

UNIT - 2

Sampling Distributions and Estimations

> Population :

- Sampling distribution of the mean (σ known & unknown)
- Sums & differences
- Central limit theorem
- Estimation
- point Estimation & Interval Estimation concerning means for large samples.

proportion

Sample proportion.

N	n	Size
μ	$\bar{x}$	mean
σ	s	S.D
p	p	proportion
Population	Sample	

> Population:

population is a set (or) collection (or) Totality of the objects under study.

> Population Size:

The number of objects or observations in the population is known as Population Size

- It is denoted by "N"

Note:

- Population is said to be Finite or Infinite depending on the size N is finite or Infinite.

> Sample:

A Finite subset of the population known as Sample

- Sample size is denoted by "n"

> Sampling:

Sampling is the process of drawing all the possible samples from a given population.

Ex 1: Population of India (population),
Population of T.S (sample)

Ex 2: Engineering colleges recognized by AICTE (population),
Engineering colleges affiliated to JNTUH (sample).

> Large Sampling:

If $n \geq 30$ then the sampling is said to be Large Sampling.

> Small Sampling:

If $n < 30$ then the sampling is said to be Small Sampling or Exact Sampling.

> Parameters:

Statistical measures or constants obtained from the population are known as population parameters (or) simply parameters.

Ex: Population mean (μ)

Population standard deviation (σ)

population proportion (P), etc.

> Statistics:

Statistical measures or constants obtained from the sample are known as sample statistics or simply statistics.

Ex: Sample mean ($\bar{x}$)

Sample standard deviation (s)

Sample proportion (p)

Note:

The word parameter refers to population and the word Statistics refers to sample

> Random Sampling:

Each member of the population has the equal chances (or) probability of being included in the sample.

> Sampling with Replacement.

Each member of the population may be chosen more than once is known as sampling with replacement.

> Sampling without Replacement:

An element of the population cannot be chosen more than once is known as sampling without replacement.

Note 1 :

For finite population (or) sampling is done without replacement. The number of samples of size n drawn from a finite population of N .

$$N = N C_n$$

Note 2:

For an infinite population (or) sampling is done with replacement. The number of samples of size n drawn from infinite population of N

$$N = N^n$$

> Population Mean:

If $x_1, x_2, x_3, \dots, x_n$ be a population size N , then population mean

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N} = \frac{\sum_{i=1}^n x_i}{N}$$

> Population variance :

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

Note : Population standard deviation

$$S.D = \sigma$$

> Sample Mean:

If $x_1, x_2, \dots, x_n$ be a sample of size ' n ', then the sample mean is defined as

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

> Sample variance:

Sample variance is given by "s"

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

Note:

Sample standard deviation
S.D = s.

Sampling Distribution

The probability distribution of a statistics ($\bar{x}$ or s or p) is called Sampling Distribution.

Sampling Distribution of means (σ - Known)

The probability Distribution of sample mean ($\bar{x}$) is known as Sampling Distribution means.

The Sample

★ Finite population (or) Sampling is done without Replacement

Consider a Finite population of size (N) with mean (μ) and standard deviation (σ) draw all the possible samples of size n without replacement from this finite population.

Let $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$ denotes the Mean and standard deviation of Sampling distribution of Means.

Therefore,

$$\begin{aligned} \mu_{\bar{x}} &= \mu \\ \sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \end{aligned}$$

★ Here, $\frac{N-n}{N-1}$ is called population Correction Factor.

Infinite population (or) Sampling is done with Replacement.

Suppose the samples are drawn from an Infinite population (or) Sampling is done with Replacement, then the mean and S.D of sampling distribution of means is given by,

$$\begin{aligned}\mu_{\bar{x}} &= \mu \\ \sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}}\end{aligned}$$

> Standard Error of Mean

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ is known as Standard Error of Mean

Imp.
Problem

- i) A population consists of 4 members 3, 7, 11, 15. Consider all possible samples of size 2 which can be drawn with Replacement and without Replacement from population. Find
- the population mean
 - the population S.D
 - the mean of sampling distribution of Means ($\mu_{\bar{x}}$)
 - S.D of sampling distribution of Means ($\sigma_{\bar{x}}$)

Verify the results.

sol. Given population is 3, 7, 11, 15

Population size $N = 4$

Population

i) Population mean $\mu = \frac{3 + 7 + 11 + 15}{4}$

$$\mu = 9$$

The population Mean is 9

ii) Population variance $\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$

population x	$x_i - \mu$ $x_i - 9$	$(x_i - \mu)^2$
3	-6	36
7	-2	4
11	2	4
15	6	36
		$\sum (x_i - \mu)^2$ $= 80$

population variance

$$\sigma^2 = \frac{80}{4} = 20$$

The population S.D

$$\sigma = \sqrt{20}$$

For Samples,

I. Sampling with replacement

No. of samples of size $n = 2$ (given) from a population of size $N = 4$ with replacement $= N^n = 4^2 = 16$

i.e., there are 16 samples of size 2 with Replacement.

