PROBABILITY AND STATISTICS

Unit-II

Sampling Distribution & Estimation

04/09/2025

▪ Sampling Distributions

- **Populations and Samples**
- **Sampling distribution of the Mean**
  (σ - known and unknown)
- **Sums and Differences**
- **The Central Limit Theorem**

- **Estimation**

  - Point Estimation

  - Interval Estimation concerning Means for Large Samples.

Population: Population is a set or collection or totality of objects under study.

Population may be animate or inanimate, actual or hypothetical

Mainly population consists a large collection of Individuals or attributes or numerical data which are of interest.

Size of the population: Number of objects or observations in the population is known as size of the population.Size of population is denoted by N.

Population may be infinite or infinite:

Population is said to be finite or infinite depending on the size N being finite or infinite.

A **sample** is a finite subset of the population.

Size of Sample is denoted by n

The sampled population is the population from which the sample is drawn.

Why Sample?

- Selecting a sample is <u>less time-consuming</u> than selecting every item in the population (**census**).

- Selecting a sample is <u>less costly</u> than selecting every item in the population.

Example1: The people of India is a Population, people of Telangana is a sample.

Example2: Engineering colleges affiliated to AICTE is a population, Autonomous Engineering colleges is a sample.

Example3:Under graduate Students in SNIST (Population), Students of a particular branch (sample).

Selecting a Sample

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

If sample size, $n \geq 30$ then Sampling is said to be **Large sampling**.

If sample size, $n<30$ then Sampling is said to be **Small sampling** or exact sampling.

Note: Samples must be representative of population, sampling should be random.

Random sampling: Random sampling is a part of the sampling technique in which each sample has an equal probability of being chosen.

Sampling from a Finite Population

Sampling with Replacement,
Each member of the Population may be chosen more than once, since the member is replaced in the population.

Thus sampling from **finite population With replacement can be considered** theoretically as a sampling from **infinite population.**

Sampling without replacement: An element of the Population can not be chosen more than once, as it is not Replaced.

Sampling distribution:

The probability distribution of a sample statistic is often called the **sampling distribution** of the statistic.

Alternatively we can consider all possible samples of size n that can be drawn from the population, and for each sample we compute the statistic. In this manner we obtain the distribution of the statistic, which is its sampling distribution.

Note: For a sampling distribution, we can compute a mean, variance, standard deviation, moments, etc.

Sampling Distributions

- **Population parameters**:

The statistical measures like population mean, population standard etc,. obtained from the population are known as population parameters or simply Parameters .

- **Notation of Population parameters:**

μ : Population mean

σ : population Standard deviation

σ^2 : Variance

p : population proportion.

Sampling Distributions

Sample statistic: Statistical quantities computed from sample observations (sample data) are known as sample statistics or briefly “statistics”.

Notations of Sample statistic

$\bar{X}$: Sample mean(some authors denoted with)

S: Sample Standard deviation

S^2 : Sample variance

P : Sample proportion

Note:

The total number of samples with size ‘n’ drawn from population of size ‘N’ is given by

- sampling **with replacement** N^n

- sampling **without replacement** ${}^{N}C_{n}$

Notations:

Sample mean= $\overline{X}$ or

Mean of sampling distribution of means= $\mu_{\overline{X}}$

Standard deviation of sampling distribution of means $= \sigma_{\overline{X}}$

Sample size =n

Population size= N

Finite population :

Consider a finite population of size N with mean μ and standard deviation σ.
Draw all possible samples of size n without replacement from this population. Suppose N>n then

Infinite population :

Suppose the samples of size n are drawn from an Infinite population or sampling is done with replacement then then

- Mean of Sampling distribution means

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

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

Standard Deviation of Sampling distribution means

Finite Population

$$\sigma_{\overline{X}} = \sqrt{\frac{N-n}{N-1}}\left(\frac{\sigma}{\sqrt{n}}\right)$$

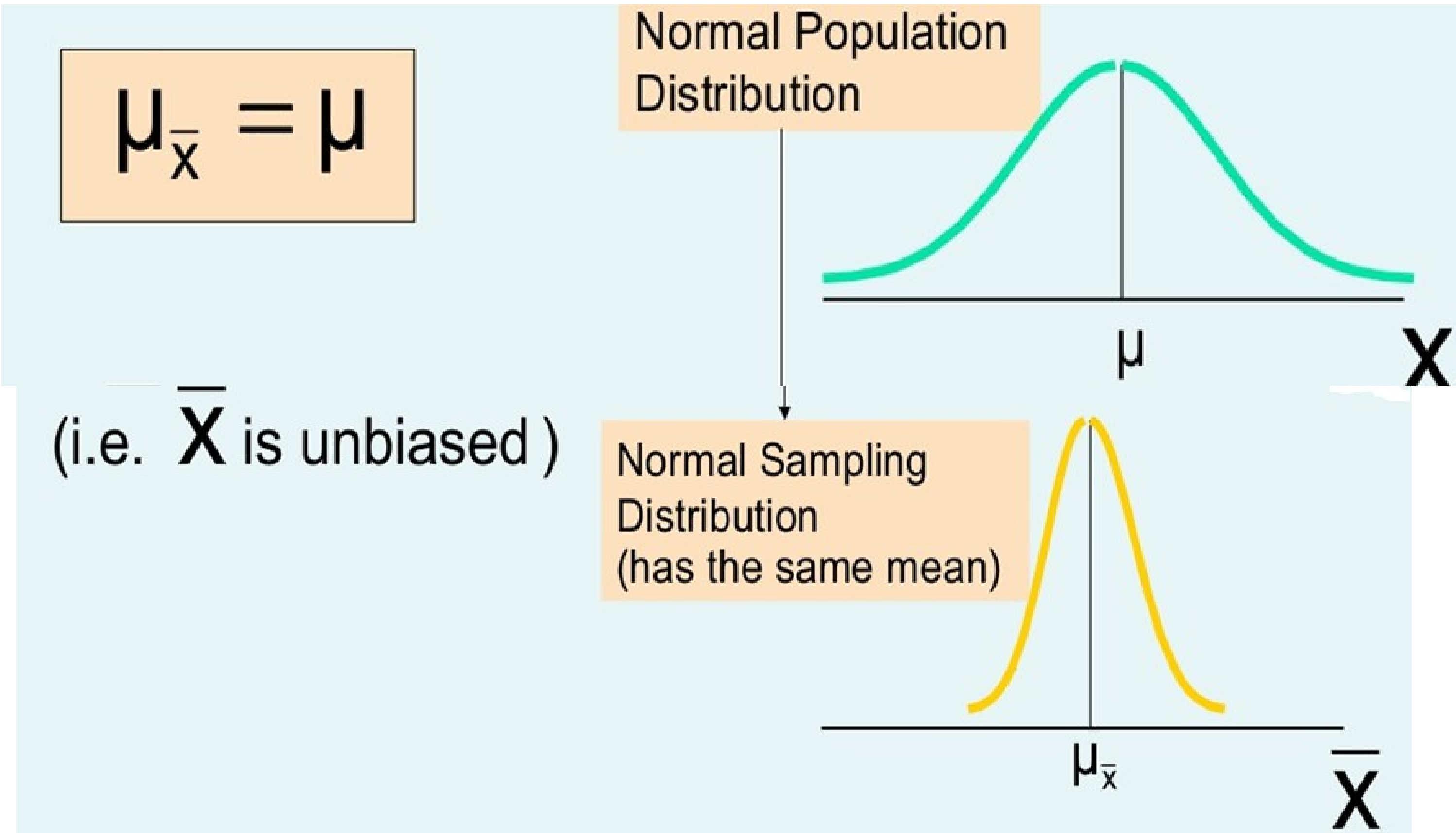
Infinite Population

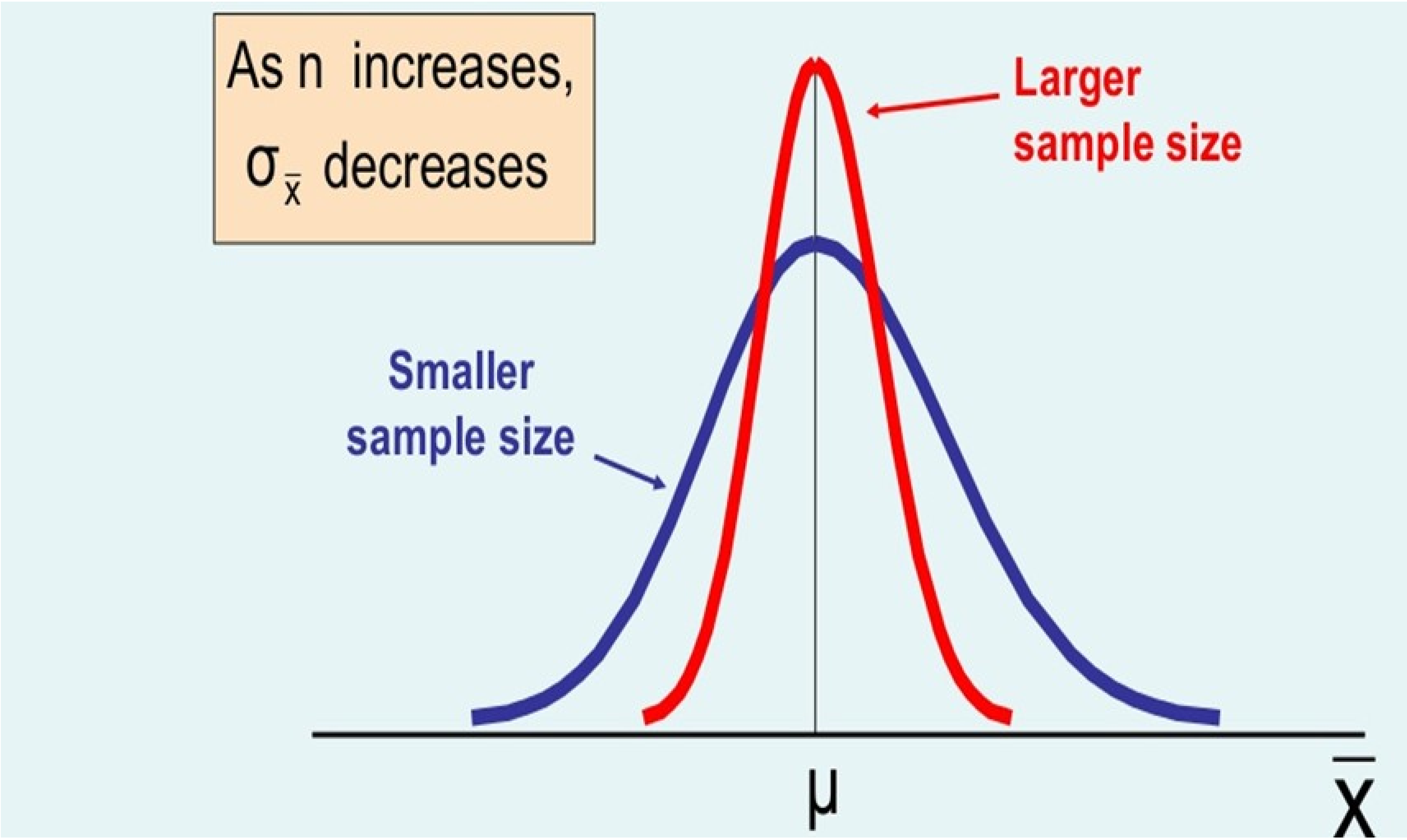
$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

- $\left(\frac{N-n}{N-1}\right)$ is the finite population correction factor.

