

Module 3 : Statistical Inference 1

Samples : with replacement $\rightarrow N^n$ ways
without replacement $\rightarrow {}^N C_n$ ways

$\{1, 2, 3, 4, 5\} \rightarrow N=5$ $n=2$ (in example)

\hookrightarrow population

\hookrightarrow taking 2 entries from population

* without replacement : $(1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5)$

$\mu_{\bar{x}} \rightarrow$ mean of samples

$\sigma_{\bar{x}}^2 \rightarrow$ variance of population samples

$\sigma_{\bar{x}} \rightarrow$ SD of samples

$\mu \rightarrow$ mean of population

$\sigma^2 \rightarrow$ variance of population

$\sigma \rightarrow$ SD of population

$\{1, 2, 3, 4, 5\} : \mu=3 \quad \sigma^2=2 \quad \sigma=\sqrt{2}$

* $\begin{matrix} 1.5 & 2 & 2.5 & 3 & 2.5 & 3 & 3.5 & 3.5 & 4 & 4.5 \\ (1, 2) & (1, 3) & (1, 4) & (1, 5) & (2, 3) & (2, 4) & (2, 5) & (3, 4) & (3, 5) & (4, 5) \end{matrix}$ mean of each entries

$\mu_{\bar{x}} = \frac{1.5 + 2 + 2.5(2) + 3(2) + 3.5(2) + 4(1) + 4.5(1)}{10} \rightarrow$ no. of times it occurs.
 $10 \rightarrow$ no. of entries

$\rightarrow \mu_{\bar{x}} = 3$

$$\sigma_{\bar{x}}^2 = \frac{(-1.5)^2 + (-0.5)^2 + (1.5)^2 + (-1)^2 + 0^2 + (-0.5)^2 + (-0.5)^2 + 0^2 + (0.5)^2 + (1)^2}{10}$$

$$\sigma_{\bar{x}}^2 = 0.75 \rightarrow \sigma_{\bar{x}} = \sqrt{0.75}$$

\bar{x}_i	F_i	$\bar{x}_i F_i$	$(\bar{x}_i - \mu_{\bar{x}})^2$	$(\bar{x}_i - \mu_{\bar{x}})^2 F_i$
1.5	1	1.5	2.25	2.25
2	1	2	1	1
2.5	2	5	0.25	0.5
3	2	6	0	0
3.5	2	7	0.25	0.5
4	1	4	1	1
4.5	1	4.5	2.25	2.25
	<u>10</u>	<u>30</u>		<u>7.5</u>

$$\mu_{\bar{x}} = \frac{\sum \bar{x}_i F_i}{\sum F_i} = \frac{30}{10} = 3$$

$$\sigma_{\bar{x}}^2 = \frac{\sum (\bar{x}_i - \mu_{\bar{x}})^2 F_i}{\sum F_i} = \frac{7.5}{10} = 0.75$$

With replacement

Without replacement

① N^n ways

${}^N C_n$ ways

② $\mu = \mu_{\bar{x}}$

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

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

$\frac{(N-n)}{N-1} \frac{\sigma^2}{n} = \sigma_{\bar{x}}^2$

3-1-24 Q. A population consists of 5 nos {2, 3, 6, 8, 12}. Consider all possible samples of size 2 that can be drawn with replacement from this population.

i. Find μ & σ (population)

ii. Find the $\mu_{\bar{x}}$ & $\sigma_{\bar{x}}$ (sample means)

iii. Considering samples without replacement, find $\mu_{\bar{x}}$ & $\sigma_{\bar{x}}$

iv. Verify $\mu = \mu_{\bar{x}}$ & $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ (with replacement)

v. Verify $\mu = \mu_{\bar{x}}$ & $\sigma_{\bar{x}}^2 = \frac{(N-n)}{N-1} \frac{\sigma^2}{n}$ (without replacement)

Ans. population: {2, 3, 6, 8, 11}, $N=5$ & $n=2$

i. μ & σ

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

$$\sigma^2 = \frac{16+9+0+4+25}{5} = \frac{54}{5} = 10.8 \rightarrow \sigma = \sqrt{10.8} = 3.28$$

ii. With replacement: N^n ways = $5^2 = 25$ samples

$$\mu_{\bar{x}} = \mu = 6$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{10.8}{2} = 5.4 \rightarrow \sigma_{\bar{x}} = 2.32$$

iii. Without replacement: ${}^N C_n$ ways = ${}^5 C_2 = 10$ samples

$$\mu_{\bar{x}} = \mu = 6$$

$$\sigma_{\bar{x}}^2 = \frac{(N-n)}{N-1} \frac{\sigma^2}{n} = \frac{(5-2)}{5-1} \frac{10.8}{2} = \frac{3}{4} (5.4) = 4.05$$

$$\sigma_{\bar{x}} = \sqrt{4.05} = 2.012$$

iv. 10 samples: {2, 3}, {2, 6}, {2, 8}, {2, 11}, {3, 6}, {3, 8}, {3, 11}, {6, 8}, {6, 11}, {8, 11}

$$\mu_{\bar{x}} = \frac{2.5+4+5+6.5+4.5+5.5+7+7+8.5+9.5}{10} = \frac{60}{10} = 6$$

$$\sigma_{\bar{x}}^2 = \frac{12.25+4+1+0.25+2.25+0.85+1+1+6.25+12.25}{10} = \frac{40.5}{10} = 4.05$$

