UNIT –IV

SAMPLING

Introduction: The totality of observations with which we are concerned , whether this number be finite or infinite constitute population. In this chapter we focus on sampling from distributions or populations and such important quantities as the sample mean and sample variance.

Def: Population is defined as the aggregate or totality of statistical data forming a subject of investigation .

EX. The population of the heights of Indian.

The number of observations in the population is defined to be the size of the population. It may be finite or infinite .Size of the population is denoted by N.As the study of entire population may not be possible to carry out and hence a part of the population alone is selected.

Def: A portion of the population which is examined with a view to determining the population characteristics is called a sample . In other words, sample is a subset of population. Size of the sample is denoted by n.

The process of selection of a sample is called Sampling. There are different methods of sampling

- Probability Sampling Methods
- Non-Probability Sampling Methods

Probability Sampling Methods:

a) **Random Sampling (Probability Sampling**):
It is the process of drawing a sample from a population in such a way that each member of the population has an equal chance of being included in the sample.
Ex: A hand of cards from a well shuffled pack of cards is a random sample.
Note : If N is the size of the population and n is the size of the sample, then
   - The no. of samples with replacement = N^n
   - The no. of samples without replacement = N_{C_n}

b) **Stratified Sampling :**
In this , the population is first divided into several smaller groups called strata according to some relevant characteristics . From each strata samples are selected at random, all the samples are combined together to form the stratified sampling.

c) **Systematic Sampling (Quasi Random Sampling):**
In this method , all the units of the population are arranged in some order . If the population size is N, and the sample size is n, then we first define sample interval denoted by = $\frac{N}{n}$. then from first k items ,one unit is selected at random. Then from

first unit every k^{th} unit is serially selected combining all the selected units constitute a systematic sampling.

Non Probability Sampling Methods:

a) **Purposive (Judgment) Sampling :**
In this method, the members constituting the sample are chosen not according to some definite scientific procedure , but according to convenience and personal choice of the individual who selects the sample . It is the choice of the individual items of a sample entirely depends on the individual judgment of the investigator.

b) **Sequential Sampling:**
It consists of a sequence of sample drawn one after another from the population. Depending on the results of previous samples if the result of the first sample is not acceptable then second sample is drawn and the process continues to take proper decision . But if the first sample is acceptable ,then no new sample is drawn.

Classification of Samples:

- **Large Samples :** If the size of the sample $n \geq 30$, then it is said to be large sample.
- **Small Samples :** If the size of the sample $n < 30$,then it is said to be small sample or exact sample.

Parameters and Statistics:

Parameter is a statistical measure based on all the units of a population. Statistic is a statistical measure based on only the units selected in a sample.
Note :In this unit , Parameter refers to the population and Statistic refers to sample.

Central Limit Theorem: If $\bar{x}$ be the mean of a random sample of size n drawn from population having mean μ and standard deviation σ , then the sampling distribution of the sample mean $\bar{x}$ is approximately a normal distribution with mean μ and SD = S.E of $\bar{x} = \frac{\sigma}{\sqrt{n}}$ provided the sample size n is large.

Standard Error of a Statistic : The standard error of statistic 't' is the standard deviation of the sampling distribution of the statistic i.e, S.E of sample mean is the standard deviation of the sampling distribution of sample mean.

Formulae for S.E:

- S.E of Sample mean $\bar{x} = \frac{\sigma}{\sqrt{n}}$ i.e, S.E $(\bar{x}) = \frac{\sigma}{\sqrt{n}}$

- S.E of sample proportion p= $\sqrt{\frac{PQ}{n}}$ i.e, S.E (p) = $\sqrt{\frac{PQ}{n}}$ where Q=1-P
- S.E of the difference of two sample means $\overline{x_1}$ and $\overline{x_2}$ i.e, S.E ($\overline{x_1} - \overline{x_2}$) = $\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$
- S.E of the difference of two proportions i.e, S.E($p_1 - p_2$) = $\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}$

Estimation :

To use the statistic obtained by the samples as an estimate to predict the unknown parameter of the population from which the sample is drawn.

Estimate : An estimate is a statement made to find an unknown population parameter.

Estimator : The procedure or rule to determine an unknown population parameter is called estimator.