- $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ is referred to as the standard error of the mean.

- Standardized sample mean $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$


$\mu_{\bar{x}} = \mu$
(i.e. $\bar{X}$ is unbiased)
Normal Population
Distribution
μ
X
Normal Sampling
Distribution
(has the same mean)
$\mu_{\bar{x}}$
$\bar{X}$


As n increases,
$\sigma_{\bar{x}}$ decreases
Larger
sample size
Smaller
sample size
μ
$\bar{X}$

- Z-value for the sampling distribution of $\bar{X}$:

$$Z = \frac{(\bar{X} - \mu_{\bar{X}})}{\sigma_{\bar{X}}} = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$$

where: $\bar{X}$ = sample mean

μ = population mean

σ = population standard deviation

n = sample size

Sampling Distributions of Mean with replacement

Ex. A population consists of FOUR numbers 2,3,4,5. Consider all possible distinct samples of size **‘2’with replacement**. Determine

(i) Mean of population

(ii) Standard deviation of population

(iii) Sampling distributions of mean

(iv) Mean of Sampling distributions of mean

(v) Standard deviation of Sampling distributions of mean

Solution:

Given, N=Population size =4
n= Sample Size=2

- (i) Mean of population:

$$\mu = \frac{2+3+4+5}{4}$$
$$= 3.5$$

- (ii) Standard deviation of population:

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}} \quad \text{Where } i=1 \text{ to } 4$$

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

$$= 1.118$$

(iii) Sampling with replacement (Infinite population):

The total number of samples of size 2 with replacement

$N^n = 4^2 = 16$

All possible samples of size 2 from the population 2,3,4,5 with replacement, we get 16 samples

S. No.	Sample	S. No.	Sample
1	2,2	9	4,2
2	2,3	10	4,3
3	2,4	11	4,4
4	2,5	12	4,5
5	3,2	13	5,2
6	3,3	14	5,3
7	3,4	15	5,4
8	3,5	16	5,5

- The sampling distributions of mean are

S. No.	Sample	Mean	S. No.	Sample	Mean
1	2,2	2	9	4,2	3
2	2,3	2.5	10	4,3	3.5
3	2,4	3	11	4,4	4
4	2,5	3.5	12	4,5	4.5
5	3,2	2.5	13	5,2	3.5
6	3,3	3	14	5,3	4
7	3,4	3.5	15	5,4	4.5
8	3,5	4	16	5,5	5

The sampling distributions of mean can also arranged in the form of frequency distribution

Sample mean: $\overline{X}_i$	**2**	**2.5**	**3**	**3.5**	**4**	**4.5**	**5**
Frequency: f_i	**1**	**2**	**3**	**4**	**3**	**2**	**1**

(iv) The Mean of sampling distributions of mean is

$$\mu_{\overline{X}} = \frac{\sum f_i \overline{X}_i}{\sum f_i}$$

$$= \frac{1(2) + 2(2.5) + 3(3) + 4(3.5) + 3(4) + 2(4.5) + 1(5)}{16}$$

$$= \frac{56}{16} = 3.5$$

(v) The Standard deviation of sampling distributions of mean is

$$\sigma_{\overline{X}} = \sqrt{\frac{\sum f_i \overline{X}_i^{\,2}}{\sum f_i} - \mu_{\overline{X}}^{\,2}}$$

$$= \sqrt{\frac{1(2^2)+2(2.5^2)+3(3^2)+4(3.5^2)+3(4^2)+2(4.5^2)+1(5^2)}{16} - 3.5^2}$$

$$= 0.7905$$

From (i) μ=3.5 (mean of population)

(iv) $\mu_{\overline{X}}$ =3.5 (Mean of sampling distributions of mean)

From (ii) σ = 1.118 (standard deviation of population)

(v) $\sigma_{\overline{X}} = 0.7905 = \frac{\sigma}{\sqrt{n}}$ =1.118/1.414

(Standard deviation of sampling distributions of mean)

- From (i) and (iv)

The Mean of sampling distributions of mean = Mean of population **is verified.**

- From (ii) and (v)

The Standard deviation of sampling distributions of mean = Standard deviation of population divided by root of sample size(n) **is verified.**

Sampling Distributions of Mean without replacement

A population consists of SIX numbers 4,8,12,16,20,24. consider all possible distinct samples of size '2'**without replacement.** Determine

(i) Mean of population

(ii) Standard deviation of population

(iii) Sampling distributions of mean

(iv) Mean of Sampling distributions of mean

(v) Standard deviation of Sampling distributions of mean.

- (i) Mean of population:

(Size of population N=6)

$$\mu = \frac{4+8+12+16+20+24}{6}$$

$$= 14$$

- (ii) Standard deviation of population:

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

$$= \sqrt{\frac{(4-14)^2 + (8-14)^2 + (12-14)^2 + (16-14)^2 + (20-14)^2 + (24-14)^2}{6}}$$

$$= 6.683$$

(iii) The all possible samples of size ‘2’ without replacement are

S. No.	Sample	S. No.	Sample
1	4,8	10	12,16
2	4,12	11	12,20
3	4,16	12	12,24
4	4,20		
5	4,24	13	16,20
		14	16,24
6	8,12		
7	8,16	15	20,24
8	8,20		
9	8,24		

- The sampling distributions of mean are

S. No.	Sample	Mean	S. No.	Sample	Mean
1	4,8	6	10	12,16	14
2	4,12	8	11	12,20	16
3	4,16	10	12	12,24	18
4	4,20	12			
5	4,24	14	13	16,20	18
			14	16,24	20
6	8,12	10			
7	8,16	12	15	20,24	22
8	8,20	14			
9	8,24	16			

(iv) The Mean of sampling distributions of mean is

Sample mean	6	8	10	12	14	16	18	20	22
f	1	1	2	2	3	2	2	1	1

$$\mu_{\overline{X}} = \frac{\sum f_i \overline{X}_i}{\sum f_i} = \frac{1(6)+1(8)+2(10)+2(12)+3(14)+2(16)+2(18)+1(20)+1(22)}{15}$$

$$= \frac{210}{15} = 14$$

(v) The Standard deviation of sampling distributions of mean is

$$\sigma_{\overline{X}} = \sqrt{E(\overline{X}^2) - \mu_{\overline{X}}^{\ 2}} = \sqrt{\frac{\sum f_i \overline{X}_i^{\ 2}}{\sum f_i} - \mu_{\overline{X}}^{\ 2}}$$

$$= \sqrt{\frac{1(6^2)+1(8^2)+2(10^2)+2(12^2)+3(14^2) +2(16^2)+2(18^2)+1(20^2)+1(22^2)}{15} - 14^2}$$

$$= 4.32$$

From (i) μ=14 (mean of population)

(iv) $\mu_{\overline{X}}$ =14 (Mean of sampling distributions of mean).

From (ii) σ = 6.83 (standard deviation of population)

(v) $\sigma_{\overline{X}} = 4.32 = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{6.83}{\sqrt{2}}\sqrt{\frac{6-2}{6-1}}$

$$Note:\ \mu_{\overline{X}} = \mu \quad and$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

- From (i) and (iv) The Mean of sampling distributions of mean = Mean of population **is verified**
- From (ii) and (v) The Standard deviation of sampling distributions of mean = ,

Is verified.

Example:

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.

σ= S.D of population = √Variance = √2.

Standard Error of mean = $\frac{\sigma}{\sqrt{n}}$

= √2/√169

= 1.41/13

= 0.1085.

Example:

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

Example: Find the value of the finite population correction factor for n=10, N=1000

Solution:

$$\text{Finite correction factor} = \frac{N-n}{N-1}$$

$$= \frac{1000-10}{1000-1} = 0.991$$

Central Limit Theorem: If $\overline{X}$ is the mean of random sample of size n taken from a population with mean μ and finite standard deviation σ, then the standardized sample mean

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

is a random variable whose distribution approaches that of standard normal distribution as $n \to \infty$.

Example: Determine the mean and standard deviation of the sampling distribution of means of 300 random samples each of size n=36 are drawn from a population with size N=1500 which is normally distributed with mean μ=22.4 and σ=0.048,if sampling is done (a) with replacement (b) without replacement.

Solution:

(a).With replacement

$$\mu_{\overline{X}} = \mu = 22.40$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{0.048}{\sqrt{36}} = 0.008$$

(b) Without replacement

$$\mu_{\overline{X}} = \mu = 22.40$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{0.048}{\sqrt{36}}\sqrt{\frac{1500-36}{1500-1}} = 0.0079$$

$$\sigma_{\overline{X}} \approx 0.008$$

Example 2: Determine the expected number of random samples having their means (a) between 22.39 and 22.41 (b) greater than 22.42 (c) less than 22.37 (d) less than 22.8 or more than 22.41. Where the sampling distribution of means of 300 data random samples each of size n=36 are drawn from a population with size N=1500 which is normally distributed with mean μ=22.4 and σ=0.048.

Solution:

N=size of Population=1500

n=sample size=36

N_s=Number of samples=300

μ=population mean= 22.4 and

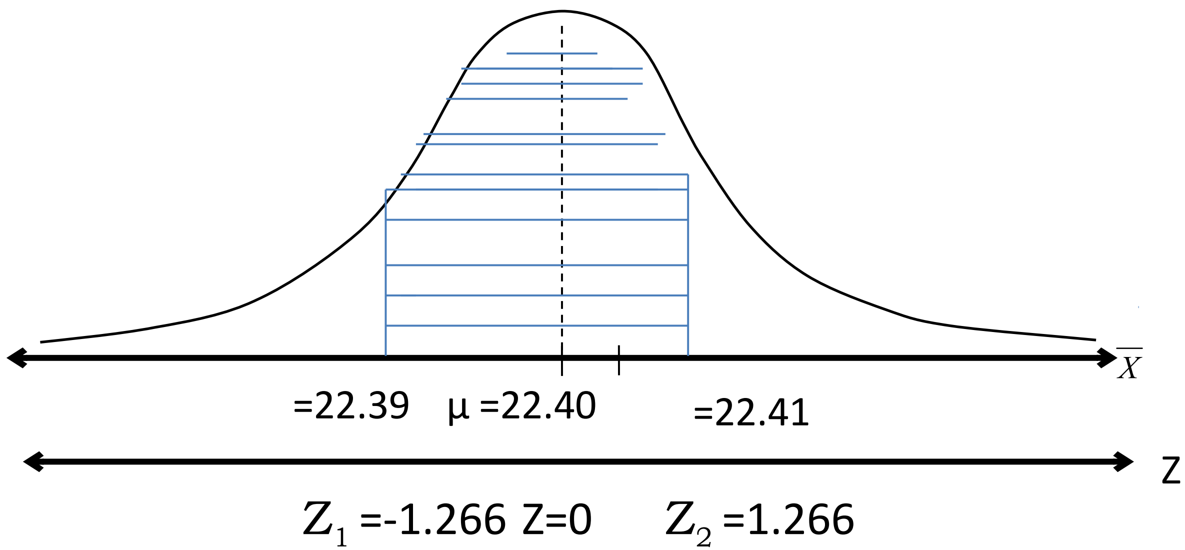
σ=population standard deviation=0.048.

