac{2000}{\sqrt{64}}$	$8500 \pm 1.96 \dfrac{2000}{\sqrt{64}}$	$8500 \pm 2.575 \dfrac{2000}{\sqrt{64}}$
$\Rightarrow \quad 8500 \pm \dfrac{3290}{8}$	$\Rightarrow \quad 8,500 \pm \dfrac{3,920}{8}$	$\Rightarrow \quad 8500 \pm \dfrac{5150}{8}$
$\Rightarrow \quad 8500 \pm 411.25$	$\Rightarrow \quad 8500 \pm 490$	$\Rightarrow \quad 8500 \pm 644$

It may be noted that the interval or limits gets wider as the desired level of confidence is increased.

3.2(d) Confidence interval for the population mean for large samples (when σ is unknown) is

$$\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Example 17.

With reference to the nanopillar height data $n = 50$, $\bar{x} = 305.58$ nm, and $s^2 = 1{,}366.86$ (hence, $s = 36.97$ nm), construct a 99% confidence interval for the population mean of all nanopillars.

Solution :

Given : $n = 50$, $\bar{x} = 305.58$, $s = 36.97$, and $z_{0.005} = 2.575$, we get

$$\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$\Rightarrow 305.58 - 2.575 \,.\, \frac{36.97}{\sqrt{50}} < \mu < 305.58 + 2.575 \,.\, \frac{36.97}{\sqrt{50}}$$

(or) $292.12 < \mu < 319.04$. We are 99% confident that the interval from 292.12 nm to 319.04 nm contains the true mean nanopillar height.

3.2(e) Confidence interval for the population mean for small samples (σ is unknown)

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\text{(or)} \quad \bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Example 18.

We know that silk fibers are very tough but in short supply. Engineers are making breakthroughs to create synthetic silk fibers that can improve everything from car bumpers to bullet-proof vests or to make artificial blood vessels. One research group reports the summary statistics

$$n = 18, \ \bar{x} = 22.6 \ , \ s = 15.7$$

for the toughness (MJ/m^3) of processed fibers.

Construct a 95% confidence interval for the mean toughness of these fibers. Assume that the population is normal.

Solution :

Given : $n = 18$ and $t_{0.025} = 2.110$ for $n - 1 = 17$ degrees of freedom.

The 95% confidence formula for μ becomes

$$\bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$\Rightarrow \quad 22.6 - 2.110 \frac{15.7}{\sqrt{18}} < \mu < 22.6 + 2.110 \,.\, \frac{15.7}{\sqrt{18}} \quad \text{(or)} \ 14.79 < \mu < 30.41 \ MJ/m^3$$

We are 95% confidence that the interval from 14.79 to 36.41 MJ/m^3 contains the mean toughness of all possible artificial fibers created by the current process.

The article does not give the original data but, since $n = 18$ is moderately large, the normal assumption is not critical unless an outlier exists.

Example 19.

A sample of 10 cam shafts intended for use in gasoline engines have average eccentricity of 1.02 and a standard deviation of 0.444. Assuming the data may be treated a random sample from a normal population, determine a 95% confidence interval for the actual mean eccentricity of the cam shaft?

Solution :

Given : $n = 10, \quad s = 0.044, \quad \bar{x} = 1.02$

$$t_{\alpha/2} = 2.26 \text{ for 9 d.f} \qquad \therefore n \text{ is small.}$$

$$\therefore E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = (2.26)\left(\frac{0.044}{\sqrt{10}}\right)$$

$$= 0.317$$

$\therefore$ 95% confidence interval for the actual mean eccentricity is

$$\bar{x} \pm t_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 1.02 \pm 0.317$$

$$= (0.7027, 1.337)$$

Example 20.

Ten ball bearings made by a certain process have a mean diameter of 0.5060 cm. with a standard deviation of 0.0040 cm. Assuming that the data may be taken as a random sample from a normal population, construct a 95% confidence interval for the actual average diameter of the ball bearings ?

Solution :

Given : $n = 10; \quad \bar{x} = 0.5060; \quad s = 0.0040.$ $\therefore$ This is a small sample.