Ex. Sample proportion is an estimate of population proportion , because with the help of sample proportion value we can estimate the population proportion value.

Types of Estimation:

- **Point Estimation:** If the estimate of the population parameter is given by a single value , then the estimate is called a point estimation of the parameter.
- **Interval Estimation:** If the estimate of the population parameter is given by two different values between which the parameter may be considered to lie, then the estimate is called an interval estimation of the parameter.

Confidence interval Estimation of parameters:

In an interval estimation of the population parameter θ, if we can find two quantities t_1 and t_2 based on sample observations drawn from the population such that the unknown parameter θ is included in the interval $[t_{1,}t_2]$ in a specified cases ,then this is called a confidence interval for the parameter θ.

Confidence Limits for Population mean μ

- 95% confidence limits are $\bar{x} \pm 1.96\ (S.E.of\ \bar{x})$
- 99% confidence limits are $\bar{x} \pm 2.58\ (S.E.of\ \bar{x})$
- 99.73% confidence limits are $\bar{x} \pm 3\ (S.E.of\ \bar{x})$
- 90% confidence limits are $\bar{x} \pm 1.645\ (S.E.of\ \bar{x})$

Confidence limits for population proportion P

- 95% confidence limits are p $\pm$ 1.96(S.E.of p)
- 99% confidence limits are p $\pm$ 2.58(S.E. of p)

- 99.73% confidence limits are $p \pm 3$(S.E.of p)
- 90% confidence limits are $p \pm 1.645$(S.E.of p)

Confidence limits for the difference of two population means μ_1 and μ_2

- 95% confidence limits are $((\overline{x_1} - \overline{x_2}) \pm 1.96$ (S.E of $((\overline{x_1} - \overline{x_2}))$
- 99% confidence limits are $((\overline{x_1} - \overline{x_2}) \pm 2.58$ (S.E of $((\overline{x_1} - \overline{x_2}))$
- 99.73% confidence limits are $((\overline{x_1} - \overline{x_2}) \pm 3$ (S.E of $((\overline{x_1} - \overline{x_2}))$
- 90% confidence limits are $((\overline{x_1} - \overline{x_2}) \pm 2.58$ (S.E of $((\overline{x_1} - \overline{x_2}))$

Confidence limits for the difference of two population proportions

- 95% confidence limits are p_1-p_2 ± 1.96 (S.E. of p_1-p_2)
- 99% confidence limits are p_1-p_2 ± 2.58 (S.E. of p_1-p_2)
- 99.73% confidence limits are p_1-p_2 ± 3 (S.E. of p_1-p_2)
- 90% confidence limits are p_1-p_2 ± 1.645 (S.E. of p_1-p_2)

Determination of proper sample size

Sample size for estimating population mean :

n= $\left(\frac{z_\alpha \sigma}{E}\right)^2$ where $z_\alpha -$ Critical value of z at α Level of significance

$\sigma -$ Standard deviation of population and

E – Maximum sampling Error = $\bar{x} - \mu$

Sample size for estimating population proportion :

$n = \frac{{z_\alpha}^2 PQ}{E^2}$ where z_α – Critical value of z at α Level of significance

P $-$ Population proportion

$Q -$ 1-P

$E -$ Maximum Sampling error = p-P

Testing of Hypothesis :

It is an assumption or supposition and the decision making procedure about the assumption whether to accept or reject is called hypothesis testing .

Def: Statistical Hypothesis : To arrive at decision about the population on the basis of sample information we make assumptions about the population parameters involved such assumption is called a statistical hypothesis .

Procedure for testing a hypothesis:

Test of Hypothesis involves the following steps:

Step1: Statement of hypothesis :

There are two types of hypothesis :

- **Null hypothesis:** A definite statement about the population parameter. Usually a null hypothesis is written as no difference , denoted by H_0.
  Ex. $H_0: \mu = \mu_0$
- **Alternative hypothesis :** A statement which contradicts the null hypothesis is called alternative hypothesis. Usually an alternative hypothesis is written as some difference , denoted by H_1.
  Setting of alternative hypothesis is very important to decide whether it is two-tailed or one – tailed alternative , which depends upon the question it is dealing.
  Ex.$H_1: \mu \neq \mu_0$ (Two – Tailed test)
  or
  $H_1: \mu > \mu_0$ (Right one tailed test)
  or
  $H_1: \mu < \mu_0$ (Left one tailed test)

Step 2: Specification of level of significance :

The LOS denoted by α is the confidence with which we reject or accept the null hypothesis. It is generally specified before a test procedure ,which can be either 5% (0.05) , 1% or 10% which means that thee are about 5 chances in 100 that we would reject the null hypothesis H_0 and the remaining 95% confident that we would accept the null hypothesis H_0 . Similarly , it is applicable for different level of significance.