S.NoSamplesSample Mean

1	3,3	3
2	3,7	5
3	3,11	7
4	3,15	9
5	7,3	5
6	7,7	7
7	7,11	9
8	7,15	11
9	11,3	7
10	11,7	9
11	11,11	11
12	11,15	13
13	15,3	9
14	15,7	11
15	15,11	13
16	15,15	15

Sampling distribution of mean with replacement.

Sample mean ($\bar{x}$)	3	5	7	9	11	13	15	total
probability $P(x = \bar{x})$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	1

iii) Mean of sampling distribution of mean's

$$\begin{aligned} \mu_{\bar{x}} &= E(\bar{x}) = \sum \bar{x}_i \cdot p(x = \bar{x}_i) \\ &= (3 \times \frac{1}{16}) + (5 \times \frac{2}{16}) + \dots + (15 \times \frac{1}{16}) \end{aligned}$$

$$\mu_{\bar{x}} = E(\bar{x}) = 9$$

iv) Variance of sampling distribution of mean's

$$\begin{aligned} \sigma_{\bar{x}}^2 &= E(\bar{x}^2) - [E(\bar{x})]^2 \\ &= \sum \bar{x}_i^2 \cdot p(x = \bar{x}_i) - (9)^2 \\ &= (3^2 \times \frac{1}{16}) + (5^2 \times \frac{2}{16}) + \dots + (15^2 \times \frac{1}{16}) - 81 \end{aligned}$$

$$\sigma_{\bar{x}}^2 = 10$$

$$\sigma_{\bar{x}} = \sqrt{10}$$

Sampling without Replacement

No. of samples of size $n=2$ from the population 3, 7, 11, 15 of size $N=4$, without replacement

$$Nc_n = 4C_2 = \frac{4 \times 3}{1 \times 2} = 6$$

S.No	Samples	Sample mean $\bar{x}$
1	3, 7	$\frac{3+7}{2} = 5$
2	3, 11	7
3	3, 15	9
4	7, 11	9
5	7, 15	11
6	11, 15	13

Sampling distribution of mean without replacement.

Sample mean ($\bar{x}$)	5	7	9	11	13	total
probability $P(\bar{x} = \bar{x}_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

mean of sampling distribution of mean's

$$\begin{aligned} \mu_{\bar{x}} &= E(\bar{x}) = \sum \bar{x}_i \cdot p(\bar{x} = \bar{x}_i) \\ &= (5 \times \frac{1}{6}) + (7 \times \frac{1}{6}) + \dots + (13 \times \frac{1}{6}) \end{aligned}$$

$$\boxed{\mu_{\bar{x}} = E(\bar{x}) = 9}$$

variance of sampling distribution of means

$$\begin{aligned} \sigma_{\bar{x}}^2 &= E(\bar{x}^2) - [E(\bar{x})]^2 \\ &= \sum \bar{x}_i^2 \cdot p(\bar{x} = \bar{x}_i) - (9)^2 \\ &= (5^2 \times \frac{1}{6}) + (7^2 \times \frac{1}{6}) + \dots + (13^2 \times \frac{1}{6}) - 81 \end{aligned}$$

$$\sigma_{\bar{x}}^2 = \frac{20}{3}$$

$$\boxed{\sigma_{\bar{x}} = \sqrt{\frac{20}{3}}}$$

Verification with Replacement.

$$i) \mu_{\bar{x}} = 9 = \mu$$

$$ii) \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{20}}{\sqrt{2}} = \sqrt{\frac{20}{2}} = \sqrt{10} = \sigma_{\bar{x}}$$

$$\boxed{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$$

without replacement

$$i) \mu_{\bar{x}} = 9 = \mu$$

$$ii) \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \Rightarrow \frac{\sqrt{20}}{\sqrt{2}} \sqrt{\frac{4-2}{4-1}} \Rightarrow \frac{\sqrt{20}}{\sqrt{2}} \frac{\sqrt{2}}{\sqrt{3}} \Rightarrow \sqrt{\frac{20}{3}} = \sigma_{\bar{x}}$$

$$\boxed{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}}$$

A population consists of 4 members 2, 3, 4, 5. Consider all possible samples of size 2 which can be drawn with replacement & without replacement from population. Find.