•

(a). For $\bar{X}_1=22.39, Z_1=(22.39-22.4)/0.0079= -1.26$

$\bar{X}_2=22.41, Z_2=(22.41-22.4)/0.0079= 1.26$

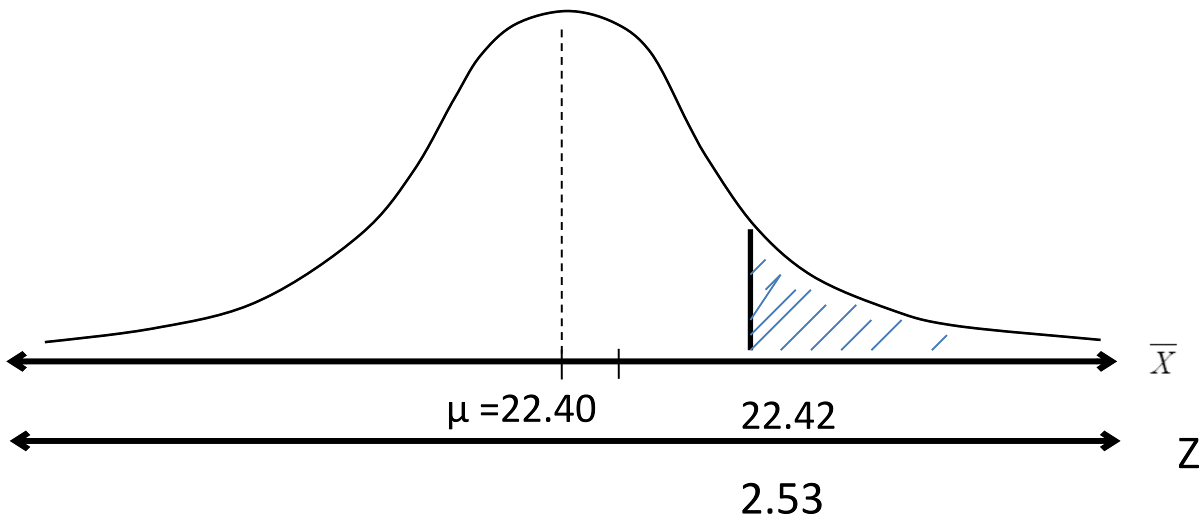
$$P(22.39< \bar{X}<22.41)=P(-1.26<Z<1.26)=2(0.3962)=0.7924$$



Expected number of samples=(Total number of samples) (Probability)

=(300)(0.7924)=238

Therefore expected number of samples who have mean lying between 22.39 to 22.41 is =(300)(0.7924)=238 (b). For $\overline{X}_1=22.42, Z_1=(22.42-22.4)/0.0079= 2.53$



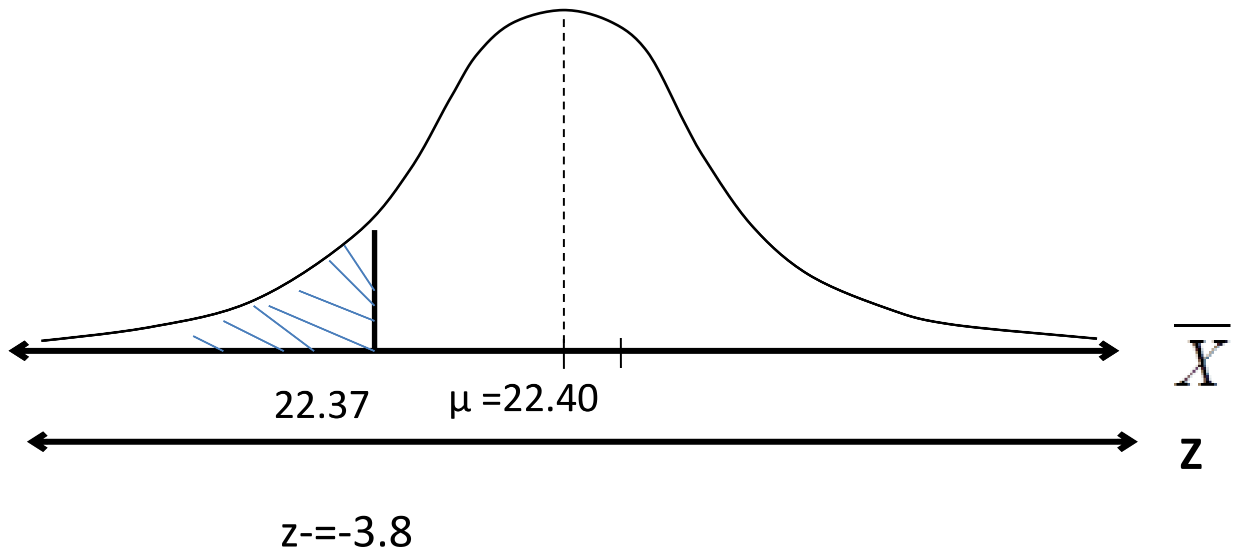
$P(\overline{X}>22.42)=P(Z>2.53)=0.5-0.4943=0.00057$

Therefore Expected number of samples who have mean greater than 22.42 is

$=(300)(0.00057)=2(\text{Approx})$

(c). For $\bar{X}_1 = 22.37, Z_1 = (22.37 - 22.4)/0.0079 = -3.8$

$P(\bar{X} < 22.37) = P(Z < -3.8) = 0.5 - 0.4949 = 0.0001$



Therefore Expected number of samples who have mean greater than 22.37 is

$= (300)(0.0001) = 0 \text{(Approx)}$

(d). For $\bar{X}_1=22.38, Z_1=(22.38-22.4)/0.0079$

$= -2.53$

For $\bar{X}_2=22.41, Z_2=(22.41-22.4)/0.0079$

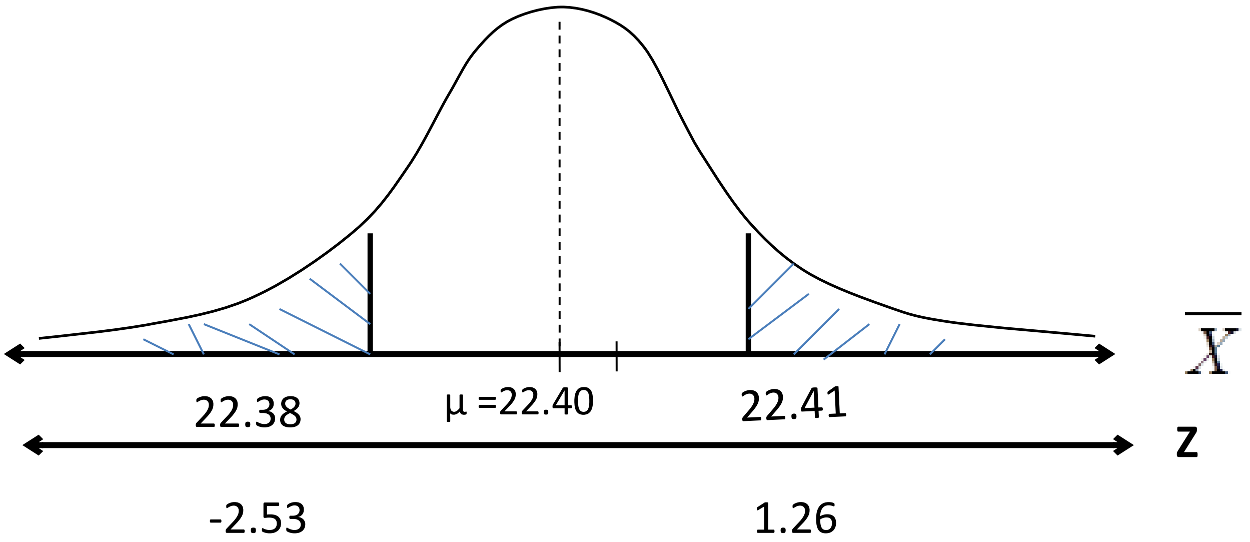
$= 1.26$

$P(\bar{X}<22.38 \text{ or } \bar{X}>22.41)$

$=P(Z<-2.53 \text{ or } Z>1.26)$

$= (0.5-0.4943)+ (0.5-0.3962)$

$=0.0057+0.1038=0.1095$



Therefore Expected number of samples =(300)(0.1095)=33(Approx)

Example

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

Solution:

Given that sample size $n = 100$

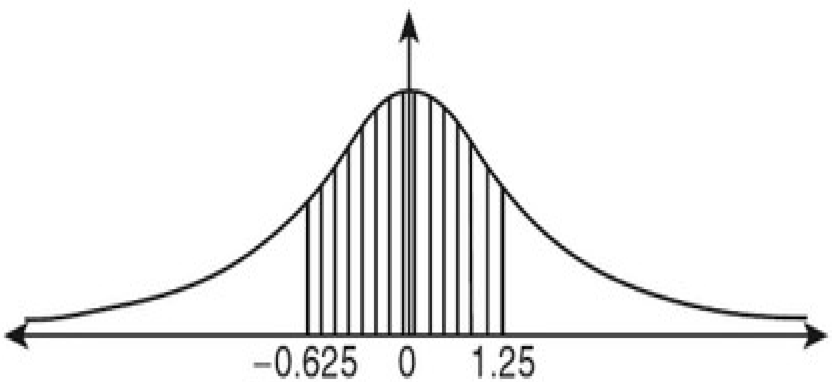
Mean $\mu = 76$, variance $\sigma^2 = 256$

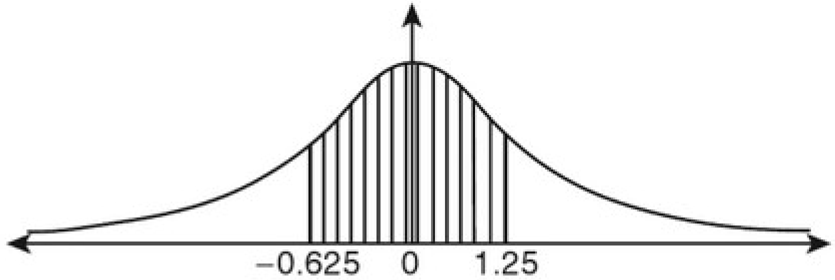
Probability that $\bar{x}$ will be between 75 and 78 is $P(75 < \bar{x} < 78)$

Using the central limit theorem, $z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

When $\bar{x} = 75$, $z_1 = \dfrac{75 - 76}{\sqrt{\dfrac{256}{100}}} = -0.625$

When $\bar{x} = 78$, $z_2 = \dfrac{78 - 76}{\sqrt{\dfrac{256}{100}}} = 1.25$





$$\therefore P(75 < \bar{x} < 78) = P\,(-0.625 < z < 1.25)$$

$$= P(-0.625 < z < 0) + P\,(0 < z < 1.25)$$

$$= P\,(0 < z < 0.625) + P(0 < z < 1.25)$$

$$= 0.234 + 0.3944 = 0.628$$

Example

A random sample of size 100 is taken from an infinite population having the mean 80 and standard deviation 20. What is the

(i) Probability that $\bar{x}$ will be greater than 85?