\Rightarrow from iii. $\mu_{\bar{x}} = 6 = \mu$

$$\sigma_{\bar{x}}^2 = 4.05 \rightarrow \sigma_{\bar{x}}^2 = \frac{10.8}{2} \frac{(N-n)}{N-1} = \frac{10.8}{2} \left(\frac{3}{4} \right) = 4.05$$

iv. Table banao aur karo :

Q. Certain cubes manu. by company have mean life of 800 hrs & SD of 60 hrs. Find probability that a random sample with replacement of 16 cubes will have a mean lifetime

i. b/w 790 & 810 hrs

ii. less than 785 hrs

iii. more than 820 hrs

iv. b/w 770 & 830 hrs

Ans $\mu = 800$ hrs $\sigma = 60$ hrs $n = 16$

$$\mu_{\bar{x}} = 800$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{4} = 15$$

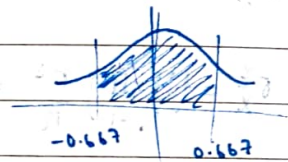
$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - 800}{15}$$

i. $P(790 < \bar{x} < 810)$

$$\bar{x} = 790 \Rightarrow z_1 = \frac{790 - 800}{15} = -\frac{10}{15} = -0.667$$

$$\bar{x} = 810 \Rightarrow z_2 = \frac{810 - 800}{15} = \frac{10}{15} = 0.667$$

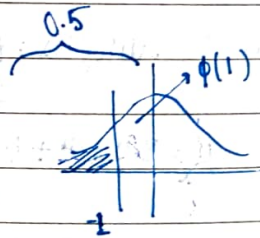
$$\therefore P(-0.667 < z < 0.667) = 2\phi(0.667) = 2(0.2486) = \underline{0.4972}$$



ii. $P(\bar{x} < 785)$

$$\bar{x} = 785 \Rightarrow z_1 = \frac{785 - 800}{15} = -\frac{15}{15} = -1$$

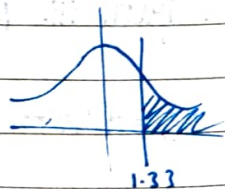
$$P(z < -1) = 0.5 - \phi(1) = 0.5 - 0.2413 = \underline{0.2587}$$



iii. $P(\bar{x} > 820)$

$$\bar{x} = 820 \Rightarrow z = \frac{820 - 800}{15} = \frac{20}{15} = 1.334$$

$$P(z > 1.334) = 0.5 - \phi(1.33) = 0.4082 = \underline{0.0918}$$



iv. $P(770 < \bar{x} < 830)$

$$\bar{x} = 770 \Rightarrow z_1 = \frac{770 - 800}{15} = -2$$

$$\bar{x} = 830 \Rightarrow z_2 = \frac{830 - 800}{15} = 2$$

$$P(-2 < z < 2) = 2\phi(2) = 2(0.4772) = \underline{0.9544}$$

Q. The weights of 1500 ball bearings are distributed with $\mu = 635$ g & $\sigma = 1.36$ g. If 300 samples of size 36 are drawn from this & population.

- Find $\mu_{\bar{x}}$ & $\sigma_{\bar{x}}$ if sampling is done WR & WoR.
- If sampling is done WR, find how many random samples would have their mean:

(a) b/w 634.76 & 635.24 g

(b) greater than 635.5 g

(c) less than 634.2 g.

Ans. $N = 1500$ $\mu = 635$ g $\sigma = 1.36$ g $n = 36$

i. WR: $\mu_{\bar{x}} = \mu = 635$ g

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{1.36^2}{36} = \frac{1.8496}{36} = 0.05137$$

$$\sigma_{\bar{x}} = 0.226$$

WoR: $\mu_{\bar{x}} = \mu = 635$

$$\sigma_{\bar{x}}^2 = \frac{(N-n)}{N-1} \frac{\sigma^2}{n} = \frac{(1500-36)}{1500-1} \frac{1.8496}{36} = \underline{0.224}$$

ii. (a) $P(634.76 < \bar{x} < 635.24)$

$$\bar{x} = 634.76 \Rightarrow z_1 = \frac{634.76 - 635}{1.36 \cdot 0.226} = \frac{-0.24}{0.30736} = -0.781$$

$$\bar{x} = 635.24 \Rightarrow z_2 = \frac{635.24 - 635}{1.36 \cdot 0.226} = \frac{0.24}{0.30736} = 0.781$$