- i) The population mean
- ii) The population S.D
- iii) The mean of sampling distribution of means
- iv) S.D of sampling distribution of means
- v) Verify the results.

Sol. For population

Given population 2, 3, 4, 5, Here $N=4$

i) Population mean $\mu = \frac{2+3+4+5}{4} = 10.25$

$$\boxed{\mu = 3.5}$$

ii) Population Variance.

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

$$= \frac{(2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2}{4}$$

$$\sigma^2 = \frac{5}{4}$$

$$\boxed{\sigma = \frac{\sqrt{5}}{2}}$$

iii)

iv)

v)

4) Find the value of population correction factor for 16/10/2

i) $n = 5, N = 200$

ii) $n = 10, N = 100$

Sol Population correction factor $= \frac{N-n}{N-1}$

i) $n = 5, N = 200$

Population correction factor $= \frac{200-5}{200-1} \Rightarrow \frac{195}{199}$

ii) $n = 10, N = 100$

Population correction factor $= \frac{100-10}{100-1} = \frac{90}{99}$

5) How many different samples can be chosen from a finite population of size 25? How many different samples of size 2 can be chosen?

Sol No. of samples of size n can be drawn from a population sample size N

i) With replacement $= N^n$

ii) Without replacement $= {}^N C_n$

$n = 2, N = 25$

i) With replacement $= 25^2 = 625$

ii) Without replacement $= {}^{25} C_2 = \frac{25 \times 24}{1 \times 2} = 300$

6) The Variance of a population is 2. The size of sample collected from the population is 169. What is the standard error of mean.

Sol. Given, $\sigma^2 = 2 \Rightarrow \sigma = \sqrt{2}$

$$n = 169$$

$$\text{Standard error of mean} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{2}}{\sqrt{169}} = \frac{\sqrt{2}}{13}$$

7) What is the effect on standard error, if a sample is taken from an infinite population. of sample size is increased from 400 to 900.

Sol.

Standard Error of Mean

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{---} \star$$

$$\text{When } n = 400, S.E_1 = \frac{\sigma}{\sqrt{400}} = \frac{\sigma}{20} \quad \text{---} \textcircled{1}$$

$$\text{When } n = 900, S.E_2 = \frac{\sigma}{\sqrt{900}} = \frac{\sigma}{30}$$

$$= \frac{2}{3} \left(\frac{\sigma}{20} \right)$$

$$= \frac{2}{3} (S.E_1) \quad \text{---} \text{(By } \textcircled{1} \text{)}$$

$$S.E_2 = \frac{S.E_1}{\frac{3}{2}}$$

Hence, The standard error is decreased by $\frac{3}{2}$

8) When a sample is taken from an infinite population. What happened to the SE of mean, with the sample size is decreased from 800 to 200.

Sol.

Standard Error of Mean = $\frac{\sigma}{\sqrt{n}}$ --- $\star$

$$\text{When } n = 800, S.E_1 = \frac{\sigma}{\sqrt{800}} = \frac{\sigma}{20\sqrt{2}} \quad \text{---} \textcircled{1}$$

$$\text{When } n = 200, S.E_2 = \frac{\sigma}{\sqrt{200}} = \frac{\sigma}{10\sqrt{2}}$$

$$= 2 \left(\frac{\sigma}{20\sqrt{2}} \right)$$

$$S.E_2 = 2 \times S.E_1 \quad \text{(By } \textcircled{1} \text{)}$$

Hence, the SE is increased by 2 times

Vimp. Central Limit Theorem $\bar{x} \rightarrow n \rightarrow \mu \rightarrow \sigma^2$

If $(\bar{x})$ is the mean of a Random sample of size (n) drawn from a population with mean (μ) and Variance (σ^2) . Then the Standardized Sample mean z is a random variable whose distribution approaches to that of Standard Normal distribution.

$$z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

- i) The mean height of students in college is 155cm and S.D is 15. What is the probability that the mean height of 36 students is less than 157cm.