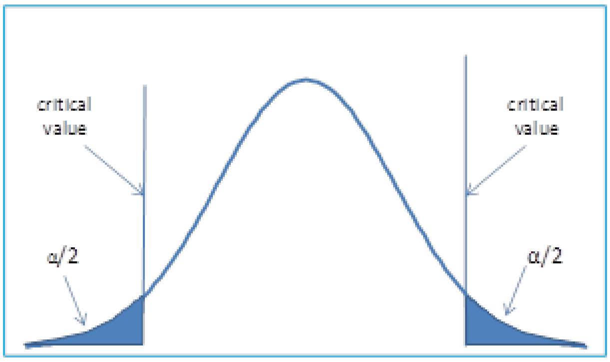
Step 3 : Identification of the test Statistic :

There are several tests of significance like z,t, F etc .Depending upon the nature of the information given in the problem we have to select the right test and construct the test criterion and appropriate probability distribution.

Step 4: Critical Region:

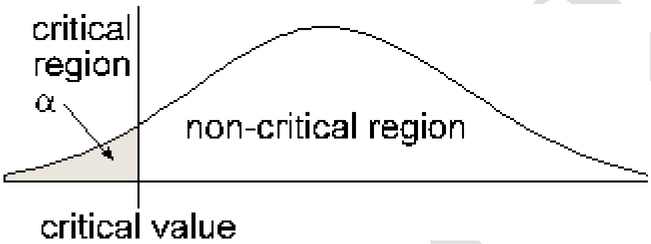
It is the distribution of the statistic .

- **Two – Tailed Test : The critical region under the curve is equally distributed on both sides of the mean.**
  If H_1 has $\neq$ sign , the critical region is divided equally on both sides of the distribution.

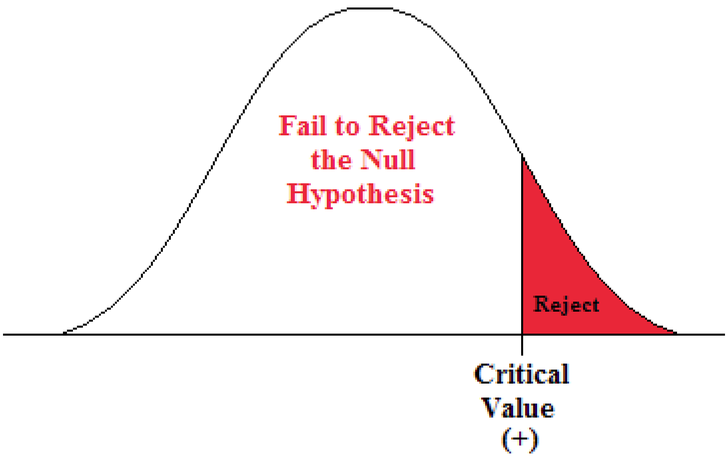


- **One Tailed Test: The critical region under the curve is distributed on one side of the mean.**

Left one tailed test: If H_1 has $<$ sign , the critical region is taken in the left side of the distribution.



Right one tailed test : If H_1 has $>$ sign , the critical region is taken on right side of the distribution.



Step 5 : Making decision:

By comparing the computed value and the critical value decision is taken for accepting or rejecting H_0

If calculated value $\leq$ critical value , we accept H_0, otherwise reject H_0.

Errors of Sampling :

While drawing conclusions for population parameters on the basis of the sample results , we have two types of errors.

- **Type I error :** Reject H_0 when it is true i.e, if the null hypothesis H_0 is true but it is rejected by test procedure .
- **Type II error :** Accept H_0 when it is false i.e, if the null hypothesis $H_0$is false but it is accepted by test procedure.

