3 Chap 3: Sampling Distribution:

(~~Parameter~~)

4 Population: Population is the aggregate or total of statistical data forming a subject of investigation

for eg: population of workers in a factory, population of nationalized bank in India etc.

5. → The no. of observations in a population is defined to be size of population and is denoted by 'N'.

→ If the no. of units in the population is finite then the population is called finite population

eg: The population of students in a clg

→ If the size of population is infinite then it is called infinite population

eg: The population of stars in the sky

Sample: Most of the times study of entire population may not be possible in such a cases we can select some units from the population i.e a part (or) subset of population which is examined with a view of to determine the population characteristic is called a sample.

→ The no. of obs in the sample is called sample size and is denoted by 'n'.

→ If the sample size 'n' is greater than or equal to 30 then the sample is called large sample.

→ If the sample size 'n' is less than 30 then the sample is called small sample.

→ **Sampling Techniques or sampling methods:** The technique of drawing samples from the population is called sampling methods. Some imp methods of sampling are discussed below

① **Simple random sampling method (SRSM):** It is the process of drawing a sample from the population in such a way each and every unit of population has an equal chance of being included in the sample. The sample obtained by this process is called random sample.

→ If each element of a population may be selected more than once then it is called simple random sampling with replacement & whereas if the element cannot be selected more than once then it is called simple random sampling without replacement

→ The no. of samples with replacement is 'N^n'

→ The no. of samples without replacement is '^NC_n'

where N = Population size
n = sample size

② **Stratified random sampling:** This method is useful when the population is heterogeneous. In this type of sampling the population is first subdivided into several groups. These groups are called stratums then a sample is selected from each stratum at randomly. The process of obtaining and examining a stratified

3 sample with a view of to estimating the characteristics of the population is known as stratified random sampling.

4 (3) Systematic sampling method:- In this method all the units of population are arranged in some order then from the first 'k' items 1 unit is selected at random. By continuing

5 this process we get systematic sample.

Parameter:- It is a statistical measure based on all the observations of population.

Population mean μ, Population standard deviation σ, Population Proportion 'P' etc are usually referred as a parameter.

Statistics:- It is a statistical measure based only on all the units selected in a sample.

sample mean $\bar{x}$, sample s.d 's', sample proportion 'p' are usually referred to as statistics

Sampling distribution of a statistic:- sampling theory is the study of relationship b/w a population and samples drawn from the population.

Sampling distribution of a statistic is the frequency distribution which is formed with various values of a statistic computed from diff. samples of the same size drawn from the same population these various values can be arranged into a frequency distribution table which is known as the sampling distribution of the statistic.

Let us consider a finite population of size 'N' and drawn all possible random samples of each of same size 'n', then we get 'k' samples. Now the sampling distribution of mean and standard deviation can be tabulated as follows

Samples	Sample units	Sample mean	Sample S.D
1	$x_{11}, x_{12} \cdots x_{1n}$	$\bar{x}_1$	s_1
2	$x_{21}, x_{22} \cdots x_{2n}$	$\bar{x}_2$	s_2
⋮	⋮	⋮	⋮
k	$x_{k1}, x_{k2} \cdots x_{kn}$	$\bar{x}_k$	s_k

Standard error of a statistic: The standard error of a statistic is the standard deviation of the sampling distribution of statistic

Formulas for standard error:

1. standard error of sample mean $\bar{x} = S.E(\bar{x}) = \frac{\sigma}{\sqrt{n}}$.

2. S.E of the diff of 2 sample mean. SE of $(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

3. S.E of sample s.d 's' $= S.E(s) = \frac{\sigma}{\sqrt{2n}}$

4. S.E of sample proportion 'p' $= S.E(P) = \sqrt{\frac{PQ}{n}}$; $Q = 1 - P$

5. S.E of diff of two proportions $= S.E(P_1 - P_2) = \sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}$

Note: ① Suppose the samples are drawn from an infinite population i.e sampling is done with replacement then mean of sampling distribution of

3 Means is equal to population mean i.e $\mu_{\bar{x}} = \mu$

$$\text{Variance}(\bar{x}) = \sigma^2/n$$

4 ② If the samples are drawn from a finite population i.e sampling is done without replacement then the mean of sampling distribution is = population mean