Sol. Given,

$$\mu = 155, \sigma = 15, n = 36$$

To find $p(\bar{x} < 157) = ?$

Using Central Limit Theorem,

$$z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

$$z = \frac{157 - 155}{\left(\frac{15}{\sqrt{36}}\right)} \Rightarrow 0.8$$

$$\text{If } p(\bar{x} < 157) = p(z < 0.8)$$

$$= 0.5 + p(z)$$

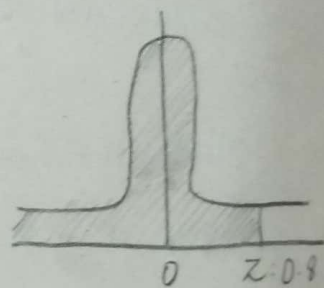
$$= 0.5 + A(0 \text{ to } z)$$

$$= 0.5 + A(0 \text{ to } 0.8)$$

$$= 0.5 + 0.2881$$

[$\because$ from table]

$$p(\bar{x} < 157) = 0.7881$$



x (A Random sample of size 100 is taken from an infinite population having mean $\mu = 76$)

2) A Random sample of size 100 is taken from an infinite population having mean $(\mu) = 76$, and variance $\sigma^2 = 256$. What is the probability that $\bar{x}$ will be between 75 and 78.

Sol. Given $\mu = 76$, $\sigma^2 = 256 \Rightarrow \sigma = 16$
 $n = 100$

To find $P(75 < \bar{x} < 78) = ?$

$$Z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

$$\text{when } \bar{x} = 75, \quad Z = \frac{75 - 76}{\left(\frac{16}{\sqrt{100}}\right)} = -0.625 \approx -0.63$$

$$\text{when } \bar{x} = 78, \quad Z = \frac{78 - 76}{\left(\frac{16}{\sqrt{100}}\right)} = 1.25$$

$$\therefore P(75 < \bar{x} < 78) = P(-0.63 < Z < 1.25)$$

$$= P(Z_1) + P(Z_2)$$

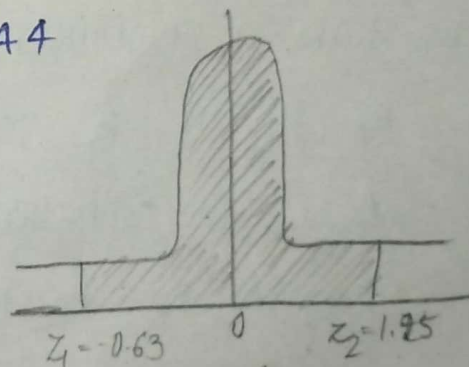
$$= A(0 \text{ to } Z_1) + A(0 \text{ to } Z_2)$$

$$= A(0 \text{ to } -0.63) + A(0 \text{ to } 1.25)$$

$$= A(0 \text{ to } 0.63) + A(0 \text{ to } 1.25)$$

$$= 0.2357 + 0.3944$$

$$= 0.6301.$$



Sampling Distribution of Sums & Differences.

- ⇒ μ_{S_1} and σ_{S_1} is the mean and the standard deviation of sampling distribution of statistics S_1 obtained from all the possible samples of size n_1 drawn from population "A".
- ⇒ Similarly μ_{S_2} and σ_{S_2} is the mean and the standard deviation of sampling distribution of statistics of S_2 obtained from all possible samples of size n_2 drawn from population "B".
- ⇒ The mean μ_{S_1} and $\mu_{S_2} \Rightarrow \mu_{S_1+S_2} = \sigma_{S_1+S_2}$ be the mean & deviation of sampling distribution of sum are given by.

$$i) \mu_{S_1+S_2} = \mu_{S_1} + \mu_{S_2}$$

$$\sigma_{S_1+S_2} = \sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2}$$

- ii) Similarly, sampling distribution of differences are given by,

$$\mu_{S_1-S_2} = \mu_{S_1} - \mu_{S_2}$$

$$\sigma_{S_1-S_2} = \sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2}$$

Problems

- i) The mean voltage of a battery is 15 & S.D is 0.2. Find the probability that 4 such batteries are connected in series. we have a combined voltage of 60.8 (or) more volts.

Sol. $\mu_A, \mu_B, \mu_C, \mu_D$ is the vol^t mean voltage of 4 battery's A, B, C, D connected in series.

$$\text{Given, } \mu_A = \mu_B = \mu_C = \mu_D = 15$$

$$\sigma_A = \sigma_B = \sigma_C = \sigma_D = 0.2$$

$$\therefore \mu = \mu_{A+B+C+D} = \mu_A + \mu_B + \mu_C + \mu_D \Rightarrow 15 + 15 + 15 + 15$$

$$\boxed{\mu = 60}$$

$$\sigma = \sigma_{A+B+C+D} = \sqrt{\sigma_A^2 + \sigma_B^2 + \sigma_C^2 + \sigma_D^2}$$

$$= \sqrt{(0.2)^2 + (0.2)^2 + (0.2)^2 + (0.2)^2}$$

$$= \sqrt{4(0.2)^2} \Rightarrow \boxed{\sigma = 0.4}$$

Let $\bar{x}$ be the mean of combined voltage of 4 batteries.