The maximum error of estimate for 95% confidence is

$$\boxed{E_{max} = t_{\alpha/2} \frac{s}{\sqrt{n}}}$$

Given : $t_{\alpha/2}$ at 95% with $(n-1)$ d.f.

i.e., $\quad t_{0.025}, 9 \text{ df} = 2.262$

$$\Rightarrow \quad E_{max} = \frac{2.262 \times 0.0040}{\sqrt{10}}$$

$$= 0.00286$$

Now the 95% confidence interval limits are

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 0.5060 \pm 0.00286$$

$$= (0.5031,\ 0.5088)$$

3.2.(f) Confidence interval for the difference between two population means for large samples (σ is known)

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Example 21.

In a certain factory there are two independent processes manufacturing the same item. The average weight in a sample of 250 items produced from one process is found to be 120 O_{zs}, with a standard deviation of 12 O_{zs}. While the corresponding figures in a sample of 400 items from the other process are 124 O_{zs} and 14 O_{zs}. Find the 99% confidence limits for the difference in the average weight of items produced by the two processes respectively.

Solution :

Given : $n_1 = 250,\ \bar{x}_1 = 120,\ S_1 = 12$

Given : $n_2 = 400,\ \bar{x}_2 = 124,\ S_1 = 14$

$$Z_{\alpha/2} = 2.58$$

$\therefore$ 99% confidence limits for $(\mu_1 - \mu_2)$ are

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = (\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$= (120 - 124) \pm 2.58 \sqrt{\frac{12^2}{250} + \frac{14^2}{400}}$$

$$= 4 \pm (2.58)\,(1.0324)$$

$$= 4 \pm 2.6635$$

$$(\mu_1 - \mu_2) = (1.34,\ 6.66)$$

3.2(g) Confidence interval for the difference between two population means for small samples.

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\frac{\sigma_1^2}{x_1} + \frac{\sigma_2^2}{x_2}}$$

$$\text{(or)}\ (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\, S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where $t_{\alpha/2}$ is the tabulated value of t with $n_1 + n_2 - 2$ degrees of freedom at α level of significance.

Example 22.

In a test given to two groups of students the marks obtained were as follows:

First group :	18	20	36	50	49	36	34	49	61
Second group :	29	28	26	35	30	44	46		

Construct a 95% confidence interval on the mean marks secured by students of the above two groups.

Solution :

Given : $n_1 = 9, \quad n_2 = 7, \quad \bar{x}_1 = \frac{333}{9} = 37\ ; \quad \bar{x}_2 = \frac{238}{7} = 34$

$$S^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum_i (x_i - \bar{x}_1)^2 + \sum_j (x_j - \bar{x}_2)^2\right]$$

$$= \frac{1}{14}[1134 + 386] = 108.57$$

$\therefore S = 10.42$ and $t_{\alpha/2}$ with $n_1 + n_2 - 2 = 14$ d.f is 1.76

$\therefore$ the 95% confidence interval is

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\, S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= (37 - 34) \pm (1.76)(10.42)\sqrt{\tfrac{1}{9} + \tfrac{1}{7}}$$

$$= 3 \pm (1.76)(5.25) = 3 \pm 9.24$$

$$= (6.24,\ 12.24)$$

Example 23.

Construct a 94% confidence interval for the difference between the mean lifetimes of two kinds of light bulbs, given that a random sample of 40 light bulbs of the first kind lasted on the average 418 hours of continuous use and 50 light bulbs of the second kind lasted on the average 402 hours of continuous use. The population standard deviations are known to be $\sigma_1 = 26$ and $\sigma_2 = 22$.

Solution :

For $\alpha = 0.06$, we find from table, $z_{0.03} = 1.88$. Therefore, the 94% confidence interval for $\mu_1 - \mu_2$ is

$$(418-402)-1.88 \cdot \sqrt{\frac{26^2}{40}+\frac{22^2}{50}} < \mu_1 - \mu_2 < (418-402)+1.88 \cdot \sqrt{\frac{26^2}{40}+\frac{22^2}{50}}$$

which reduces to

$$6.3 < \mu_1 - \mu_2 < 25.7$$

Hence, we are 94% confident that the interval from 6.3 to 25.7 hours contains the actual difference between the mean lifetimes of the two kinds of light bulbs. The fact that both confidence limits are positive suggests that on the average the first kind of light bulbs is superior to the second kind.