(ii) Probability that $\bar{x}$ will lie between 75 and 85?

Solution: Given that sample size = 100

Mean $= \mu = 80$

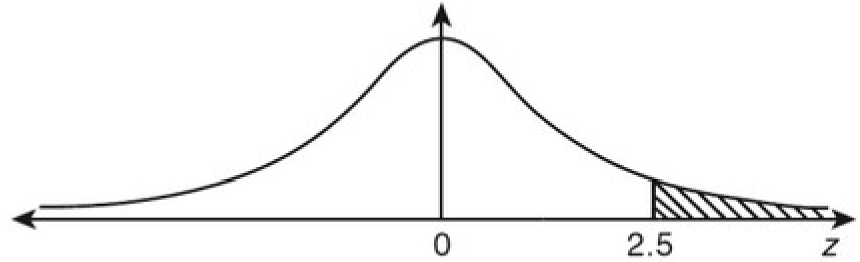
Standard deviation, $\sigma = 20$

The standard normal variant, $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

(i) Probability that $\overline{x}$ will be greater than 85

$$= P(\overline{x} > 85),$$

When $\overline{x} = 85$, $z = \dfrac{85-80}{\dfrac{20}{\sqrt{100}}} = \dfrac{5}{2} = 2.5$



$$P(\overline{x} > 85) = P(z > 2.5)$$

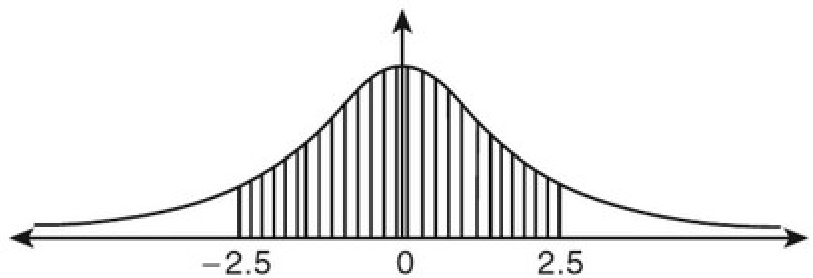
$$= P(0 < z < \infty) - P(0 < z < 25)$$

$$= 0.5 - 0.498$$

$$= 0.0065$$

(ii) Probability that $\overline{x}$ will be between 75 and 85 $= P(75 < \overline{x} < 85)$

When $\overline{x} = 75$, $z = \dfrac{75-80}{\dfrac{20}{\sqrt{100}}} = \dfrac{-5}{2} = -2.5$



$$P(75 < \overline{x} < 85) = P(-2.5 < \overline{x} < 2.5)$$

$$= 2 \times P(0 < \overline{x} < 2.5)$$

$$= 2 \times 0.4938 = 0.\ 9876.$$

Example:

The mean height of students of a class is 155 cms and standard deviation is 15. What is the probability that the mean height of 36 students is less than 157 cm?

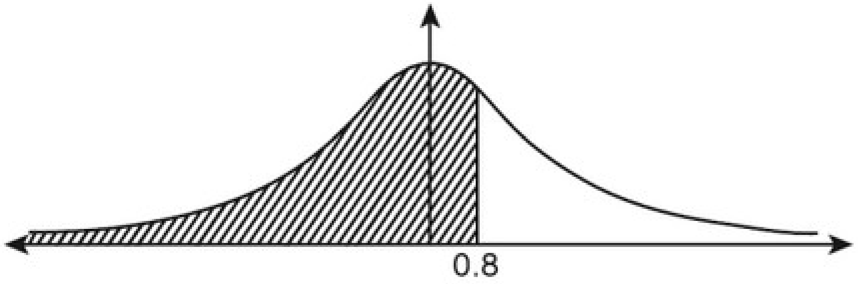
Solution: Mean height of students $\mu = 155$ cm

Standard deviations, $\sigma = 15$

Probability that mean height of 36 students is less than 157 cm $= P(\bar{x} < 157)$

When $\bar{x} = 157$, $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{157 - 155}{\frac{15}{\sqrt{36}}} = \dfrac{+2}{2.5} = 0.8$

$$\begin{aligned} P(\bar{x} < 157) &= p\ (z < 0.8) \\ &= P(-\infty < z < 0) + P(0 < z < 0.8) \\ &= 0.5 + 0.2881 \\ &= 0.7881 \end{aligned}$$



Example:

An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a random sample of 16 bulbs will have an average life of

(i) less than 775 hours?

(ii) Will be between 775 and 825 hours.

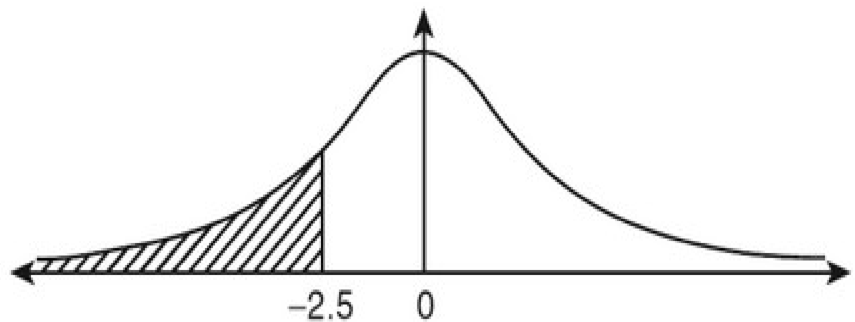
Solution: Mean life of light bulbs = 800 hours

Standard deviation of light bulbs $\sigma = 40$ hours

Probability that bulbs have an average life of less than 775 hours = $P(\bar{x} < 775)$

(i) When $\bar{x} = 775$, $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{775 - 800}{\frac{40}{\sqrt{16}}} = -2.5$

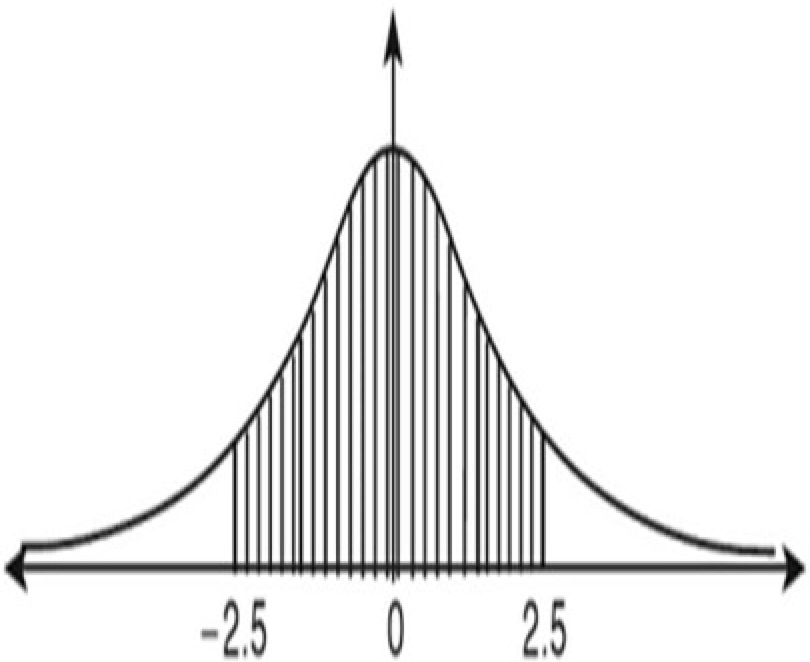
$$\begin{aligned} P(\bar{x} < 775) &= P(z < -2.5) \\ &= P(Z > 2.5) \\ &= -P(0 < z < \infty) \\ &\quad -P(0 < z < 2.5) \\ &= 0.5 - 0.4938 = 0.0062 \end{aligned}$$



(ii) Probability that average life will be between 775 and 825 hours = $P(775 < \bar{x} < 825)$

When $\bar{x} = 825$,

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{825 - 800}{\frac{40}{\sqrt{16}}} = 2.5$$



$$P(775 < \bar{x} < 825) = P(-2.5 < z < 2.5)$$

$$= 2P(0 < z < 2.5)$$

$$= 2(0.4932) = 0.9876$$

Sampling Distribution of the mean (σ Unknown)

For the large sample (of size $n \geq 30$):we can substitute the sample Standard deviation S in place of σ.

The sample standard deviation S is calculated using the sample mean $\bar{X}$ by the following formula

$$S^2 = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)^2}{n-1}$$

For small size $n<30$,when σ is unknown it can be substituted by S, provided we make an assumption that the sample is drawn from a normal population.


THANK
YOU

PROBABILITY AND STATISTICS

Unit-II

PPT-2

Sampling Distribution & Estimation

(2021-22) For ECM, CSE(A).

Dr.B.Vijayabaskerreddy

Faculty of Mathematics
Department of Science & Humanities
SNIST

04/09/2025

Sampling distribution of differences and sums

Let μ_{S_1} and σ_{S_1} be the mean and standard deviation of sampling distribution of statistic S_1 obtained by computing S_1 for all possible samples of size n_1 drawn from population A.

Let μ_{S_2} and σ_{S_2} be the mean and standard deviation of sampling distribution of statistic S_2 obtained by computing S_2 for all possible samples of size n_2 drawn from population B.

The mean $\mu_{S_1 - S_2}$ and the standard deviation $\sigma_{S_1 - S_2}$ of the sampling distribution of differences are given by

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

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

Assuming that the samples are independent.