5 i.e $\mu_{\bar{x}} = \mu$

$$V(\bar{x}) = \left(\frac{N-n}{N-1}\right) \cdot \frac{\sigma^2}{n}$$

Here $\frac{N-n}{N-1}$ is known as finite population correction factor

Prob

1. Find the value of finite population correction factor for $n = 10$ & $N = 100$

$$\left(\frac{N-n}{N-1}\right) = \left(\frac{100-10}{100-1}\right) = \frac{90}{99} = 0.9091$$

2. A population consists of 5 numbers 2, 3, 6, 8, 11 consider all possible samples of size 2 which can be drawn with replacement from this population. Find ① The population mean ② The S.D of population ③ The mean of sampling distribution of means ④ The S.D of sampling distribution of means

sol $N = 5$, $n = 2$

① Population mean $= \mu = \frac{2+3+6+8+11}{5} = 6 = \mu$

② $\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} = \sqrt{\frac{(2-6)^2+(3-6)^2+(6-6)^2+(8-6)^2+(11-6)^2}{5}} = 3.2863$

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

(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)

Sample mean ($\bar{x}$) – 2, 2.5, 4, 5, 6.5
2.5, 3, 4.5, 5.5, 7
4, 4.5, 6, 7, 8.5 = 150
5, 5.5, 7, 8, 9.5
6.5, 7, 8.5, 9.5, 11

$(x-\mu)^2$ = 16, 12.25, 4, 1, 0.25
12.25, 9, 2.25, 0.25, 1
4, 2.25, 0, 1, 6.25 = 135
1, 0.25, 1, 4, 12.25
0.25, 1, 6.25, 12.25, 25

③ The mean of sampling distribution of means $\mu_{\bar{x}} = \frac{\Sigma\bar{x}}{25} = \frac{150}{25} = 6$

$\mu_{\bar{x}} = \mu$

④ The S.D of sampling distribution of means $\sigma_{\bar{x}} = \sqrt{\frac{\Sigma(\bar{x}-\mu)^2}{25}}$

$$= \sqrt{\frac{135}{25}} = 2.3238$$

$$\therefore \sigma \neq \sigma_{\bar{x}}$$

Prob

3. A Population consists of 5, 10, 14, 18, 13, 24 consider all possible samples of size 2 which can be drawn without replacement from this population then find
① Population mean ② S.D of Population ③ The mean of sampling distribution of means ④ S.D of sampling distribution of means.

sol

$N = 6, \ n = 2$

① Population mean $= \mu = \frac{5+10+14+18+13+24}{6} = 14$.

② $\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} = \frac{(5-14)^2+(10-14)^2+(14-14)^2+(18-14)^2+(13-14)^2+(24-14)^2}{6}$

$= 5.9722$

③ In case of without replacement method the total possible samples $= NCn = 6C_2 = 15$

③ Samples :

(5, 10) : (5, 14) (5, 18) (5, 13) (5, 24)

(10, 14) (10, 18) (10, 13) ~~(10, 24)~~ (10, 24)

(14, 18) (14, 13) (14, 24)

(18, 13) (18, 24)

(13, 24)

sample means:

7.5	9.5	11.5	9	14.5
12	14	11.5	17	
16	13.5	19		
15.5	21			
18.5				

$x = 210$

$(\bar{x}-\mu)^2 =$

42.25	20.25	6.25	25	0.25
4	0	6.25	9	
4	0.25	25		
2.25	49			
20.25				

$= 214$

③ The mean of sampling distribution of mean $\mu\bar{x} = \frac{\Sigma\bar{x}}{15} = \frac{210}{15}$

$= 14$

$\mu x = \mu$

④ The S.D of sampling distribution of mean $\sigma\bar{x} = \sqrt{\frac{\Sigma(\bar{x}-\mu)^2}{15}}$

$= \sqrt{\frac{214}{15}} = 3.7771$

$\therefore \sigma \neq \sigma\bar{x}$

4. If the population is 3, 6, 9, 15, 27 list all possible samples of size 3 that can be taken without replacement from this population. find each ① Mean of population ② S.D of population ③ The mean of a sampling distribution of means ④ S.D of sampling distribution of means

Sol

(a) $N = 5, \quad n = 3$

① Population mean $= \mu = \dfrac{3+6+9+15+27}{5} = 12$

② $\sigma = \sqrt{\dfrac{\Sigma(x-\mu)^2}{N}} = \sqrt{\dfrac{(3-12)^2+(6-12)^2+(9-12)^2+(15-12)^2+(27-12)^2}{5}} = 8.48583$