3.2(h) Interval estimation of variance

If s^2 is the value of the variance of a random sample of size n from a normal population, then

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,\, n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\, n-1}}$$

is a $(1-\alpha)\,100\%$ confidence interval for σ^2.

Example 24.

In 16 test runs the gasoline consumption of an experimental engine had a standard deviation of 2.2 gallons. Construct a 99% confidence interval for σ^2, which measures the true variability of the gasoline consumption of the engine.

Solution :

Given : $n = 16$, $s = 2.2$, $\chi^2_{0.005,15} = 32.801$ and $\chi^2_{0.995,15} = 4.601$.

$$\frac{(n-1)\,s^2}{\chi^2_{\alpha/2,\,n-1}} < \sigma^2 < \frac{(n-1)\,s^2}{\chi^2_{1-\alpha/2,\,n-1}}$$

$$\frac{15\,(2.2)^2}{32.801} < \sigma^2 < \frac{15\,(2.2)^2}{4.601}$$

(i.e.,) $$2.21 < \sigma^2 < 15.78$$

3.2(i) Interval estimation of ratio of two variance

If s_1^2 and s_2^2 are the values of the variances of independent random samples of sizes n_1 and n_2 from normal population, then

$$\left[\frac{s_1^2}{s_2^2}\,\frac{1}{f_{\alpha/2}\,n_1-1,\,n_2-1} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\,f_{\alpha/2}\,n_2-1,\,n_1-1\right]$$

is a $(1-\alpha)\,100\%$ confidence interval for $\dfrac{\sigma_1^2}{\sigma_2^2}$

25.

A study has been made to compare the nicotine contents of two brands of cigarettes. Ten cigarettes of Brand A and an average nicotine content of 3.1 milligrams with a standard deviation of 0.5 milligram, while eight cigarettes of Brands B had an average nicotine content of 2.7 milligrams with a standard deviation of 0.7 milligram. Assuming that the two sets of data are independent random samples from normal populations with equal variances, construct a 95% confidence interval for the difference between the mean nicotine contents of the two brands of cigarettes. Find a 98% confidence interval for $\dfrac{\sigma_1^2}{\sigma_2^2}$

Solution :

Given : $n_1 = 10$, $n_2 = 8$, $s_1 = 0.5$, $s_2 = 0.7$

$f_{0.01,9,7} = 6.72$, $f_{0.01,7,9} = 5.61$

$$\left[\frac{s_1^2}{s_2^2}\frac{1}{f_{\alpha/2}n_1 - 1, n_2 - 1} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\alpha/2} n_2 - 1, n_1 - 1\right]$$

$$\Rightarrow \qquad \frac{0.25}{0.49}\,\frac{1}{6.72} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{0.25}{0.49}\,5.61$$

$$\Rightarrow \qquad 0.076 < \frac{\sigma_1^2}{\sigma_2^2} < 2.862$$

Since the interval obtained here includes the possibility that the ratio is 1, there is no real evidence against the assumption of equal population variance.

MISCELLANEOUS PROBLEMS

Example 26.

A random sample of 50 sales invoices was taken from a large population of sales invoices. The average value was found to be Rs. 2,000 with a standard deviation of Rs. 540. Find a 90% confidence interval for the true mean value of all the sales.

Solution :

The information given is $\overline{X} = 2{,}000$, $s = 540$, $n = 64$, and $\alpha = 10\%$

Therefore,

$$s_{\overline{x}} = \frac{s}{\sqrt{n}} = \frac{540}{\sqrt{64}} = 67.50 \text{ and } z_{\alpha/2} = 1.64 \text{ (from normal table)}$$

The required confidence interval of population mean μ is given by :

$$\overline{X} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} = 2000 \pm 1.64\,(67.50) = 2000 \pm 110.70$$

Hence, the mean of the sales invoices for the whole population is likely to fall between Rs. 1,889.30 and Rs. 2,110.70 i.e., $1{,}889.30 \le \mu < 2110.70$

Example 27.

A random sample of 100 observations yields sample mean $\overline{X} = 150$ and sample variance $s^2 = 400$. Compute 95% and 99% confidence interval for the population mean.

