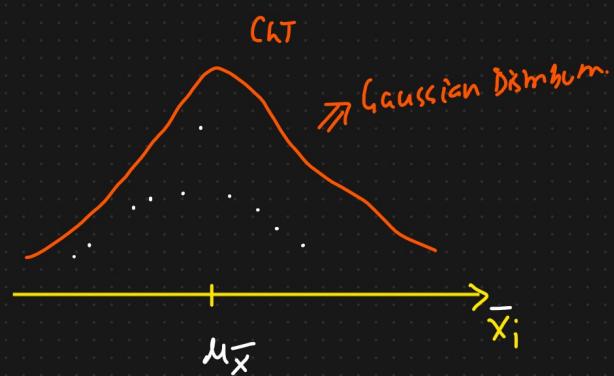
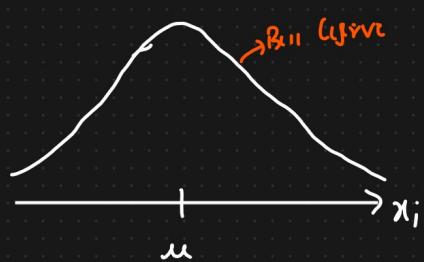


Agenda

- ① Central Limit Theorem ✓
- ② Z-score And Z-stats ✓
- ③ Z-test, t-test { Solve problems Assignment } Hypothesis Testing.

① Central Limit Theorem

$$① X \sim N(\mu, \sigma)$$

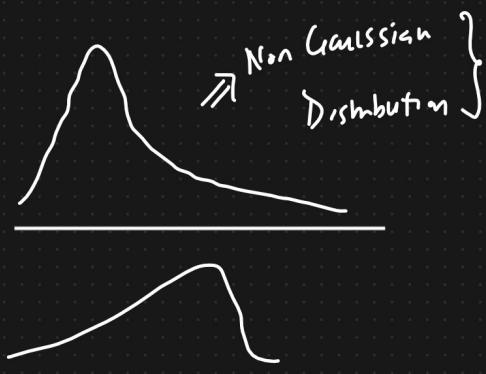


$$\begin{aligned} S_1 &= \{x_1, x_2, x_3, \dots, x_n\} = \bar{x}_1 \text{ of the } n \text{ samples} \\ S_2 &= \{x_2, x_3, x_4, x_5, \dots, x_n\} = \bar{x}_2 \text{ of the } n \text{ samples} \\ S_3 &= \{ \dots \} = \bar{x}_3 \\ S_4 &= \{ \dots \} = \bar{x}_4 \\ S_5 &= \{ \dots \} = \bar{x}_5 \\ &\vdots \\ S_m &= \{ \dots \} = \bar{x}_m \end{aligned}$$

n = Sample size

$$n \geq 30$$

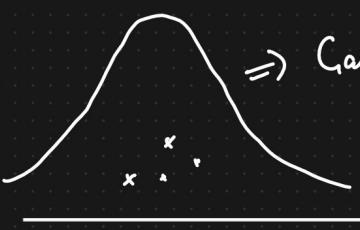
$$② X \not\sim N(\mu, \sigma)$$



$$\begin{aligned} S_1 &= \{ \dots \} = \bar{x}_1 \\ S_2 &= \{ \dots \} = \bar{x}_2 \\ S_3 &= \{ \dots \} = \bar{x}_3 \\ S_4 &= \{ \dots \} = \bar{x}_4 \\ &\vdots \\ S_m &= \{ \dots \} = \bar{x}_m \end{aligned}$$

1, 2, 3, 4, 5

\Downarrow
 $n = \text{sample size} \Rightarrow \underline{\text{any value}}$



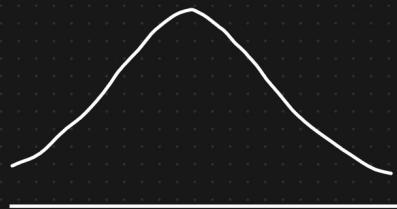
\Rightarrow Gaussian/ Normal Distribution

(mean will follow gaussian distribution)

① Central Limit Theorem

The central limit theorem relies on the concept of a sampling distribution, which is the probability distribution of a statistic for a large number of samples taken from a population.

The central limit theorem says that the sampling distribution of the mean will always be normally distributed, as long as the sample size is large enough. Regardless of whether the population has a normal, Poisson, binomial, or any other distribution, the sampling distribution of the mean will be normal.