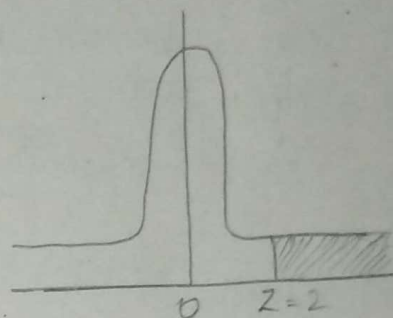
To find $p(\bar{x} \geq 60.8) = ?$

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

$$\text{when } \bar{x} = 60.8, Z = \frac{60.8 - 60}{0.4}$$

$$\boxed{Z = 2}$$

$$\begin{aligned} p(\bar{x} \geq 60.8) &= p(Z > 2) \\ &= 0.5 - p(Z) \\ &= 0.5 - A(0 \text{ to } Z) \\ &= 0.5 - A(0 \text{ to } 2) \\ &= 0.5 - 0.4772 \\ &= 0.0228 \end{aligned}$$



2) The mean lifetime of light bulbs produced by a company is 1500 hours and S.D of 150 hours. Find the probability that

a) atleast 5000h

b) atleast 4200h if there 3 bulbs are connected such that when 1 bulb burns out, another bulb will glow long. Assume that lifetime are normally distributed

Let μ_a, μ_b, μ_c be lighting of bulbs

$$\text{mean : } \mu_a = \mu_b = \mu_c = 1500$$

$$\text{S.D : } \sigma_a = \sigma_b = \sigma_c = 150$$

$$\mu_{a+b+c} = \mu_a + \mu_b + \mu_c$$

$$= 1500 + 1500 + 1500$$

$$\mu_{a+b+c} = 4500$$

$$\begin{aligned}\sigma_{a+b+c} &= \sqrt{(\sigma_a)^2 + (\sigma_b)^2 + (\sigma_c)^2} \\ &= \sqrt{(150)^2 + (150)^2 + (150)^2} \\ &= \sqrt{3(150)^2} \\ &= 259.8\end{aligned}$$

$$\sigma_{a+b+c} = 260$$

Let $\bar{x}$ be the mean of the combined glowing bulb using N.D

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

a) at least 5000h

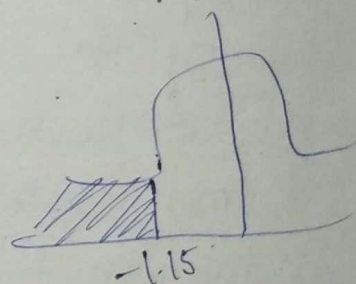
$$p(x > 5000)$$

$$Z = \frac{5000 - 4500}{260} = \frac{500}{260} = 1.92$$

$$\begin{aligned}p(x > 1.92) &= 0.5 - A(0 \text{ to } 1.92) \\ &= 0.5 - 0.4726 \\ &= 0.0274\end{aligned}$$

b) $p(x < 4200)$

$$\begin{aligned}Z &= \frac{4200 - 4500}{260} \\ &= \frac{-300}{260} = -1.15\end{aligned}$$



$$\begin{aligned}p(Z < -1.15) &= 0.5 - p(Z) \\ &= 0.5 - A(0 \text{ to } -1.15) \\ &= 0.5 - A(0 \text{ to } 1.15) \\ &= 0.5 - 0.3749 \\ &= 0.1251\end{aligned}$$