In case of without replacement the total possible samples

$= {}^{N}C_{n} = {}^{5}C_{3} = \dfrac{5!}{(5-3)!\,3!} = 10$

samples: (3, 6, 9) (3, 6, 15) (3, 6, 27) (3, 9, 15) (3, 9, 27) (3, 15, 27)

(6, 9, 15) (6, 9, 27) (6, 15, 27)

(9, 15, 27)

~~$(\bar{x}-\mu)^2$~~

$\bar{x}$ = 6 8 12 9 13 15

10 14 16 = 120

17

$(\bar{x}-\mu)^2$ = 36 16 0 9 1 9

4 4 16 = 120

25

③ $\mu_{\bar{x}} = \dfrac{\Sigma\bar{x}}{10} = \dfrac{120}{10} = 12$

$\mu_{x} = \mu$

(4) $\sigma \bar{x} = \sqrt{\frac{\Sigma(\bar{x}-\mu)^2}{10}} = \sqrt{\frac{120}{10}} = 3.4641$

$\sigma \neq \sigma x$

5. what is the effect on standard error on mean of the sample size is taken from an infinite population of sample size is increased from 400 - 900

$n_1 = 400$; $n_2 = 900$

w.k.T

$S.E(\bar{x}) = \sigma / \sqrt{n}$

$S.E(\bar{x})_{n_1} = \sigma / \sqrt{n_1} = \frac{\sigma}{\sqrt{400}} = \frac{\sigma}{20}$

$S.E(\bar{x})]_{n_2} = \frac{\sigma}{\sqrt{n_2}} = \frac{\sigma}{\sqrt{900}} = \frac{\sigma}{30}$

$S.E(\bar{x})]_{n_1} = \frac{\sigma}{20} = \frac{\sigma}{30} \times \frac{3}{2}$

$= \frac{3}{2} (S.E(\bar{x}) n_2$

$S.E(\bar{x})]_{n_1} = \frac{3}{2} \times S.E(\bar{x})]_{n_2}$

6. when the sample is taken from an infinite population what happed to the standard error of mean if the sample size is decreased from 800 to 200

$n_1 = 800$, $n_2 = 200$

$S.E(\bar{x}) = \frac{\sigma}{\sqrt{n}}$

$$S.E(\bar{x})]_{n_1} = \frac{\sigma}{\sqrt{800}} = \frac{\sigma}{20\sqrt{2}}$$

$$S.E(\bar{x})]_{n_2} = \frac{\sigma}{\sqrt{200}} = \frac{\sigma}{10\sqrt{2}}$$

$$= 2 \cdot \frac{\sigma}{20\sqrt{2}} = 2, S.E(\bar{x})]_{n_1}$$

$$S.E(\bar{x})_{n_2} = 2\, S.E(\bar{x})_{n_1}$$

Estimation:-

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

Estimator:- The processure to determine an unknown population parameter is called estimator

Types of estimation:- Basically there are two kinds of estimates to determine the statistic of polulation parameters namely point estimation & interval estimation

Point estimation:- If an estimate of population parameter is given by a single value then the estimate is called a point estimation

Properties of good estimator:- A good estimator is 'y' which is closed to the true value of parameter, the imp properties of a good estimator are unbiased ness, consistancy and efficiency

Unbiased estimates:- A statistic is said to be an unbiased estimator of the corresponding parameter if the mean of sampling distribution of the statistic is equal to the corresponding population parameter otherwise the statistic is called a biased estimator of the corresponding parameter if t is a statistic and θ be the corresponding parameter and if $E(t) = \theta \Rightarrow E[\text{statistic}] = \text{Parameter}$ the 't' is an unbiased estimator of θ.

Consistancy estimator:- An estimator t of a parameter θ is said to be consistant if it is satifies the following 2 conditions

$$E(t) = \theta \quad \& \quad V(t) \to 0 \text{ as } n \to \infty$$

Efficient estimator:- A statistic t_1 is said to be more efficient unbiased estimator of the parameter θ than the statistic 't_2' if it satisfies the following 2 conditions.