DECISION TABLE

	H_0 **is accepted**	H_0 **is rejected**
H_0 **is true**	**Correct Decision**	**Type I Error**
H_0 **is false**	**Type II Error**	**Correct Decision**

Problems:

1.**If the population is 3,6,9,15,27**

a) **List all possible samples of size 3 that can be taken without replacement from finite population**
b) **Calculate the mean of each of the sampling distribution of means**
c) **Find the standard deviation of sampling distribution of means**

Sol: Mean of the population , $\mu = \frac{3+6+9+15+27}{5} = \frac{60}{5}$ =12

Standard deviation of the population ,

$$\sigma = \sqrt{\frac{(3-12)^2+(6-12)^2+(9-12)^2+(15-12)^2+(27-12)^2}{5}}$$

$$= \sqrt{\frac{81+36+9+9+225}{5}} = \sqrt{\frac{360}{5}} = 8.4853$$

a) Sampling without replacement :
The total number of samples without replacement is $N_{C_n} = 5_{C_3}$=10
The 10 samples are (3,6,9), (3,6,15), (3,9,15), (3,6,27), (3,9,27), (3,15,27), (6,9,15), (6,9,27), (6,15,27), (9,15,27)

b) Mean of the sampling distribution of means is
$\mu_{\bar{x}} = \frac{6+8+9+10+12+13+14+15+16+17}{10} = \frac{120}{10} = 12$

c) $\sigma^2 = \frac{(6-12)^2+(8-12)^2+(9-12)^2+(10-12)^2+(12-12)^2+(13-12)^2+(14-12)^2+(15-12)^2+(16-12)^2+(17-12)^2}{10}$

$= 13.3$

$\therefore\ \sigma_{\bar{x}} = \sqrt{13.3} = 3.651$

2.**A population consist of five numbers 2,3,6,8 and 11. Consider all possible samples of size two which can be drawn with replacement from this population .Find**

- **a) The mean of the population**
- **b) The standard deviation of the population**
- **c) The mean of the sampling distribution of means and**
- **d) The standard deviation of the sampling distribution of means**

Sol: a) Mean of the Population is given by

$$\mu = \frac{2+3+6+8+11}{5} = \frac{30}{5} = 6$$

b) Variance of the population is given by

$$\sigma^2 = \sum \frac{(x_i - \bar{x})^2}{n}$$

$$= \frac{(2-6)^2+(3-6)^2+(6-6)^2+(8-6)^2+(11-6)^2}{5}$$

$$= \frac{16+9+0+4+25}{5} = 10.8 \qquad \therefore\ \sigma = 3.29$$

c) Sampling with replacement

The total no.of samples with replacement is $N^n = 5^2 = 25$

∴ List of all possible samples with replacement are

$$\begin{Bmatrix} (2,2),(2,3),(2,6),(2,8),(2,11),(3,2),(3,3)(3,6),(3,8),(3,11) \\ (6,2),(6,3),(6,6),(6,8),(6,11),(8,2),(8,3),(8,6),(8,8),(8,11) \\ (11,2),(11,3),(11,6),(11,8),(11,11) \end{Bmatrix}$$

Now compute the arithmetic mean for each of these 25 samples which gives rise to the distribution of means of the samples known as sampling distribution of means

The samples means are

$$\begin{pmatrix} 2\ ,2.5\ ,4\ ,5\ ,6.5 \\ 2.5\ ,3\ ,4.5,\ ,5.5,7 \\ 4,4.5,6,7,8.5 \\ 5,5.5,7,8,9.5 \\ 6.5,7,8.5,9.5,11 \end{pmatrix}$$

And the mean of sampling distribution of means is the mean of these 25 means

$$\mu_{\bar{x}} = \frac{\text{sum of all above sample means}}{25} = \frac{150}{25} = 6$$

d) The variance of the sampling distribution of means is obtained by subtracting the mean 6 from each number in sampling distribution of means and squaring the result ,adding all 25 numbers thus obtained and dividing by 25.

$$\sigma^2 = \frac{(2-6)^2+(2.5-6)^2+(4-6)^2+(5-6)^2+\cdots\ldots(11-6)^2}{25} = \frac{135}{25} = 5.4$$

$\therefore \sigma = \sqrt{5.4} = 2.32$

3.When a sample is taken from an infinite population , what happens to the standard error of the mean if the sample size is decreased from 800 to 200