Q3) Let $U_1 = \{2, 7, 9\}$, $U_2 = \{3, 8\}$. Find

a) μ_{U_1}

f) σ_{U_2}

b) μ_{U_2}

g) $\sigma_{U_1+U_2}$

c) $\mu_{U_1+U_2}$

h) $\sigma_{U_1-U_2}$

d) $\mu_{U_1 - U_2}$

i) $\mu_{U_1+U_2} = \mu_{U_1} + \mu_{U_2}$

e) σ_{U_1}

j) $\mu_{U_1-U_2} = \mu_{U_1} - \mu_{U_2}$

Sol. Given,

$$U_1 = \{2, 7, 9\}$$

k) $\sigma_{U_1+U_2} = \sqrt{\sigma_{U_1}^2 + \sigma_{U_2}^2}$

$$U_2 = \{3, 8\}$$

a) $\mu_{U_1} = \frac{2+7+9}{3} = 6$

$$\boxed{\mu_{U_1} = 6}$$

b) $\mu_{U_2} = \frac{3+8}{2} = 5.5$

$$\boxed{\mu_{U_2} = 5.5}$$

c) $\mu_{U_1+U_2} = \{5, 10, 12, 10, 15, 17\}$

$$\mu_{U_1+U_2} = \frac{5+10+12+10+15+17}{6} = 11.5$$

$$\boxed{\mu_{U_1+U_2} = 11.5}$$

d) $U_1 - U_2 = \{-1, 4, 6, -6, -1, 1\}$

$$\mu_{U_1-U_2} = \frac{-1+4+6-6-1+1}{6} = 0.5$$

$$\boxed{\mu_{U_1-U_2} = 0.5}$$

e) Variance of U_1 i.e., $\sigma_{U_1}^2 = \frac{\sum (x_i - \mu_{U_1})^2}{N}$

$$\sigma_{U_1}^2 = \frac{(2-6)^2 + (7-6)^2 + (9-6)^2}{3} = 8.6$$

$$\sigma_{U_1} = 2.93$$

f) Variance of U_2 i.e., $\sigma_{U_2}^2 = \frac{\sum (x_i - \mu_{U_2})^2}{N}$

$$\sigma_{U_2}^2 = \frac{(3-5.5)^2 + (8-5.5)^2}{2} = 6.25$$

$$\sigma_{U_2} = 2.5$$

g) Variance of $U_1 + U_2$ i.e., $\sigma_{U_1+U_2}^2 = \frac{\sum (x_i - \mu_{U_1+U_2})^2}{N}$

$$\sigma_{U_1+U_2}^2 = \frac{(5-11.5)^2 + (10-11.5)^2 + \dots + (17-11.5)^2}{6} = 14.91$$

$$\sigma_{U_1+U_2} = 3.86$$

h) Variance of $U_1 - U_2$ i.e., $\sigma_{U_1-U_2}^2 = \frac{\sum (x_i - \mu_{U_1-U_2})^2}{N}$

$$\sigma_{U_1-U_2}^2 = \frac{(-1-0.5)^2 + (4-0.5)^2 + \dots + (1-0.5)^2}{6} = 14.91$$

$$\sigma_{U_1-U_2} = 3.86$$

i) Verification

a) $\mu_{U_1+U_2} = 11.5 = 6 + 5.5 = \mu_{U_1} + \mu_{U_2}$

$$\mu_{U_1+U_2} = \mu_{U_1} + \mu_{U_2}$$

$$b) \mu_{U_1 - U_2} = 0.5 = 6 - 5.5 = \mu_{U_1} - \mu_{U_2}$$

$$\boxed{\mu_{U_1 - U_2} = \mu_{U_1} - \mu_{U_2}}$$

Estimation

Estimation is a procedure of making judgements about the unknown value of population parameters by using sample observations (results)

Note:

Estimation is 2 types

- 1) Point Estimation
- 2) Interval Estimation

⇒ Point Estimation:

A Point estimates of a population parameter is a single value used to estimate the population parameter.

Ex: The Sample mean $\bar{x}$ is a point estimate of the population mean μ .

⇒ Interval Estimation:

Consider a large sample of size n drawn from a population with mean (μ) and variance (σ^2). Then Large sample confidence interval is given by

$$\bar{x} - z_{\left(\frac{\alpha}{2}\right)} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\left(\frac{\alpha}{2}\right)} \cdot \frac{\sigma}{\sqrt{n}}$$

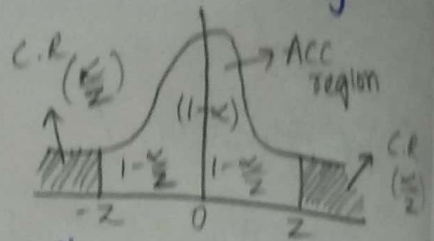
(or)

$$\left(\bar{x} - z_{\left(\frac{\alpha}{2}\right)} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\left(\frac{\alpha}{2}\right)} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

1) Here $Z_{(\frac{\alpha}{2})} \cdot \frac{\sigma}{\sqrt{n}}$ is known as Maximum Error of Estimation

2) A $(1-\alpha)100\%$ confident that the maximum Error of Estimation by

$$E_{\max} = Z_{(\frac{\alpha}{2})} \cdot \frac{\sigma}{\sqrt{n}}$$



3) If α , σ and $E_{\max}$ is given, then the sample size n is

$$n = \left(\frac{Z_{(\frac{\alpha}{2})} \cdot \sigma}{E_{\max}} \right)^2$$

Problems

1) Assuming that $\sigma = 20$. How large a random sample be taken to assert with probability 0.95 that a sample mean will not differ from true mean by more than 0.3.0 points?