① t_1 & t_2 are both unbiased estimators of θ

② $V(t_1) < V(t_2)$

Max error of an estimate for large samples

Max error of an estimate of the population mean in case of large sample is given by

$$\text{Max error } (ME) = Z(\alpha/2)\left[\sigma/\sqrt{n}\right]$$

In case of small samples the max error of an estimate is given by

$$\text{Max error } (ME) = t(\alpha/2)\,(s/\sqrt{n})$$

3. Here $z(\alpha/2)$ is the table values of z from the normal tables.

n = size of samples

4. σ = Population s.d

$t(\alpha/2)$ is the table values of t ~~from~~ distribution critical values

5.

s is s.D of samples

* Max error of Proportion is $M.E = z(\alpha/2) \cdot \sqrt{\frac{PQ}{n}}$

where $Q = 1 - P$.

Here P is Proportion of success of an event

* In case of small samples the max error of Population mean is given by $M.E = t(\alpha/2)\, s/\sqrt{n}$

* Interval estimation - An estimate of a population parameter is given by two diff values b/w which the parameter may be considered to lies then the estimate is called an interval estimate of the parameter.

For large samples the interval estimation of Population mean is given by

$$\mu = \left[\bar{x} - z(\alpha/2)\, \sigma/\sqrt{n} \; ; \; \bar{x} + z(\alpha/2)\, \sigma/\sqrt{n}\right]$$

The interval estimation of difference of two Population mean is given by

$$\left[(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \; ; \; (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

The interval estimation of population proportion (P) is given by

$$P = \left[p - z(\alpha/2)\sqrt{\frac{PQ}{n}} \; ; \; p + z_{\alpha/2}\sqrt{\frac{PQ}{n}}\right]$$

lower limit   upper limit

The interval estimation of diff of two population proportions $P_1 - P_2$

$$P_1 - P_2 = \left[(p_1 - p_2) - z_{\alpha/2}\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \; ; \; (p_1 - p_2) + z_{\alpha/2}\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right]$$

where $P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$, $Q = 1 - P$

where $\bar{x}$ = sample mean

σ = population standard deviation

$z_{\alpha/2}$ = The table value of 'z' at α% level from the normal tables

p = sample proportion

P = population proportion

n = large sample size n_1, n_2 are also large sample sizes

In case of small samples the interval estimation of population mean is given by

$$\mu = \left(\bar{x} - t(\alpha/2)\, s/\sqrt{n} \; ; \; \bar{x} + t(\alpha/2)\, s/\sqrt{n}\right)$$

The interval estimation of difference of two population means

$$\mu_1 - \mu_2 = \left((\bar{x}_1 - \bar{x}_2) - t(\alpha/2)\, s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \; ; \; (\bar{x}_1 - \bar{x}_2) + t(\alpha/2)\, s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$$

where $\bar{x}$ is the sample mean

s is the sample s.d

3 $t_{\alpha/2}$ = The table value of t at α% level from the t table values where the Population S.D can be estimated through the samples by using following formula

4

$$\boxed{S^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2} \quad (\text{or}) \quad \frac{\Sigma(x_{1i}-\bar{x}_1)^2 + \Sigma(x_{2i}-\bar{x}_2)^2}{(n_1+n_2)-2}}$$

5 Some Results on estimation :-

1. S.T the sample mean $\bar{x}$ is an unbiased estimator of the Population mean μ

Proof :- Let $x_1, x_2, \dots, x_n$ be the samples of size n drawn from the Population with mean μ and variance σ^2

i.e expectation of x_1 is μ & $E(x_1) = \mu$; $E(x_2) = \mu$; $\dots E(x_n) = \mu$

i.e $E(x) = \text{Avg} = \text{mean}$

$V(x_1) = \sigma^2$; $V(x_2) = \sigma^2$, $\dots V(x_n) = \sigma^2$

$E(\text{statistic}) = E(\text{sample mean})$

$= E(\bar{x})$

$= E\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right)$

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

$= \frac{1}{n}\left[\mu + \mu + \dots \mu (n \text{ times})\right]$

$= \frac{1}{\cancel{n}} \cdot \cancel{n}\mu = \mu$ = Population mean

$\therefore$ sample mean $\bar{x}$ is an unbiased estimator of the Population mean μ

2. S.T for a random sample of size n taken from an infinite population

$s^2 = \frac{1}{n}\Sigma(x_i - \bar{x})^2$ is not an unbiased estimator of a population variance σ^2 but $S^2 = \frac{1}{n-1}\Sigma(x_i-\bar{x})^2$ is an unbiased estimator of population variance σ^2