Sampling distribution of sum of statistics has mean $\mu_{S_1+S_2}$ and standard $\sigma_{S_1+S_2}$ given by

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

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

For example, for infinite population

$$(i)\ \mu_{\overline{X}_1+\overline{X}2}=\mu_{\overline{X}_1}+\mu_{\overline{X}_2}=\mu_1+\mu_2$$

$$(ii)\ \sigma_{\overline{X}_1+\overline{X}2}=\sqrt{\sigma^2_{\overline{X}_1}+\sigma^2_{\overline{X}_2}}=\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}$$

Example: Let $U_1=\{2,7,9\}, U_2=\{3,8\}$.

Find

$$(a)\mu_{U_1} \quad (b)\mu_{U_2} \quad (c)\mu_{U_1+U_2} \quad (d)\mu_{U_1-U_2}$$

$$(e)\sigma_{U_1} \quad (f)\sigma_{U_2} \quad (g)\sigma_{U_1+U_2} \quad (h)\sigma_{U_1-U_2}$$

Verify that

$$(i)\mu_{U_1+U_2} = \mu_{U_1} + \mu_{U_2} \quad (j)\mu_{U_1-U_2} = \mu_{U_1} - \mu_{U_2}$$

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

$$(l)\sigma_{U_1-U_2} = \sqrt{\sigma^2{}_{U_1} + \sigma^2{}_{U_2}}$$

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

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

U_1

$$(c)\ U_1 + U_2 = \{2+3, 7+3, 9+3, 2+8, 7+8, 9+8\}$$

$$(d)\, U_1 - U_2 = \{2-3, 7-3, 9-3, 2-8, 7-8, 9-8\}$$

$$(e)\, \sigma^2_{U_1} = \text{Variance of population } U_1 \text{ with mean } 6$$

$$= \frac{(2-6)^2 + (7-6)^2 + (9-6)^2}{3} = \frac{26}{3} = 8.66$$

$$\sigma_{U_1} = \sqrt{8.66} = 2.9439$$

$(f)\ \sigma^2_{U_2}$ = Variance of population U_2 with mean 5.5

$$= \frac{(3-5.5)^2 + (8-5.5)^2}{2} = \frac{26}{3} = 6.25$$

$$\sigma_{U_2} = \sqrt{6.25} = 2.5$$

$(g)\ \sigma^2_{U_1+U_2}$ = Variance of population $U_1 + U_2$ with mean 11.5

$$= \frac{(5-11.5)^2 + (10-11.5)^2 + (12-11.5)^2 + (10-11.5)^2 + (15-11.5)^2 + (17-11.5)^2}{6} = 14.9166$$

$$\sigma_{U_1+U_2} = \sqrt{14.9166} = 3.86220$$

$(h)\ \sigma^2_{U_1-U_2}$ = Variance of population $U_1 - U_2$ with mean 0.5

$$= \frac{(-1-0.5)^2 + (4-0.5)^2 + (6-0.5)^2 + (-6-0.5)^2 + (-1-0.5)^2 + (1-0.5)^2}{6} = 14.9166$$

$$\sigma_{U_1-U_2} = \sqrt{14.9166} = 3.86220$$

.

Therefore .

.

$$(k)\quad \sigma_{U_1+U_2} = 3.86220$$

$$\sqrt{\sigma^2_{U_1} + \sigma^2_{U_2}} = \sqrt{8.66 + .25} = 3.8620$$

$$\text{Therefore } \sigma_{U_1+U_2} = \sqrt{\sigma^2_{U_1} + \sigma^2_{U_2}}$$

$$(l)\quad \sigma_{U_1-U_2} = 3.86220$$

$$\sqrt{\sigma^2_{U_1} + \sigma^2_{U_2}} = \sqrt{8.66 + .25} = 3.8620$$

$$\text{Therefore } \sigma_{U_1-U_2} = \sqrt{\sigma^2_{U_1} + \sigma^2_{U_2}}$$

Let U_1 be a variable that stands for any of the elements of the population 3, 7, 8 and U_2 be a variable that stands for any of the elements of the population 2, 4. Compute (a) μ_{U1}, (b) μ_{U2}, (c) μ_{U1-U2}, (d) σ_{U1}, (e) σ_{U2}, and (f) σ_{U1-U2}.

(a) μ_{U1} = mean of population $U_1 = \frac{1}{3}\ (3 + 7 + 8) = 6$

(b) μ_{U2} = mean of population $U_2 = \frac{1}{2}\ (2 + 4) = 3$

(c) The population consisting of the differences of any member of U_1 and any member of U_2 is

$$\begin{array}{ccccccc} 3-2 & 7-2 & 8-2 & \quad & 1 & 5 & 6 \end{array}$$

or

$$\begin{array}{ccccccc} 3-4 & 7-4 & 8-4 & \quad & -1 & 3 & 4 \end{array}$$

Thus $\mu_{U1-U2} = \text{mean of } (U_1 - U_2) = \dfrac{1+5+6+(-1)+3+4}{6} = 3$

This illustrates the general result $\mu_{U1-U2} = \mu_{U1} - \mu_{U2}$, as seen from parts (a) and (b).

(d) $\sigma_{U1}^2 = \text{variance of population } U_1 = \dfrac{(3-6)^2+(7-6)^2+(8-6)^2}{3} = \dfrac{14}{3}$

or $\sigma_{U1} = \sqrt{\dfrac{14}{3}}$

(e) σ_{U2}^2 = variance of population $U_2 = \dfrac{(2-3)^2+(4-3)^2}{2} = 1$ or $\sigma_{U2} = 1$

(f) σ_{U1-U2}^2 = variance of population $(U_1 - U_2)$

$$= \frac{(1-3)^2+(5-3)^2+(6-3)^2+(-1-3)^2+(3-3)^2+(4-3)^2}{6} = \frac{17}{3}$$

or $\sigma_{U1-U2} = \sqrt{\dfrac{17}{3}}$

This illustrates the general result for independent samples, $\sigma_{U1-U2} = \sqrt{\sigma_{U1}^1 + \sigma_{U2}^2}$, as seen from parts (d) and (e).

Example: The mean voltage of a battery is 15 and standard deviation is 0.2. Find the probability that four such batteries connected in series will have a combined voltage of 60.8 or more volts.

Solution:

Let the sample of four batteries A,B,C and D connected in series

Let μ_A , $\mu_{\mathbf{B}}$ μ_C and μ_D denote the mean voltage of A, B, C and D respectively.

$$Given\ \mu_A = \mu_{\mathbf{B}} = \mu_C = \mu_D = 15$$

$$\mu = \mu_{A+B+C+D} = \mu_A + \mu_{\mathbf{B}} + \mu_C + \mu_D$$

Let σ_A , $\sigma_{\mathbf{B}}$ σ_C and σ_D denote the standard deviation of A, B, C and D respectively.

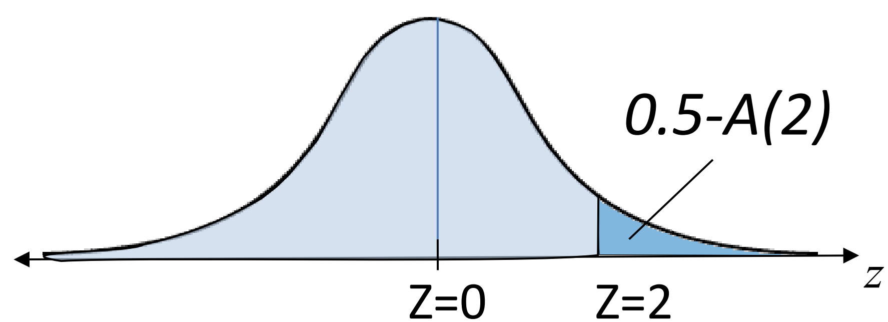
$$\sigma_A = \sigma_{\mathbf{B}} = \sigma_C = \sigma_D = 0.2$$

Let X be the combined voltage of the series.
Given X_1=60.8 .

X_1=60.8 in standard units

$$Z_1 = \frac{X_1 - \mu}{\sigma} = \frac{60.8 - 60}{0.4} = 2$$

Probability that the combined voltage is more than 60.8 is given by
$P(X \geq 60.8) = P(Z \geq 2)$
$= 0.5 - \text{Area}(2)$
$= 0.5 - 0.4772 = 0.0228$



Example

A certain type of electric light bulb has a mean lifetime of 1500 hours and a standard deviation of 150 hours. Three bulbs are connected so that when one burns out, another will go on. Assuming the lifetimes are normally distributed, what is the probability that lighting will take place for

(a) at least 5000 hours,

(b) at most 4200 hours?

Denote the electric light lifetimes as L_1, L_2, and L_3. Then

$$\mu = ¿\ \mu_{L_1+L_2+L_3} = \mu_{L_1} + \mu_{L_2} + \mu_{L_3} = 1500 + 1500 + 1500 = 4500 \text{ hours}$$

$$\sigma = ¿\ \sigma_{L_1+L_2+L_3} = \sqrt{\sigma^2_{L_1} + \sigma^2_{L_2} + \sigma^2_{L_3}} = \sqrt{3(150)^2} = 260 \text{ hours}$$

Let X be that lighting will take place for certain hours.

(a) Let X_1=5000 hours

X_1=5000 hours in standard units

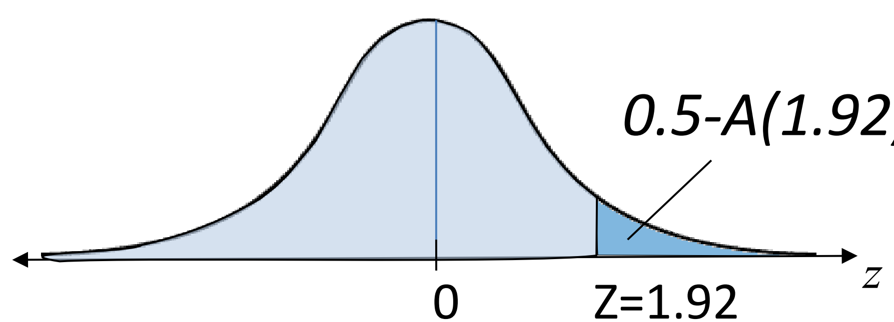
$$Z_1 = \frac{X_1 - \mu}{\sigma} = \frac{5000 - 4500}{260} = 1.92$$

P(X>5000)=P(Z>1.92)

=Area under normal curve to right of Z=1.92

=0.5-A(1.92)=0.5-4726

=0.0274



(b) Let X_2=4200 hours

X_2=4200 hours in standard units

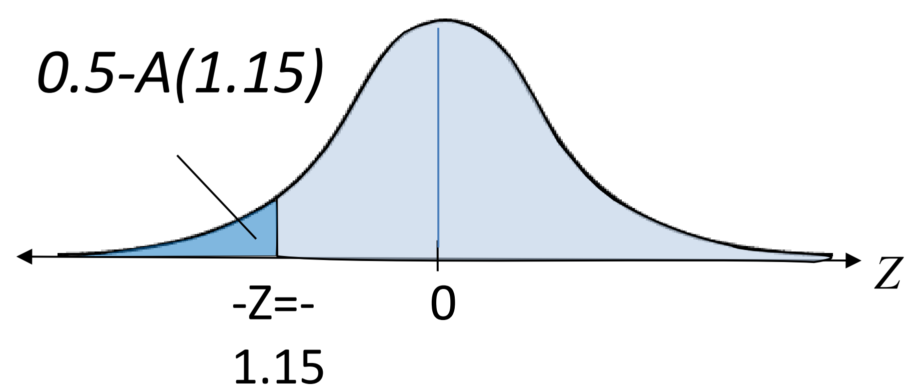
$$Z_2 = \frac{X_2 - \mu}{\sigma} = \frac{5000 - 4200}{260} = -1.15$$

P(X<4200)=P(Z<-1.15)

=Area under normal curve to Left of Z=-1.15

=0.5-A(1.15)=0.5-0.3749

=0.1251



Example:
Two distances are measured as 27.3 inches and 15.6 inches, with standard deviations (standard errors) of 0.16 inches and 0.08 inches, respectively.
Determine the mean and standard deviation of
(a) the sum,
(b) the difference of the distances.