Sol. Given that $\sigma = 20$

$$E_{\max} = 3$$

Also given $(1-\alpha)100\% = 95\%$

$$\Rightarrow (1-\alpha) = 0.95$$

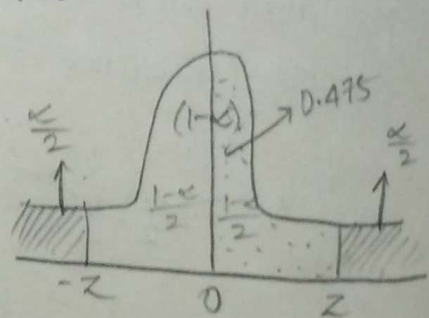
$$\Rightarrow \left(\frac{1-\alpha}{2} \right) = \frac{0.95}{2} = 0.475$$

For the area 0.475,
from table,

$$Z_{(\frac{\alpha}{2})} = 1.96$$

$$n = \left(\frac{Z_{(\frac{\alpha}{2})} \cdot \sigma}{E_{\max}} \right)^2 = \left(\frac{1.96 \times 20}{3} \right)^2$$

$$n = 171$$



Sample size never be
in fraction

- 2) What is the maximum Error one can expect to make with probability 0.90 when using the mean of a Random Sample of size 64 to Estimate the mean of population with $\sigma^2 = 2.56$.

Sol.

Given, $n = 64$

$$\sigma^2 = 2.56$$

$$\rightarrow \sigma = 1.6$$

$$(1 - \alpha) 100\% = 90\%$$

$$\Rightarrow (1 - \alpha) = 0.90$$

$$\Rightarrow \left(\frac{1 - \alpha}{2} \right) = \frac{0.90}{2} = 0.45$$

$$\boxed{Z\left(\frac{\alpha}{2}\right) = 1.65}$$

($\because$ from table)

$$E_{\max} = \frac{Z\left(\frac{\alpha}{2}\right) \sigma}{\sqrt{n}}$$

$$= \frac{1.65 \times 1.6}{\sqrt{64}}$$

$$\boxed{E_{\max} = 0.33}$$

Problems on Interval Estimation

- i) The mean and S.D of a population are 11.795 and 14.054 respectively. What one can assert that 95% Confident about the maximum Error. If $\bar{x} = 11.795$ and $n = 50$. And also construct 95% confidence interval for the true mean.

Sol.

Given, $n = 50$

$$\bar{x} = 11.795$$

$$\sigma = 14.054$$

$$i) (1-\alpha) 100\% = 95\%$$

$$(1-\alpha) = 0.95$$

$$\left(\frac{1-\alpha}{2}\right) = \frac{0.95}{2} = 0.475$$

($\therefore$ from table)

$$\boxed{Z\left(\frac{\alpha}{2}\right) = 1.96}$$

$$E_{\max} = \frac{Z\left(\frac{\alpha}{2}\right) \cdot \sigma}{\sqrt{n}} = \frac{1.96 \times 14054}{\sqrt{50}} = 3895.57$$

$$E_{\max} \approx 3896$$

ii) Confidence interval for True mean

$$= \left(\bar{x} - Z\left(\frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}}, \bar{x} + Z\left(\frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}} \right)$$

$$= (11,795 - 3896, 11,795 + 3896)$$

$$= (7899, 15691)$$

2) A Random Sample of size 81 was taken whose variance is 20.25 and mean is 32. Construct 98% confidence interval.

Sol. Given $n = 81$

$$\bar{x} = 32$$

$$\sigma^2 = 20.25 \Rightarrow \sigma = 4.5$$

$$i) (1-\alpha) 100\% = 98\%$$

$$(1-\alpha) = 0.98$$

$$\left(\frac{1-\alpha}{2}\right) = \frac{0.98}{2} = 0.49$$

$$\boxed{Z\left(\frac{\alpha}{2}\right) = 2.33}$$

$$E_{\max} = \frac{Z(\frac{\alpha}{2}) \cdot \sigma}{\sqrt{n}} = \frac{2.33 \times \frac{4.5}{\sqrt{81}}}{\sqrt{81}} = \frac{2.33 \times 0.5}{9} = 0.12925 \Rightarrow 1.165$$

$$E_{\max} \approx 1.165$$

C. I for true mean,

$$= (\bar{x} - Z(\frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}}, \bar{x} + Z(\frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}})$$

$$= (32 - 5.2425, 32 + 5.2425)$$

$$= (32 - 1.165, 32 + 1.165)$$

$$= (30.835, 33.165)$$

3) Construct a 99% confidence interval for the true mean weight of 36 persons on diet control after one month had a mean weight loss of 3.42 kgs with S.D of 0.68 kgs

Sol.