Sol: Proof: Let $x_1, x_2 \ldots x_n$ be the ~~with~~ samples of size n drawn from the population with mean μ variance σ^2

i.e $E(x_1) = \mu$; $E(x_2) = \mu$; ..... $E(x_n) = \mu$

i.e $E(x) = $ Avg = Mean

$V(x_1) = \sigma^2$; $V(x_2) = \sigma^2$, ... $V(x_n) = \sigma^2$

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

$= \frac{1}{n}[\Sigma(x^2 + \bar{x}^2 - 2x_i\bar{x})]$

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

$= \frac{1}{n}[\Sigma x_i^2 + n\bar{x}^2 - 2\bar{x}.n\bar{x}]$ $\quad \therefore \bar{x} = \frac{\Sigma x_i}{n}$

$= \frac{1}{n}[\Sigma x_i^2 + n\bar{x}^2 - 2n\bar{x}^2]$

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

$= \frac{1}{n}\Sigma x_i^2 - \bar{x}^2$

$E(s^2) = E\left(\frac{\Sigma x_i^2}{n}\right) - (E(\bar{x}^2))$

$= \frac{1}{n} \cdot E[\Sigma x_i^2] - E(\bar{x}^2)$

Now W.K.T

$V(x) = \frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$

3 $E(x_i^2) - [E(x)]^2$

$\sigma^2 = E[x_i^2] - \mu^2$

$E(x_i^2) = \sigma^2 + \mu^2$ — ①

4

$V(\bar{x}) = E(\bar{x}^2) - [E(\bar{x})]^2$

5 $\frac{\sigma^2}{n} = E(\bar{x}^2) - \mu^2$

$E(\bar{x}^2) = \frac{\sigma^2}{n} - \mu^2$ — ②

now substitute ① & ② in,

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

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

$= E(x_i^2) - E(\bar{x}^2)$

$= \sigma^2 + \cancel{\mu^2} - \frac{\sigma^2}{n} - \cancel{\mu^2}$

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

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

$\neq \sigma^2$

$\therefore s^2$ is not an unbiased estimator of σ^2

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

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

$= \frac{n}{n-1} \cdot s^2$

$E(S^2) = \frac{n}{n-1} \cdot E(s^2)$

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

s^2 is an unbiased estimator of σ^2

Note: In case of large samples the table of z i.e $z(\alpha/2)$ at (95% level is 1.96) & at (99% level is 2.58) & at (90% level is 1.645)

Prob 1. A random sample of size 100 has a S.D of 5 what can you say about the max error with 95% confidence & also construct the confidence limits with the sample mean is 23

Sol: Given that n = 100 which is greater than or equal to 30 the prob is related to large sample prob.

$n = 100$

$S.D = \sigma = 5$

at 95% confidence $z(\alpha/2) = 1.96$

$z_{\alpha/2} = 1.96$

$\alpha = 95\%$

$= 0.95$

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

for large sample

M.E of Actual mean $\mu = z(\alpha/2) \frac{\sigma}{\sqrt{n}}$

$= 1.96 \times \frac{5}{\sqrt{100}} = 0.98$

confidence in 95% for $\mu = \left(\bar{x} - z_{\alpha/2}(\sigma/\sqrt{n}) ; \bar{x} + z_{\alpha/2}(\sigma/\sqrt{n})\right)$

$= 23 - 0.98 = 22.02$

$= 23 + 0.98 = 23.98$

8. What is the max error one can accept to make with the probability 0.90 when using the mean of a random sample of size 64 to estimate the mean of population with variance 2.56 and also construct confidence limits for μ when the sample mean is 32.

Sol

$n = 64$, $\alpha = 90$

$\therefore$ It is large sample.

Given $\sigma^2 = 2.56$

$SD(\sigma) = \sqrt{\sigma^2} = \sqrt{2.56} = 1.6$

$\bar{x} = 32$

$\therefore$ The table value of z at 90% level is $Z_{\alpha/2}$ is 1.645

$\left[\begin{array}{l} \alpha = 0.90 \\ \frac{\alpha}{2} = 0.45 \\ \text{from the table} \end{array}\right]$

M.E of A.M $\mu = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$= 1.645 \times \frac{1.6}{\sqrt{64}} = 0.329$

90% confidence limits for $\mu = \left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \,;\, \bar{x} + Z_{\alpha/2}(\sigma/\sqrt{n})\right)$

$= 31.671 \,;\, 32.329$

9. A sample size of 100 was taken whose variance is 2254 mean is 54 find the max error to estimate the population mean and also construct 95% confidence interval for the mean.