Solution:

If the distances are denoted by D1 and D2, then

(a)

$$\mu_{D_1+D_2} = \mu_{D_1} + \mu_{D_2} = 27.3 + 15.6 = 42.9 \text{ inches}$$

$$\sigma_{D_1+D_2} = \sqrt{\sigma_{D_1}^2 + \sigma_{D_2}^2} = \sqrt{(0.16)^2 + (0.08)^2} = 0.18 \text{ inches}$$

(b)

$$\mu_{D_1-D_2} = \mu_{D_1} + \mu_{D_2} = 27.3 + 15.6 = 11.7 \text{ inches}$$

$$\sigma_{D_1-D_2} = \sqrt{\sigma_{D_1}^2 + \sigma_{D_2}^2} = \sqrt{(0.16)^2 + (0.08)^2} = 0.18 \text{ inches}$$

Example:

The electric light bulbs of manufacturer **A** have a mean lifetime of 1400 hours with a standard deviation of 200 hours, while those of manufacturer **B** have a mean lifetime of 1200 hours with a standard deviation of 100 hours. If random samples of 125 bulbs of each brand are tested,

what is the probability that a) The brand A bulb will have a mean lifetime that is at least 160 hours more than the brand B bulbs? b) The brand A bulb will have a mean lifetime that is at least 250 hours more than the brand B bulbs?

Solution

Let $\bar{X}_A$ and $\bar{X}_B$ denote the mean lifetimes of samples A and B, respectively. Then

$$\mu_{\bar{X}_A-\bar{X}_B} = \mu_{\bar{X}_A} - \mu_{\bar{X}_B} = 1400 - 1200 = 200 \text{ hours}$$

$$\sigma_{\bar{X}_A-\bar{X}_B} = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = \sqrt{\frac{(100)^2}{125} + \frac{(200)^2}{125}} = 20 \text{ hours}$$

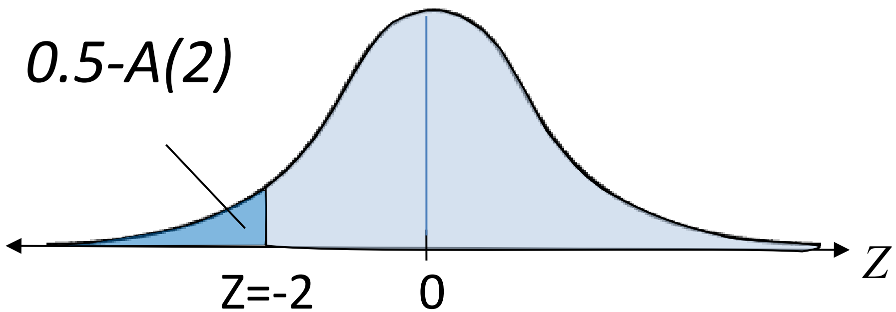
The standardized variable for the difference in means that

$$Z = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_{\bar{X}_A - \bar{X}_B})}{\sigma_{\bar{X}_A - \bar{X}_B}} = \frac{(\bar{X}_A - \bar{X}_B) - 200}{20}$$

and is very nearly normally distributed.

(a) The difference 160 hours in standard units = $(160 - 200)/20 = -2$.

$$\text{Required probability} = (\text{area under normal curve to right of } z = -2)$$
$$= 0.5000 + 0.4772 = 0.9772$$



(b) The difference 250 hours in standard units = $(250 - 200)/20 = 2.50$.

$$\text{Required probability} = (\text{area under normal curve to right of } z = 2.50)$$
$$= 0.5000 - 0.4938 = 0.0062$$

Example: Ball bearings of a given brand mean weigh 0.50 oz with a standard deviation of 0.02 oz. What is the probability that two lots, of 1000 ball bearings each, will differ in weight by more than 2 oz?

Solution:

Let $\bar{X}_1$ and $\bar{X}_2$ denote the mean weights of ball bearings in the two lots. Then

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = 0.50 - 0.50 = 0$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(0.02)^2}{1000} + \frac{(0.02)^2}{1000}} = 0.000895$$

The standardized variable for the difference in means is

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{0.000895}$$ and is very nearly normally

distributed.

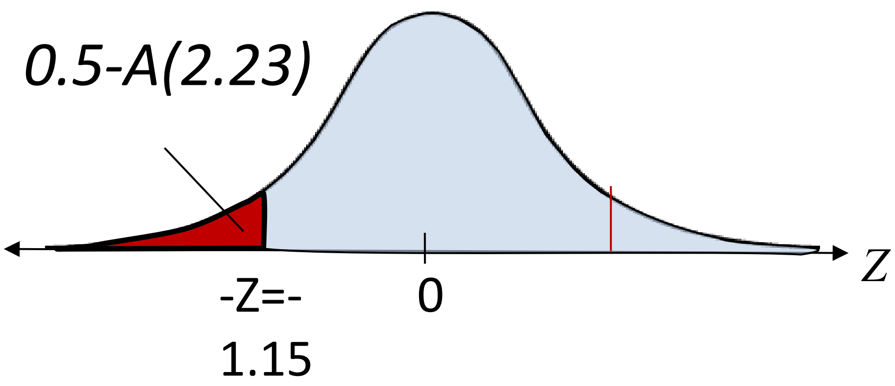
A difference of 2 oz in the lots is equivalent to a difference of 2/1000=0.002 oz in the means.

This can occur either if

$$\bar{X}_1 - \bar{X}_2 \geq 0.002 \text{ or } \bar{X}_1 - \bar{X}_2 \leq -0.002, \text{i.e.,}$$

$$Z \geq \frac{0.002 - 0}{0.000895} = 2.23 \qquad \text{or} \qquad Z \leq \frac{-0.002 - 0}{0.000895} = -2.23$$

$P(Z \geq 2.23 \text{ or } Z \leq -2.23)$



$= P(Z \geq 2.23) + P(Z \leq 2.23)$

$= 2(0.5000 - 0.4871)$

$= 0.0258$

Example:

Let $\overline{X}_1$ and $\overline{X}_2$ be the average drying times of two types of oil paints with sample sizes $n_1 = n_2 = 18$. Find $P(\overline{X}_1 - \overline{X}_2 > 1)$ assuming that $\sigma_1 = \sigma_2 = 1$ and the mean drying times are equal for the two types of oil paints.

Solution We have $\sigma^2_{\overline{X}_1 - \overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{1}{18} + \frac{1}{18} = \frac{1}{9}$

$$P(\overline{X}_1 - \overline{X}_2 > 1) = P\left(z > \frac{1 - (\mu_1 - \mu_2)}{\sigma_{\overline{X}_1 - \overline{X}_2}}\right)$$

$$= P\left(z > \frac{1}{\sqrt{1/9}}\right)$$

$$P(z > 3) = 1 - P(0 < z < 3) = 1 - 0.9987$$

$$= 0.0013$$

Thanks for watching this video

PROBABILITY AND STATISTICS

Unit-II

PPT-3

Sampling Distribution & Estimation

(2021-22) For ECM, CSE(A).

Dr.B.Vijayabaskerreddy

Faculty of Mathematics
Department of Science & Humanities
SNIST

04/09/2025

Estimation :

The method of obtaining the most likely value of the population parameter using statistic is called estimation.

Estimator:

Any sample statistic which is used to estimate an unknown population parameter is called an estimator.

Estimate :

When we observe a specific numerical value of our estimator, we call that value is an estimate. In other words, an **estimate is a specific observed value of a statistic**.

Types of Estimation:

- **Point Estimation:** Point estimation of a parameter is a statistical estimation where the parameter is estimated by the single number from the sample data.

- **Interval Estimation:** Interval estimation is a statistical estimation where the parameter is estimated by an Interval obtained from the sample data.

Confidence Interval for Population Mean when σ known(for large sample)

If $\bar{X}$ is the mean of a random sample of size n from the population with known variance σ^2, $(1-\alpha)100\%$ confidence interval for μ is given by

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

Where $Z_{\alpha/2}$ is the Z-value leaving an area of $\alpha/2$ to the right.

Error of Estimate

Since the sample mean estimate vary rarely equals to the mean of population μ, a point estimate is generally accompanied with a statement of error which gives difference between estimate and the quantity to be estimated, the estimator .Thus error = $\bar{X}-\mu$.