Solution :

Given : $n = 100$, $\overline{X} = 150$, $S^2 = 400 \Rightarrow S = 20$

$$S.E_{\overline{X}} = \frac{S}{\sqrt{n}} = \frac{20}{\sqrt{100}} = 2 \text{ [For large sample, } \sigma = S]$$

At 95% confidence level, the value of $z_{\alpha/2} = 1.96$

At 95% confidence level, the value of $z_{\alpha/2} = 2.58$

(1) 95% confidence interval or limits for μ are :

$$\overline{X} \pm 1.96\, S.E_{\overline{X}}$$

Putting the values, we get,

$$150 \pm 1.96 \times 2 = 150 \pm 3.92 = 153.92 \text{ or } 146.08$$

Hence, $146.08 < \mu < 153.92$

(2) 99% confidence interval or limits for μ are :

$$\overline{X} \pm 258 .S.E_{\overline{X}}$$

$$= 150 \pm 2.58 \times 2$$

$$= 150 \pm 5.16$$

$$= 155.16 \text{ (or) } 144.84$$

Hence, $144.84 < \mu < 155.16$

EXERCISE

1. If t is an unbiased estimator θ, show that t^2 is a biased estimator of θ^2, but if t is a consistent estimator of θ, then t^2 is also a consistent estimator of θ^2.
2. If t_n is a sample statistic that $E(t_n) = \theta_n$ and Var $(t_n) = \delta_n$ and if $\theta_n \to \theta$ and $\delta_n \to 0$ as $n \to \infty$. Then t_n is a consistent estimator of θ.

3. If t is the most efficient estimator and t' is an unbiased estimator with efficiency e, obtain the efficiency of $\lambda t + (1 - \lambda) t'$, where λ is a constant.

4. Show that the sample covariance matrix S is an unbiased estimator of Σ.

5. Specify and explain the qualities of good estimator.

[A.U N/D 2013 (MBA)]

6. A random sample of size 100 has mean 15, the population variance being 25. Find the interval estimate of the population mean with a confidence level of

   (1) 99% and (2) 95%

[A.U. N/D 2013 (MBA)]

7. What are the properties of an estimator ? [A.U A/M 2015 (MBA)]

8. Estimate the population mean at 95% confidence interval from the following sample data drawn from normal distribution.

X	F
10	6
20	10
30	4

[A.U A/M 2017 (R-13)]

9. Suppose that the heights of 100 male students at XYZ university represent a random sample of the heights of all male students at the university. Find (1) 95% (2) 99% confidence intervals for estimating the mean height of the XYZ university students.

[A.U A/M 2017 (R-13)]

10. Estimate α and β in the Gamma distribution with the density function

$$f(x; \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha - 1} . e^{-\beta x} ; x \geq 0 ; \alpha , \beta > 0$$

11. If $(x_1, x_2, ..., x_n)$ is a random sample from a population with density function $f(x; \theta, \mu) = \theta e^{-\theta (x - \mu)}$; $x > \mu$. Find the method of moments estimators of θ and μ.

12. Find the maximum likelihood estimators of a and b for the uniform population, the density function of which is given by

$$f(x; a, b) = \frac{1}{b - a}, a \leq a \leq b$$

13. Find the maximum likelihood estimator of the parameter θ on the population given by $f(x, \theta) = \dfrac{1}{\theta^{P}\,\Gamma(p)}\, x^{p-1}\, e^{-x/\theta}$; $x \geq 0$ and $p\,(>0)$ is known. Also find its variance.

14. The daily number of accidential disconnects with a server follows a Poisson distribution. One five days

2	5	3	3	7

accidental disconnects are observed.

(a) Obtain the maximum likelihood estimate of λ.

(b) Find the maximum likelihood estimate of the probability that 3 or more accidental disconnects will occur.

UNIT-III.

PART-A

1. **Write the difference between Point Estimate and Interval Estimate of population parameters. [A.U M/J 2013, M/J 2014, Jan. 2015 (MBA)]**

Ans.

Point Estimate of Population Parameters	Interval Estimate of Population Parameters
A point estimate of a population parameter is a single value of a statisic.	An interval estimate is defined by two numbers, between which a population parameter is said to lie.
For example, the sample mean ($\overline{X}$) is a point estimate of the population mean μ. Similarly, the sample proportion p is a point estimate of the population proportion P.	For example, $a < X < b$ is an interval estimate of the population mean μ. It indicates that the population mean is greater than a but less than b.