$n = 300$

$\sigma^2 = 225$, $\sigma = 15$

$z(\alpha/2) = 1.96$

$\bar{x} = 54$

$M.E$ of $A.M$ dl $= z(\alpha/2) \cdot \frac{\sigma}{\sqrt{n}}$

$= 1.96 \times \frac{15}{\sqrt{300}} = 1.6974$

95% confidence limits for $\mu = \left(\bar{x} - z(\alpha/2) \cdot \frac{\sigma}{\sqrt{n}} ; \bar{x} + z(\alpha/2)\left(\frac{\sigma}{\sqrt{n}}\right)\right)$

$= (52.3026 ; 55.6974)$

4. It is decided to estimate the mean no. of hrs to continue until a certain computer will first require repair if it can be assumed that $\sigma = 48$hrs how larger the sample will be needed so that one will be able to assert 90% confidence the sample mean of almost 10hrs

Sol Given $\sigma = 48$

$z(\alpha/2) = 1.645$

Max error $= 10$

$M.E = z(\alpha/2) \cdot \frac{\sigma}{\sqrt{n}}$

$\sqrt{n} = \frac{z(\alpha/2) \cdot \sigma}{M.E}$ $\Rightarrow n = \left(\frac{z(\alpha/2) \cdot \sigma}{M.E}\right)^2 = \left[\frac{1.65 \times 48}{10}\right]^2 = 62.7268$

$= 62$

3 5. In a study of an automobile insurance a random sample of 80 body repairs cost had a mean of 472.36 rs & S.D of 62.35rs. If $\bar{x}$ is used as a Point estimate to the true avg repair cost with what confidence we can assert the max error do not exceed 10rs

4

Sol
Given $n=80$

$\sigma = 62.35$

5 $\bar{x} = 472.36$

$M.E = 10$

$$M.E = Z(\alpha/2)\frac{\sigma}{\sqrt{n}}$$

$$Z(\alpha/2)\cdot\sigma = \sqrt{n}\cdot M.E, \Rightarrow Z(\alpha/2) = \frac{\sqrt{n}\times M.E}{\sigma} = \frac{\sqrt{80}\times 10}{62.35} = 1.4345$$

$$Z(\alpha/2) = 1.44$$

$(\alpha/2) = 0.425$ from table

$\alpha = 0.8502$

$\therefore$ 85% confidence to estimate the population mean

Prob
1. A company wated to introduce a new plan of work and a survey was conducted for this purpose out of 500 workers in 1 group 62% of favoured the new plan and another group of sample of 400 workers 41% were against new plan construct 95% confidence limits for the difference of two proportions

Sol
Given $n_1 = 500$

$n_2 = 400$

P_1 = Proportion of favoured new plan

$= 0.62$

P_2 = Proportion of favoured of new plan in 2nd group

$= 1 - 0.41 = 0.59$

Population proportion $P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$

$= \frac{500(0.62) + 400(0.59)}{900} = 0.6067$

$Q = 1 - P = 1 - 0.6067 = 0.3933$

At 95% level $Z(\alpha/2) = 1.96$

95% Confidence limits to difference of two proportions

$$\left[(P_1 - P_2) - Z(\alpha/2)\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} ; (P_1 - P_2) + Z(\alpha/2)\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right]$$

$$\left[(0.62 - 0.59) - 1.96\sqrt{0.2386\left(\frac{1}{500} + \frac{1}{400}\right)} ; + \right]$$

$$\left[-0.0342 ; 0.0942\right]$$

2. A machine produced 20 defective articles in a batch of 400 after over halting it produced 10 defectives in a batch of 300 construct 99% confidence limits for the difference of 2 proportions

Sol $n_1 = 400$ $\quad P_1 = \frac{x_1}{n_1} = 0.05$

$n_2 = 300$

$x_1 = 20$ $\quad P_2 = \frac{x_2}{n_2} = 0.033$

$x_2 = 10$

at 99% concentration $Z(\alpha/2)$ is 2.58

3 $$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{400(0.05) + 300(0.0333)}{700} = 0.0427$$

$$Q = 1 - P = 1 - 0.0428 = 0.9572$$

4

~~At 99%~~

$$\left[(P_1 - P_2) - z_{(\alpha/2)} \sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \; ; \; + \right]$$