Sol: The standard error of mean = $\frac{\sigma}{\sqrt{n}}$

Sample size = n .let n= n_1=800

Then $S.E_1 = \frac{\sigma}{\sqrt{800}} = \frac{\sigma}{20\sqrt{2}}$

When $n_1$is reduced to 200

let n= n_2=200

Then $S.E_2 = \frac{\sigma}{\sqrt{200}} = \frac{\sigma}{10\sqrt{2}}$

$\therefore S.E_2 = \frac{\sigma}{10\sqrt{2}} = 2\left(\frac{\sigma}{20\sqrt{2}}\right) = 2\,(S.E_1)$

Hence if sample size is reduced from 800to 200, S. E. of mean will be multiplied by 2

4.The variance of a population is 2 . The size of the sample collected from the population is 169. What is the standard error of mean

Sol: n= The size of the sample =169

$\sigma =$ S.D of population = $\sqrt{\text{Variance}} = \sqrt{2}$

Standard Error of mean = $\frac{\sigma}{\sqrt{n}} = \frac{\sqrt{2}}{\sqrt{169}} = \frac{1.41}{13} = 0.185$

5.The mean height of students in a college is 155cms and standard deviation is 15 . What is the probability that the mean height of 36 students is less than 157 cms.

Sol: μ = Mean of the population

= Mean height of students of a college = 155cms

n = S.D of population = 15cms

$\overline{x}$ = mean of sample = 157 cms

Now $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{157-155}{\frac{15}{\sqrt{36}}} = \frac{12}{15} = 0.8$

$\therefore P(\bar{x} \leq 157) = P(z < 0.8) = 0.5 + P(0 \leq z \leq 0.8)$

= 0.5 +0.2881 = 0.7881

Thus the probability that the mean height of 36 students is less than 157 = 0.7881

6.A random sample of size 100 is taken from a population with σ = 5.1 . Given that the sample mean is $\bar{x}$ = 21.6 Construct a 95% confidence limits for the population mean .

Sol: Given $\bar{x}$ = 21.6

$z_{\alpha/2}$ = 1.96, n = 100 , σ = 5.1

$\therefore$ Confidence interval = $\left(\bar{x} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} , \bar{x} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \right)$

$\bar{x} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 21.6 - \frac{1.96 \text{ x } 5.1}{10} = 20.6$

$\bar{x} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} = 21.6 + \frac{1.96 \text{ x } 5.1}{10} = 22.6$

Hence (20.6,22.6) is the confidence interval for the population mean μ

7.It is desired to estimate the mean time of continuous use until an answering machine will first require service . If it can be assumed that σ = 60 days, how large a sample is needed so that one will be able to assert with 90% confidence that the sample mean is off by at most 10 days.

Sol: We have maximum error (E) = 10 days , σ = 60 days and $z_{\alpha/2}$ = 1.645

$\therefore$ n = $\left[\frac{z_{\alpha/2}.\sigma}{E}\right]^2 = \left[\frac{1.645 \text{ x } 60}{10}\right]^2 = 97$

8.A random sample of size 64 is taken from a normal population with $\mu = 51.4$ and σ = 6.8.What is the probability that the mean of the sample will a) exceed 52.9 b) fall between 50.5 and 52.3 c) be less than 50.6

Sol: Given n = the size of the sample = 64

μ = the mean of the population = 51.4

σ = the S.D of the population = 6.8

a) P($\bar{x}$ exceed 52.9) = P($\bar{x}$ > 52.9)

Z = $\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{52.9-51.4}{\frac{6.8}{\sqrt{64}}} = 1.76$

$\therefore$ P($\bar{x}$ > 52.9) = P(z > 1.76)

= 0.5 – P(0 < z < 1.76)

= 0.5 – 0.4608 = 0.0392

b) P($\bar{x}$ fall between 50.5 and 52.3)

i.e, P(50.5 < $\bar{x}$ < 52.3) = P($\overline{x_1} < \bar{x} < \overline{x_2}$)

$z_1 = \frac{\overline{x_1}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{50.5-51.4}{0.85} = -1.06$

$z_2 = \frac{\overline{x_2}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{52.3-51.4}{0.85} = 1.06$