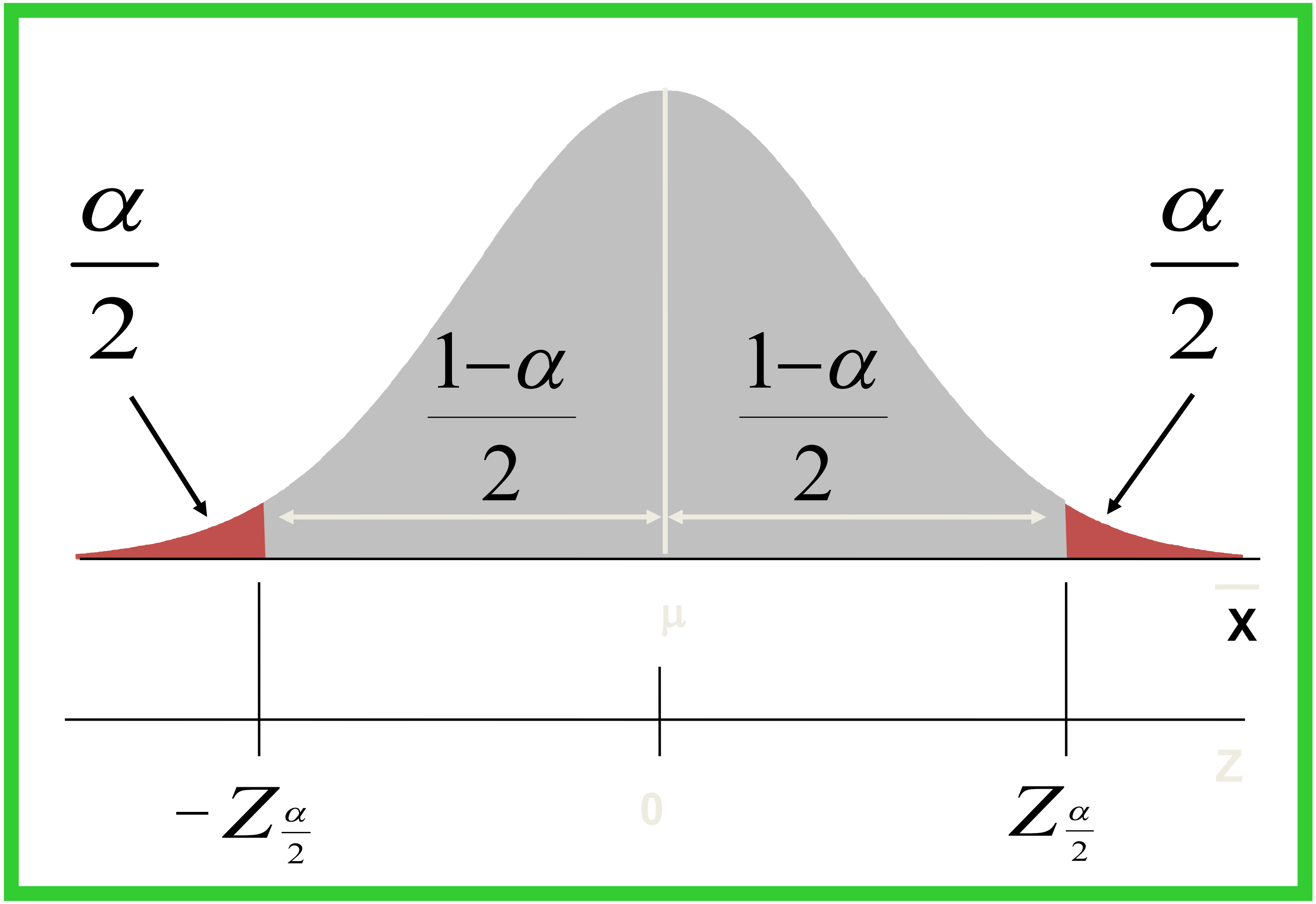
- The **maximum error** of estimate E with (1-α) probability is given by

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

- **Sample size:** when α,E,σ are known, the sample size n is given by

$$n = \left(\frac{Z_{\alpha/2}\sigma}{E} \right)^2$$

Distribution of Sample Means for (1-α)100% Confidence



Confidence Interval Estimate of the Population Mean for large sample when σ unknown

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$\bar{X}$: sample mean

S: sample standard deviation

n: sample size

Confidence Interval Estimate of the Population Mean for small sample(n<30) when σ unknown

$$\bar{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$\bar{X}$: sample mean

S: sample standard deviation

n: sample size

- **for small sample(n<30) when σ unknown**

The **maximum error** of estimate E with (1-α) probability is given by

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

Here t-distribution is with n-1 degrees of freedom

Example : The efficiency expert of a computer company tested 40 engineers to estimate the average time it takes to assemble a certain computer component , getting mean of 12.73 minutes and standard deviation of 2.06 minutes

(a).If $\bar{X}$ =12.73 is used as a point estimate of the actual average time required to perform the task, determine the maximum error with 99% confidence.

(b) Construct 98% confidence interval for the true average time it takes to do the job.

(c) with what confidence can we assert that the sample mean does not differ from the true mean by more than 30 seconds.

Solution:

Given data

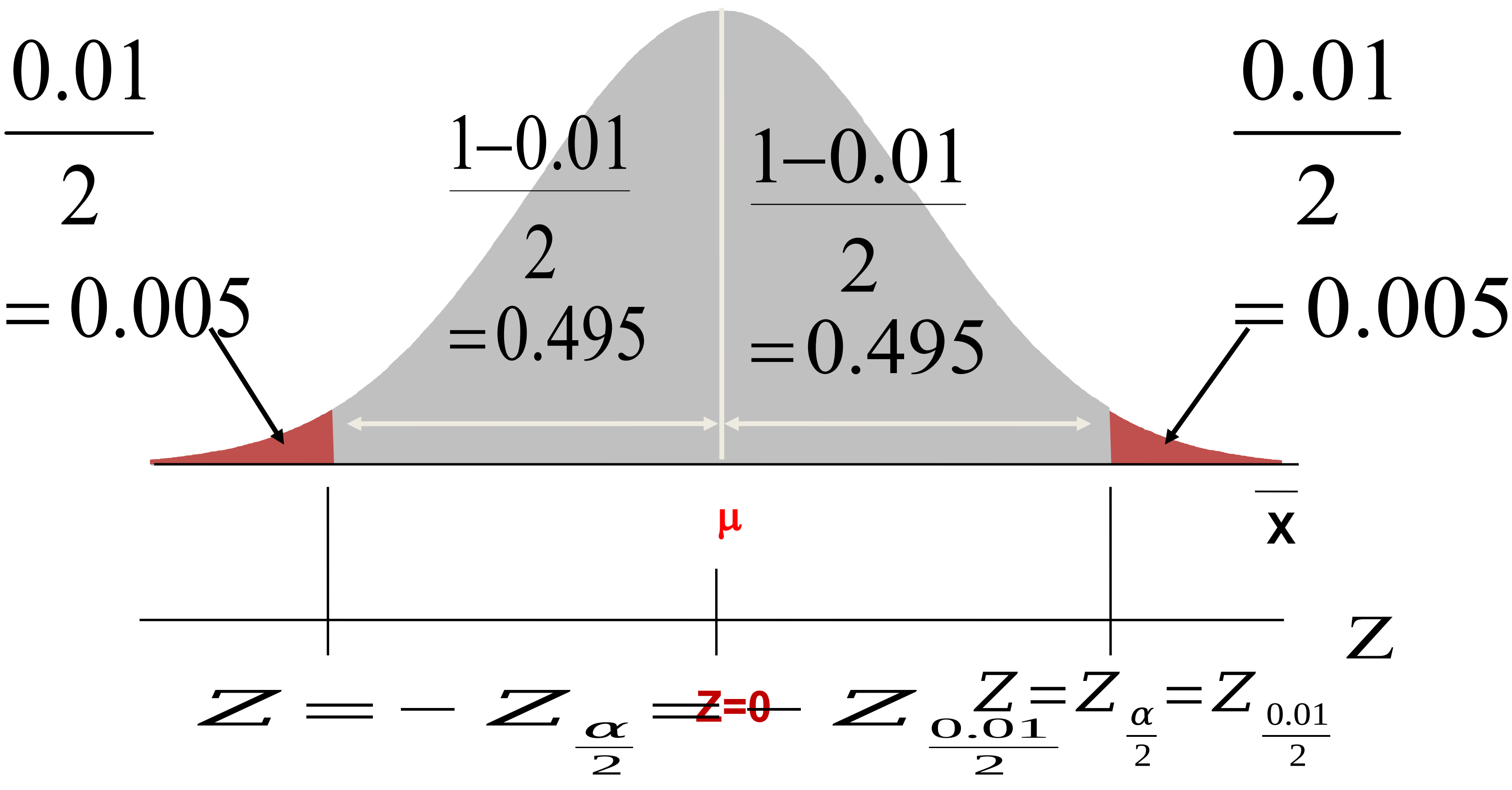
Sample mean = $\bar{X}=12.73.$

Sample standard deviation $=S=2.06.$

(a).For 99% confidence ,

$(1-\alpha)=99/100=0.99$ implies $\alpha=0.01$

And $[(1-\alpha)/2]=0.495$



$\frac{0.01}{2} = 0.005$
$\frac{1-0.01}{2} = 0.495$
$\frac{1-0.01}{2} = 0.495$
$\frac{0.01}{2} = 0.005$
μ
$\bar{X}$
Z
$Z = -Z_{\frac{\alpha}{2}} = -Z_{\frac{0.01}{2}}$
Z=0
$Z = Z_{\frac{\alpha}{2}} = Z_{\frac{0.01}{2}}$

$A(Z_{\alpha/2})=0.495$

$=(0.4949+0.4951)/2$

Implies

$Z_{\alpha/2}= (2.57+2.58)/2$

$=2.575$

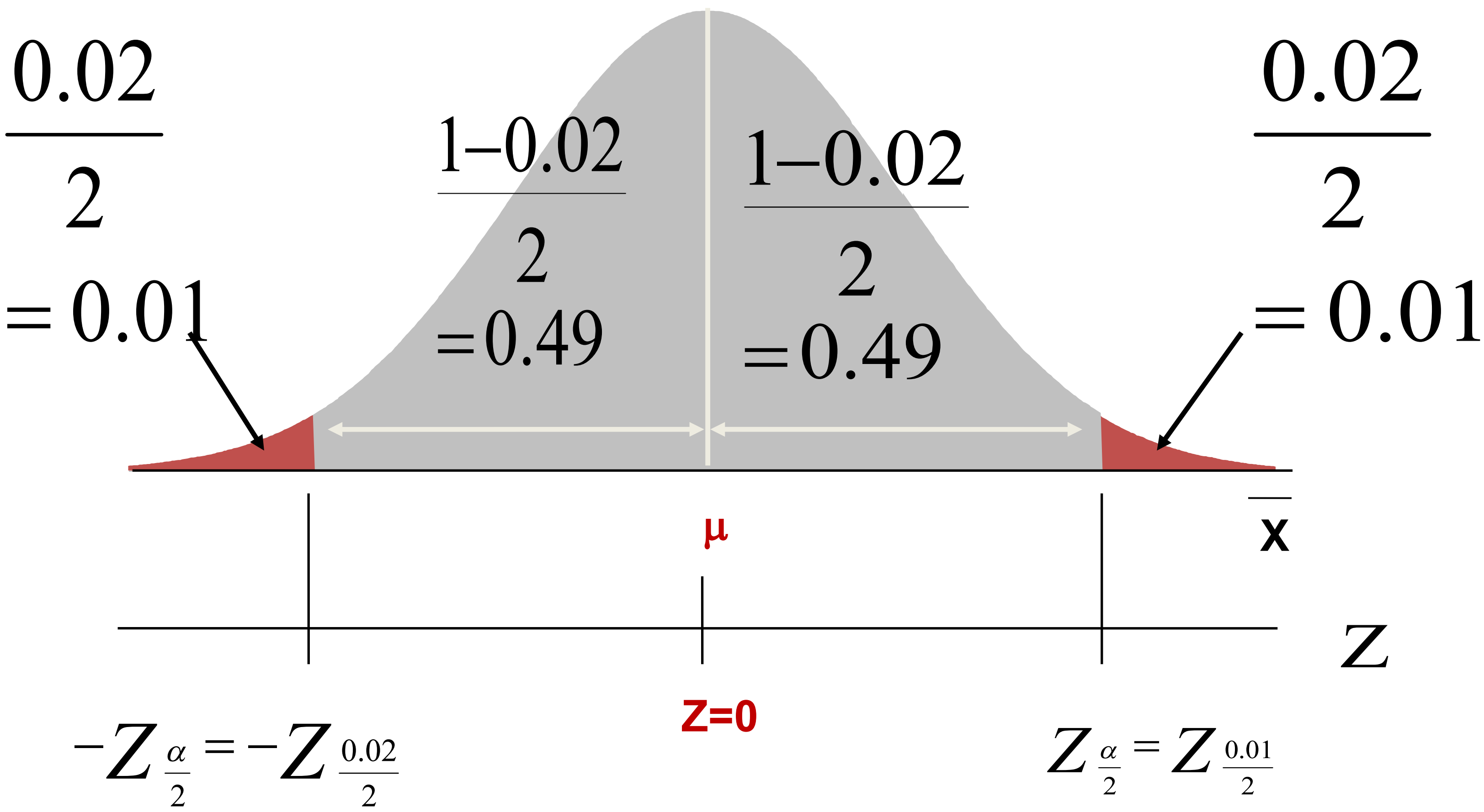
Maximum error estimate

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = z_{\alpha/2}\frac{S}{\sqrt{n}} = (2.575)\left(\frac{2.06}{\sqrt{40}}\right) = 0.8387$$

(b). For 98% confidence ,

$(1-\alpha)=98/100=0.98$ implies $\alpha=0.02$

and $[(1-\alpha)/2]=0.49$.