Given, $n = 36$

$$\bar{x} = 3.42$$

$$\sigma = 0.68$$

$$i) (1 - \alpha) 100\% = 99\%$$

$$(1 - \alpha) = 0.99$$

$$\left(\frac{1 - \alpha}{2}\right) = \frac{0.99}{2} = 0.495$$

$$\boxed{Z(\frac{\alpha}{2}) = 2.58}$$

$$E_{\max} = \frac{Z(\frac{\alpha}{2}) \cdot \sigma}{\sqrt{n}} = \frac{2.58 \times 0.68}{\sqrt{36}}$$

$$= 0.2924$$

C.I for true mean

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

$$= (3.42 - 0.2924, 3.42 + 0.2924)$$

$$= (3.1276, 3.7124)$$

$$\Rightarrow (3.12, 3.71)$$

4) A Random sample of size 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 .

Sol.

Given, $n = 100$

$$\bar{x} = 21.6$$

$$\sigma = 5.1$$

$$i) (1 - \alpha) 100\% = 95\%$$

$$(1 - \alpha) = 0.95$$

$$\left(\frac{1 - \alpha}{2}\right) = \frac{0.95}{2} = 0.475$$

$$\boxed{z_{\left(\frac{\alpha}{2}\right)} = 1.96} \quad (\because \text{from table})$$

$$E_{\max} = \frac{z_{\left(\frac{\alpha}{2}\right)} \sigma}{\sqrt{n}} = \frac{1.96 \times 5.1}{\sqrt{100}} = 0.9996$$

$$\boxed{E_{\max} = 0.9996}$$

C.I for true mean,

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

$$= (21.6 - 0.9996, 21.6 + 0.9996)$$

$$= (20.6, 22.6)$$

Problems

- 3) A Random sample of size 100 has a S.D of 5. What can you say about the maximum error with 95% confidence.

Given that $\sigma = 5$

$$n = 100$$

$$(1 - \alpha) 100\% = 95\%$$

$$(1 - \alpha) = 0.95$$

$$\left(\frac{1 - \alpha}{2}\right) = \frac{0.95}{2} = 0.475$$

From table

$$Z\left(\frac{\alpha}{2}\right) = 1.96$$

$$E_{\max} = \frac{Z\left(\frac{\alpha}{2}\right) \sigma}{\sqrt{n}} \Rightarrow \frac{1.96 \times 5}{\sqrt{100}} \Rightarrow 0.98$$

$$E_{\max} = 0.98$$

- 4) To estimate the average amount of time visitors take to move from one building to another in an office complex, the mean of a random sample of size n is used. Given $\sigma = 1.40$ mins, determine how large n should be the sample size if it is ascertained with 99% confidence that the error E is at most 0.25.

Sol.

Given, $\sigma = 1.40$

$$E_{\max} = 0.25$$

$$(1 - \alpha) 100\% = 99\%$$

$$(1 - \alpha) = 0.99$$

$$\left(\frac{1 - \alpha}{2}\right) = \frac{0.99}{2} = 0.495$$

$$\boxed{Z\left(\frac{\alpha}{2}\right) = 2.58}$$

$$n = \left(\frac{Z\left(\frac{\alpha}{2}\right) \sigma}{E_{\max}} \right)^2 = \left(\frac{2.58 \times \frac{1.40}{0.25}}{0.25} \right)^2$$

$$= 208.744$$

$$\boxed{n = 208}$$

5) Using the mean of a Random sample of size 150 to estimate the mean mechanical aptitude of mechanics of a large workshop and assuming $\sigma = 6.2$. What can we assert with 0.99 probability about the max Size of Error.

Given, $n = 150$

$$\sigma = 6.2$$

$$(1 - \alpha) = 0.99$$

$$\left(\frac{1 - \alpha}{2} \right) = \frac{0.99}{2} = 0.495$$

$$\boxed{Z\left(\frac{\alpha}{2}\right) = 2.58}$$

$$E_{\max} = \frac{Z\left(\frac{\alpha}{2}\right) \sigma}{\sqrt{n}}$$

$$= \frac{2.58 \times 6.2}{\sqrt{150}}$$

$$= 1.3060$$

$$\boxed{E_{\max} = 1.30}$$