$$\therefore P(-0.781 < z < 0.781) = 2\phi(0.781) = 2 \times 0.3554 = 0.7108$$

$$\therefore \text{No. of samples} = 300 \times 0.7108 = \underline{213}$$

(b) $P(\bar{x} > 635.5)$

$$\bar{x} = 635.5 \Rightarrow z = \frac{635.5 - 635}{1.36 \cdot 0.226} = \frac{0.5}{0.30736} = 1.627$$

$$\therefore P(z > 2.212) = 0.5 - \phi(2.212)$$

$$\therefore \text{No. of samples} = 300 \times$$

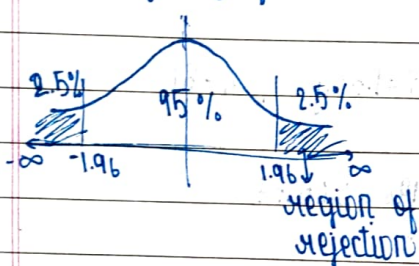
$$c) P(\bar{x} < 634.2)$$

$$\bar{x} = 634.2 \Rightarrow z_1 = \frac{634.2 - 635}{0.226} = -3.54$$

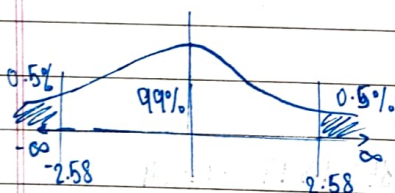
$$P(z < -3.54) = 0.5 - \phi(3.54)$$

5-01-2024

Level of Significance



95% of area under bell curve
5% : ~~of~~ level of significance
 $z_{0.05}$



99%
1% level of significance
 $z_{0.01}$

$$|z| \leq 1.96 \Rightarrow -1.96 \leq z \leq 1.96$$

$$\Rightarrow -1.96 \leq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq 1.96 \Rightarrow -1.96 \leq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq 1.96$$

$$\Rightarrow (-1.96) \frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq (1.96) \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow (-1.96) \frac{\sigma}{\sqrt{n}} - \bar{x} \leq -\mu \leq \frac{\sigma}{\sqrt{n}} (1.96) - \bar{x}$$

$$\Rightarrow (1.96) \frac{\sigma}{\sqrt{n}} + \bar{x} \geq \mu \geq \frac{\sigma}{\sqrt{n}} (-1.96) + \bar{x}$$

$$-1.96 \frac{\sigma}{\sqrt{n}} + \bar{x} \leq \mu \leq 1.96 \frac{\sigma}{\sqrt{n}} + \bar{x} \Rightarrow \mu = \pm 1.96 \frac{\sigma}{\sqrt{n}} + \bar{x}$$

confidence limits.

Q. A random sample of 400 items chosen from an infinite pop. is found to have $\bar{x}=82$ & $\sigma_x=18$. Find 95% confidence limits of mean of population. Also find 99% confidence limits

Ans. $\mu = \bar{x} = 82$ $\sigma_x = 18 = \sigma$ $n = 400$ $\bar{x} = 82$
 $\mu = \pm 1.96 \left(\frac{18}{\sqrt{400}} \right) + 82$

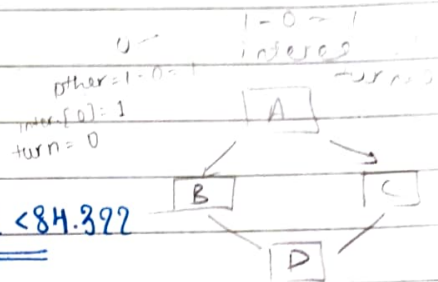
\bar{x} = sample mean
 (mean of one of the samples)
 sample of mean μ
 (mean of all the means of samples)

$\Rightarrow \mu = 80.236$ or $83.764 \Rightarrow 80.236 < \mu < 83.764$

99%:

$\mu = \pm 2.58 \left(\frac{18}{\sqrt{400}} \right) + 82$

$\mu = 79.678$ or $84.322 \Rightarrow 79.678 < \mu < 84.322$



Q. The life of a certain computer is approx. normally dist. with mean 800 hrs & SD of 40 hrs. If a random sample of 30 computers has an avg life 788 hrs. Test the hypothesis that avg life is 800 hrs at 5% LOS & 1% LOS.

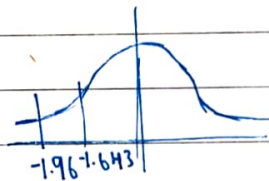
Ans. $\mu = 800$ $\sigma = 40 = \sigma_x$ $\bar{x} = 788$ $n = 30$
 5% LOS:

$\mu = \pm 1.96 \left(\frac{40}{\sqrt{30}} \right) + 788$

$\mu = \pm 2.58 \left(\frac{40}{\sqrt{30}} \right) + 788$
 $769.15 < \mu < 806.84$

$\Rightarrow \mu = 802.313$ or 773.686

$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{788 - 800}{40/\sqrt{30}} = -1.643$



95% accept and 5% reject

99% accept and 1% reject