P(50.5 < $\bar{x}$ < 52.3) = P(-1.06 < z < 1.06)

= P(-1.06 < z < 0) + P(0 < z < 1.06)

= P(0 < z < 1.06) + P(0 < z < 1.06)

= 2(0.3554) = 0.7108

c) P($\bar{x}$ will be less than 50.6) = P($\bar{x}$ < 50.6)

Z = $\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{50.6-51.4}{\frac{6.8}{\sqrt{64}}}$ = -0.94

∴P(z < -0.94) = 0.5 - P(0.94 < z < 0)

= 0.5 - P(0 < z < 0.94) = 0.50-0.3264

= 0.1736

9.The mean of certain normal population is equal to the standard error of the mean of the samples of 64 from that distribution . Find the probability that the mean of the sample size 36 will be negative.

Sol: The Standard error of mean = $\frac{\sigma}{\sqrt{n}}$

Sample size , n =64

Given mean , μ = Standard error of the mean of the samples

$\mu = \frac{\sigma}{\sqrt{64}} = \frac{\sigma}{8}$

We know z = $\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x}-\frac{\sigma}{8}}{\frac{\sigma}{6}}$

= $\frac{6\bar{x}}{\sigma}$ - $\frac{3}{4}$

If Z < 0.75, $\bar{\bar{x}}$ is negative

P(z < 0.75) = P($-\infty < z < 0.75$)

= $\int_{-\infty}^{0} \emptyset(z)\, dz + \int_{0}^{0.75} \emptyset(z) dz$ = 0.50 + 0.2734

= 0.7734

10.The guaranteed average life of a certain type of electric bulbs is 1500hrs with a S.D of 10 hrs. It is decided to sample the output so as to ensure that 95% of bulbs do not fall short of the guaranteed average by more than 2% . What will be the minimum sample size ?

Sol : Let n be the size of the sample

The guaranteed mean is 1500

We do not want the mean of the sample to be less than 2% of (1500) i.e, 30 hrs

So 1500 – 30 = 1470

∴ $\bar{x} > 1470$

$\therefore |z| = \left|\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| = \left|\frac{1470-1500}{\frac{120}{\sqrt{n}}}\right| = \frac{\sqrt{n}}{4}$

From the given condition , the area of the probability normal curve to the left of $\frac{\sqrt{n}}{4}$ should be 0.95

$\therefore$ The area between 0 and $\frac{\sqrt{n}}{4}$ is 0.45

We do not want to know about the bulbs which have life above the guranteed life .

$\therefore \frac{\sqrt{n}}{4} = 1.65$ i.e., $\sqrt{n} = 6.6$

$\therefore$ n = 44

11.A normal population has a mean of 0.1 and standard deviation of 2.1 . Find the probability that mean of a sample of size 900 will be negative .

Sol : Given $\mu = 0.1$, $\sigma = 2.1$ and n = 900

The Standard normal variate

$Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x}-\mu}{\frac{2.1}{\sqrt{900}}} = \frac{\bar{x}-0.1}{0.07}$

$\therefore \bar{x} = 0.1 + 0.007\, z$ where $z \sim N(0,1)$

$\therefore$ The required probability , that the sample mean is negative is given by

$P(\bar{x} < 0) = P(0.1 + 0.07\, z < 0)$

$= P(0.07\, z < -0.1)$

$= P\left(z < \left(\frac{-0.1}{0.07}\right)\right.$

$= P(z < -1.43)$

$= 0.50 - P(0 < z < 1.43)$

$= 0.50 - 0.4236 = 0.0764$

12.In a study of an automobile insurance a random sample of 80 body repair costs had a mean of Rs 472.36 and the S.D of Rs 62.35. If $\bar{\bar{x}}$ is used as a point estimator to the true average repair costs , with what confidence we can assert that the maximum error doesn't exceed Rs 10.