5 $$\left[(0.05 - 0.0333) - 2.58\sqrt{0.0409\left(\frac{1}{400} + \frac{1}{300}\right)} \; ; \; + \right]$$

$$\left[-0.0232 \; ; \; 0.0566 \right]$$

~~-0.0934~~ ~~0.1679~~

3. A sample of size 10 was taken from a population S.D of 0.03 find the max error with 99% confidence and also construct 99% limits for true avg with sample mean 20

Sol

$n = 10$ (small sample)

$\bar{x} = 20$

$S.D = \sigma = 0.03$

$1 - \alpha = 99\% = 0.99$

$\alpha = 0.01$

$\frac{\alpha}{2} = 0.005$

$t_{(\alpha/2)}^{(n-1)} = t_{(0.05)}^{(9)} = 1.833$

$M.E = t_{(\alpha/2)} \left[s/\sqrt{n} \right] = 0.0174$

99% confidence interval $= \left[\bar{x} - t_{(\alpha/2)} (s/\sqrt{n}) \; ; \; \bar{x} + t_{(\alpha/2)} (s/\sqrt{n}) \right]$

$$\left[20 - 1.833\left(\frac{0.03}{\sqrt{10}}\right) ; +\right]$$

$$(19.9826 \quad ; \quad 20.0174)$$

Prob

1. A random sample from a company very extensive files shows that the orders for a certain kind of missionary were filled with respectively in 10, 12, 19, 14, 15, 18, 11, 13 days. Use 1% level find max error and confidence limits for true avg

sol

$n = 8$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{10+12+19+14+15+18+11+13}{8} = 14$$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(10-14)^2 + (12-14)^2 + \cdots + (15-14)^2 + (18-14)^2 + (11-14)^2 + (13-14)^2}{7}}$$

$= 3.2071$

$\alpha = 0.01 \quad , \frac{\alpha}{2} = 0.005$

$$t^{(n-1)}_{(\alpha/2)} = t^{(7)}_{(0.005)} = 3.499 \text{(From table)}$$

$$M.E = t_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}} = 3.9674$$

$$1\% \text{ confidence limits} = \left[\bar{x} - t_{(\alpha/2)}\frac{s}{\sqrt{n}} ; \bar{x} + t_{(\alpha/2)}\frac{s}{\sqrt{n}}\right]$$

$$= 10.0326 \quad ; \quad 17.9674$$

2. A random sample of 10 students have the following IQ :

3 70, 120, 110, 101, 88, 83, 95, 98, 107, 100 find the max error at 5% level and also construct confidence interval for true avg

Sol

4 $n = 10$

$\bar{x} = \frac{\Sigma x}{n} = 97.2$

5 $s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$

$$= \sqrt{\frac{(70-97.2)^2 + (120-97.2)^2 + (110-97.2)^2 + (101-97.2)^2 + (88-97.2)^2 + (83-97.2)^2 + (95-97.2)^2 + (98-97.2)^2 + (107-97.2)^2 + (100-97.2)^2}{9}}$$

$= 14.2735$

$(\cancel{13.8413 + 3.4258} = (17.3271))$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$t^{(n-1)}_{(\alpha/2)} \quad t^{9}_{0.025} = 2.262$

$M.E = t_{(\alpha/2)} \cdot \frac{s}{\sqrt{n}} = 10.2099$

5% confidence limits $= (\bar{x} - t_{(\alpha/2)}\, s/\sqrt{n} \;;\; +)$

$= (86.9901 \;;\; 107.4099)$

3. Two independent samples of 8 and 7 items respectively had the following values find the confidence limits at 5% level for the diff of two mugs

(x) sample 1: 11 11 13 11 15 9 12 14 $\bar{x} = 12$

(y) sample 2: 9 11 10 13 9 8 10 $\bar{y} = 10$

$n_1 = 8$, $n_2 = 7$

$(x-\bar{x})^2$: 1 1 1 1 9 9 0 4 = 26

$(y-\bar{y})^2$: 1 1 0 9 1 4 0 = 16

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2}{n_1+n_2-2}} = 1.7974$$

$$t_{\alpha/2}^{(n_1+n_2-2)} = t_{0.025}^{(13)} = 2.160$$

5% level confidence limits = $\left[(\bar{x}-\bar{y}) - t_{\alpha/2}\, S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} ;\ (\bar{x}-\bar{y}) + t_{\alpha/2}\, S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]$

$= [-0.0093,\ 4.0093]$