2. **What is an Estimator ?**

Ans. Estimator (or) Point estimator is a procedure for producing an estimate of a parameter of interest. An estimator is usually a function of only sample data values, and when these data values are available, it results in an estimator of the parameter of interest.

3. **Define a point estimator.**

Ans. A point estimator of some population parameter θ is a single numerical values $\overline{\theta}$ of a statistic $\overline{\theta}$. The statistic $\overline{\theta}$ is called the point estimator.

4. **What are the characteristics that should be satisfied by a good estimator?**

Ans.

(i) Consistency (ii) Unbiasedness (iii) Efficiency and (iv) Sufficiency

5. **What are the commonly used methods of point estimation ?**

Ans.

1. Methods of maximum likelihood estimator.
2. Method of minimum variance
3. Method of moments

4. Method of least squares

5. Method of minimum chi-square.

6. Method of inverse probability

6. Define the minimum variance unbiased estimator (MVUE).

Ans. If all the unbiased estimators of θ are considered the one with the smallest variance is called the minimum variance unbiased estimator.

7. Define Population Estimation. **[A.U N/D 2019 MBA]**

Ans. Population estimates can describe the total population size as well as demographic characteristics such as age, sex, or education level.

Population estimates are dependent on the demographic components of change : Mortality, Fertility and Migration.

8. What is called interval estimation ? **[A.U N/D 2018 (R17) MBA]**

Ans. In an estimate of a population parameter θ given by two distinct numbers between which the parameter may be considered to lie, then the estimate is called an interval estimate of the parameter θ.

9. Define point estimate and interval estimate.

Ans. See Text Book Page No. 3.52 Q.No. 1

10. Define unbiased estimator. **[A.U A/M 2017 MBA]**

Ans. See Text Book Page No. 3.4

11. What is an example of a robust statistic?

Ans. Using the same dataset : 50, 52, 55, 56, 59, 59, 60 if we changed [illegible] 1000, the median is entirely unaffected. It's still 56 in both cases. [illegible]tly, the median is a robust statistic for central tendency while the [illegible] is not.

12. How do you calculate robust mean ?

Ans. It is given by $(x_1 + x_2 + \ldots + x_n)/n$. You can change the calculated value of the mean by an arbitrary large amount, simply by changing one of the data points by a large amount. Therefore, the breakdown point is just $1/n$.

13. Define maximum likelihood estimator of θ.

Ans. The maximum likelihood estimator of θ is the value of θ that maximizes the likelihood function $L(\theta) = P_{i=1}^{n} f(x_i, \theta)$

14. Mention the properties of the maximum likelihood estimator.

Ans. Under very general and not restrictive conditions, when the sample size n is large and if $\overline{\theta}$ is the maximum likelihood estimator of the parameter θ,

(i) $\overline{\theta}$ is an approximately unbiased estimator of θ, $[E(\overline{\theta}) = \theta]$.

(ii) The variance of $\overline{\theta}$ is nearly as small as the variance that could be obtained with any other estimator.

(iii) $\overline{\theta}$ has an approximate normal distribution.

(iv) Invariance property

If $\overline{\theta}_1, \overline{\theta}_2, ..., \overline{\theta}_k$ are the maximum likelihood estimators of the parameters $\theta_1, \theta_2, ..., \theta_k$, then the maximum likelihood estimator of any function $h(\theta_1, \theta_2, ..., \theta_k)$ of these parameters is the same functions $h(\overline{\theta}_1, \overline{\theta}_2, ..., \overline{\theta}_k)$ of the estimators $\overline{\theta}_1, \overline{\theta}_2, ..., \overline{\theta}_k$

15. Define the population moment estimators.

Ans. Let $X_1, X_2, ..., X_n$ be a random sample from either a probability mass function or probability density function with m uknown parameters $\theta_1, \theta_2, ..., \theta_m$. The moment estimator $\overline{\theta}_1, \overline{\theta}_2, ..., \overline{\theta}_n$ are found by equating the first m population moments to the first m sample moments and solving the resulting equations for the uknown parameters.