Sol : Size of a random sample , n = 80

The mean of random sample , $\bar{x}$ = Rs 472.36

Standard deviation , σ = Rs 62.35

Maximum error of estimate , E_{max} = Rs 10

We have $E_{max} = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

i.e., $Z_{\alpha/2} = \frac{E_{max} \cdot \sqrt{n}}{\sigma} = \frac{10\sqrt{80}}{62.35} = \frac{89.4427}{62.35} = 1.4345$

$\therefore Z_{\alpha/2} = 1.43$

The area when z = 1.43 from tables is 0.4236

$\therefore \frac{\alpha}{2} = 0.4236$ i.e ., $\alpha = 0.8472$

$\therefore$ confidence = (1- α) 100% = 84.72 %

Hence we are 84.72% confidence that the maximum error is Rs. 10

13.If we can assert with 95% that the maximum error is 0.05 and P = 0.2 find the size of the sample.

Sol : Given P =0.2 , E = 0.05

We have Q = 0.8 and $Z\alpha/_2$= 1.96 (5% LOS)

We know that maximum error , E = $Z\alpha/_2 \sqrt{\frac{PQ}{n}}$

$\Rightarrow$ 0.05 = 1.96 $\sqrt{\frac{0.2 \text{ x } 0.8}{n}}$

$\Rightarrow$ Sample size , n = $\frac{0.2 \text{ x } 0.8 \text{ x } (1.96)^2}{(0.05)^2}$ = 246

14.The mean and standard deviation of a population are 11,795 and 14,054 respectively . What can one assert with 95 % confidence about the maximum error if $\bar{x}$ = 11,795 and n = 50. And also construct 95% confidence interval for true mean .

Sol: Here mean of population , μ = 11795

S.D of population , σ = 14054
$\bar{x}$ = 11795
$n =$ sample size = 50 , maximum error = $Z\alpha/_2 \cdot \frac{\sigma}{\sqrt{n}}$
$Z\alpha/_2$ for 95% confidence = 1.96
$\text{Max. error}, E = Z\alpha/_2 \cdot \frac{\sigma}{\sqrt{n}}$ = 1.96 .$\frac{14054}{\sqrt{50}}$ = 3899

$\therefore$ Confidence interval = ($\bar{x} - Z\alpha/_2 \cdot \frac{\sigma}{\sqrt{n}}$, $\bar{x} + Z\alpha/_2 \cdot \frac{\sigma}{\sqrt{n}}$)
= (11795-3899, 11795+3899)
= (7896, 15694)

15.Find 95% confidence limits for the mean of a normally distributed population from which the following sample was taken 15, 17 , 10 ,18 ,16 ,9, 7, 11, 13 ,14.

Sol: We have $\bar{x}$ = $\frac{15+17+10+18+16+9+7+11+13+14}{10}$ = 13

$S^2 = \sum \frac{(x_i - \bar{x})^2}{n-1}$

$= \frac{1}{9}$ $[(15-13)^2 + (15-13)^2 + (15-13)^2 + (15-13)^2 + (15-13)^2 + (15-13)^2 + (15-13)^2 + (15-13)^2 + (15-13)^2 + (15-13)^2]$

$= \frac{40}{3}$

Since $Z\alpha/_2$ = 1.96 , we have

$$Z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 1.96 \cdot \frac{\sqrt{40}}{\sqrt{10}.\sqrt{3}} = 2.26$$

∴ Confidence limits are $\bar{x} \pm Z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$ = 13 $\pm$ 2.26 = (10.74 , 15.26)

16.A random sample of 100 teachers in a large metropolitan area revealed mean weekly salary of Rs. 487 with a standard deviation Rs.48. With what degree of confidence can we assert that the average weekly of all teachers in the metropolitan area is between 472 to 502 ?

Sol: Given $\mu = 487$, $\sigma = 48$, $n = 100$

$$Z = \frac{\bar{\bar{x}} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{\bar{x} - 487}{\frac{48}{\sqrt{100}}} = \frac{\bar{x} - 487}{4.8}$$

Standard variable corresponding to Rs. 472 is

$$Z_1 = \frac{472 - 487}{4.8} = -3.125$$

Standard vaiable corresponding to Rs. 502

$$Z_2 = \frac{502 - 487}{4.8} = 3.125$$

Let $\bar{x}$ be the mean salary of teacher . Then

P (472 < $\bar{x}$ < 502) = P (-3.125 < z < 3.125)

= 2 (0 < z < 3.125)

$= 2 \int_0^{3.125} \emptyset(z)dz$

= 2 (0.4991) = 0.9982

Thus we can ascertain with 99.82 % confidence