$\frac{0.02}{2} = 0.01$
$\frac{1-0.02}{2} = 0.49$
$\frac{1-0.02}{2} = 0.49$
$\frac{0.02}{2} = 0.01$
μ
$\overline{X}$
Z
Z=0
$-Z_{\frac{\alpha}{2}} = -Z_{\frac{0.02}{2}}$
$Z_{\frac{\alpha}{2}} = Z_{\frac{0.01}{2}}$

A($Z_{\alpha/2}$)=0.49 implies $Z_{\alpha/2}$=2.33(Approx)

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = (2.33)\left(\frac{2.06}{\sqrt{40}}\right) = 0.758915$$

(Since σ is replaced with S)

98% confidence interval is

$$\left(\overline{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

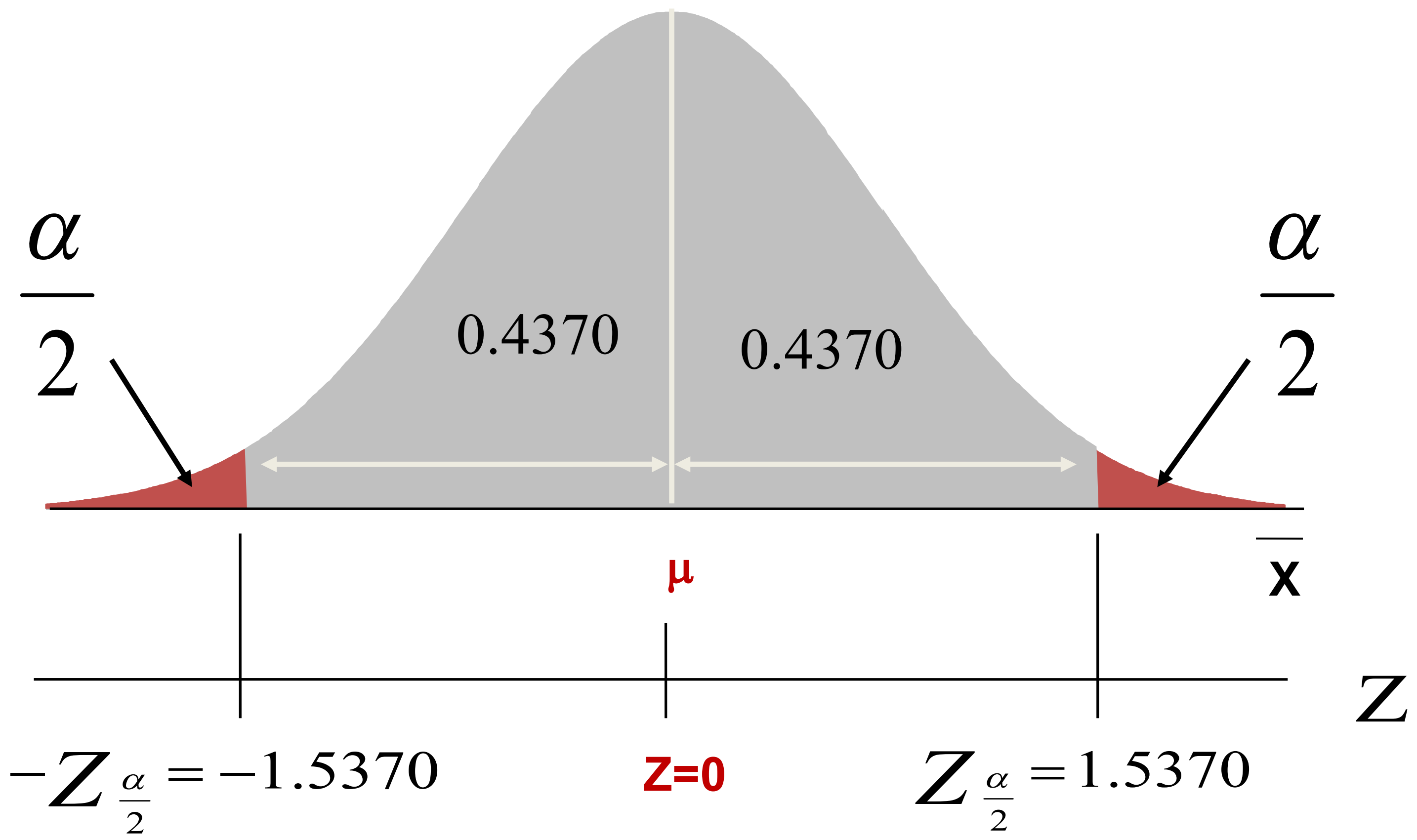
$$\left(12.73 - 0.7589,\ 12.73 + 0.7589\right)$$

(Since σ is replaced with S)

(C) $$30\text{seconds} = \frac{1}{2}\text{minutes} = E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = Z_{\alpha/2}\left(\frac{2.06}{\sqrt{40}}\right)$$

$$(\text{since } \sigma = \text{S})$$

$$\text{implies } \frac{1}{2} = Z_{\alpha/2}\left(\frac{2.06}{\sqrt{40}}\right) \Rightarrow Z_{\alpha/2} = 1.5350$$



$\frac{\alpha}{2}$
0.4370
0.4370
$\frac{\alpha}{2}$
μ
$\bar{x}$
Z
$-Z_{\frac{\alpha}{2}} = -1.5370$
Z=0
$Z_{\frac{\alpha}{2}} = 1.5370$

The area corresponding to $Z_{\alpha/2}$ =1.5350 is 0.4370.

i.e $Z_{\alpha/2}$ =1.5350 gives Area($Z_{\alpha/2}$)=0.4370

Then the area between $-Z_{\alpha/2}$ to $Z_{\alpha/2}$ is $(1-\alpha)=2(0.4370)=0.8740$.

Thus we can ascertain with $(1-\alpha)100\%= 87.4\%$ confidence.

Example: 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 σ=1.40 minutes, determine how large should be the sample size if it is ascertained with 99% confidence that the error E is at most 0.25.

Solution:

For 99% confidence ,(1-α)=99/100 =0.99 implies α=0.01.
And [(1-α)/2]=0.495. A($Z_{\alpha/2}$)=0.495 =(0.4949+0.4951)/2. Implies
$Z_{\alpha/2}$= (2.57+2.58)/2
=2.575

$$n = \left(\frac{Z_{\alpha/2}\sigma}{E}\right)^2 = \left[\frac{(2.575)(1.40)}{0.25}\right]^2 = 207.98 \approx 208$$

Example: Find the degree of confidence to assert that the average salary of school teachers is between Rs.272 and Rs.302 if a random sample of 100 such teachers revealed a mean salary of Rs.287 with standard deviation of Rs.48.

Solution:

Standard variable corresponding to Rs.272 is

Standard variable corresponding to Rs.302 is

Let $\bar{X}$ be the mean salary of teachers, then

$P(272<X\ \bar{X}<302)$

$=P(-3.125<Z<3.125)$

$=2$ [Area under the curve from $Z=0$ to 3.125]

$=2(4.99)$

$=0.9982.$

Thus we can ascertain with 99.82% confidence

Example: A random sample of size 100 is taken from a population with $\boldsymbol{\sigma}$ = 5.1 . Given that the sample mean is $\overline{\boldsymbol{x}}$ = 21.6 Construct a 95% confidence limits for the population mean .

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

$$z\alpha/_2 = 1.96,\ \mathrm{n} = 100\ ,\ \sigma = 5.1$$

$$\therefore \text{Confidence interval} = (\ \bar{x} - z\alpha/_2 . \frac{\sigma}{\sqrt{n}}\ ,\ \bar{x} + z\alpha/_2 . \frac{\sigma}{\sqrt{n}})$$

$$\bar{x} - z\alpha/_2 . \frac{\sigma}{\sqrt{n}} = 21.6 - \frac{1.96 \text{ x } 5.1}{10} = 20.6$$

$$\bar{x} + z\alpha/_2 . \frac{\sigma}{\sqrt{n}} = 21.6 + \frac{1.96 \text{ x } 5.1}{10} = 22.6$$

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

Example: 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 $\boldsymbol{\sigma}$ = 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 , $\sigma = 60$ days and $z\alpha/_2 = 1.645$

$$\therefore n = \left[\frac{z\alpha/_2 . \sigma}{E}\right]^2 = \left[\frac{1.645 \text{ x } 60}{10}\right]^2 = 97$$

Example: 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 $\bar{x}$ 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., $Z_{\alpha/2} = \frac{E_{max} \cdot \sqrt{n}}{\sigma} = \frac{10\sqrt{80}}{62.35} = \frac{89.4427}{62.35} = 1.4345$

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

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

Example: 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 $\mu = 487$, $\sigma = 48$, $n = 100$

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

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

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

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

$$= 2(0 < z < 3.125)$$

$$= 2(0.4991) = 0.9982$$

Thus we can ascertain with 99.82 % confidence.

Example: 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 , $\mu = 11795$

S.D of population , $\sigma = 14054$

$\bar{x} = 11795$

$n =$ sample size $= 50$, maximum error $= Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$Z_{\alpha/2}$ for 95% confidence $= 1.96$

Max. error , $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{14054}{\sqrt{50}} = 3899$

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

$= (11795-3899, 11795+3899)$

$= (7896, 15694)$

Thanks
for
watching
this video