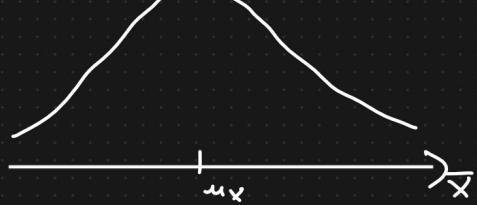


$$X \sim N(\mu, \sigma)$$

σ = population std

μ = population mean

Sampling Distribution of the mean



$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

standard

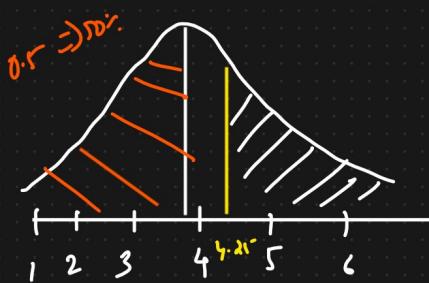
error.

n = Sample size

①

$$X \sim N(4, 1)$$

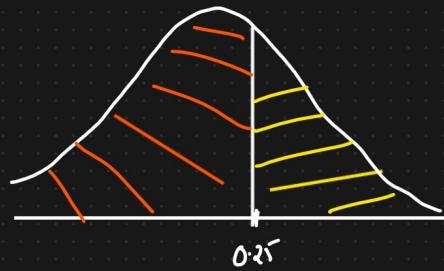
$$\lambda_i = 4.25$$



$$Z_{score} = \frac{4.25 - 4}{\sigma} = 0.25$$

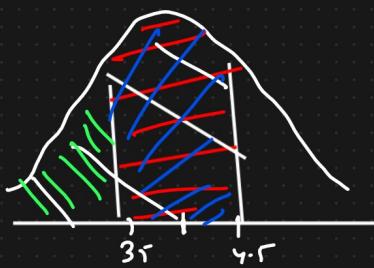
Download Z-table

Q) What percentage of score falls above 4.25?



$$1 - 0.59871 = 0.4013 = 40.13\%$$

Q) What percentage of score lies between 3.5 to 4.5?

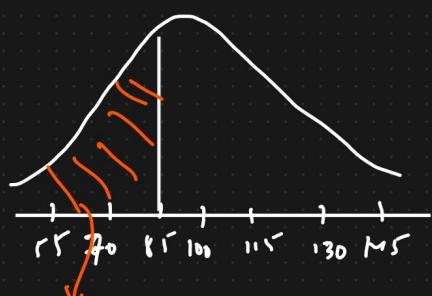


$$Z\text{-Score} = \frac{4.5 - 4}{1} = +0.5 = 0.69146$$

$$\begin{aligned} Z\text{-Score} &= \frac{3.5 - 4}{1} = -0.5 = 0.30854 \\ &= 0.3829 \\ &= 38.29\% \end{aligned}$$

Q) In India the average IQ is 100, with a standard deviation of 15. What is the percentage of the population would you expect to have an IQ lower than 85?

Ans). $\mu = 100$ $\sigma = 15$



$$0.1586 = 15.86\%$$

$$Z\text{-Score} = \frac{85 - 100}{15} = -1$$

Q) $IQ > 85$

$$1 - 0.1586 = 84.13\%$$

$| Z >, IQ \leq 100 | \Rightarrow \text{Internal Assignment}$

① Hypothesis Testing And Statistical Analysis

- ① Z-test
- ② t-test
- ③ Chi Square
- ④ ANOVA

① Z-test

With a $\sigma = 3.9$

①) The average heights of all residents in a city is 168cm. A doctor believes the mean to be different. He measured the height of 36 individuals and found the average height to be 169.5 cm.

- (a) State null and Alternate Hypothesis
- (b) At a 95% confidence level, is there enough evidence to reject the null hypothesis.

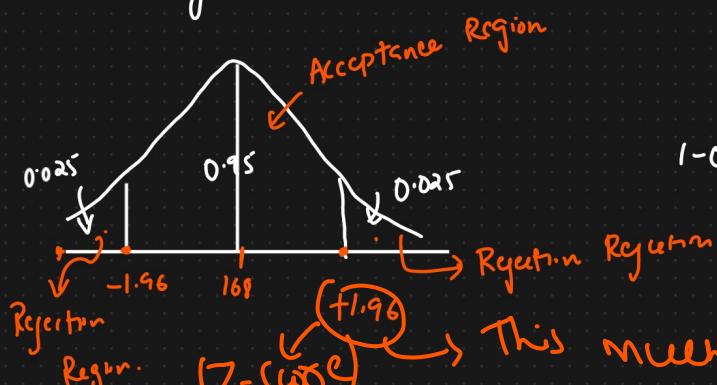
Ans) $\mu = 168 \text{ cm}$ $\sigma = 3.9$ $n = 36$ $\bar{x} = 169.5$ ($I = 0.95$) $\alpha = 1 - C.I$
 $= 0.05\%$.

① Null Hypothesis $H_0 : \mu = 168 \text{ cm}$

Alternate Hypothesis $H_1 : \mu \neq 168 \text{ cm}$

1 tailed test
 If mean is greater than 168

② Decision Boundary band C.I.



2 Tailed Test

$$1 - 0.025 = 0.9750$$

This much Standard deviation away from Mean (Z-table)

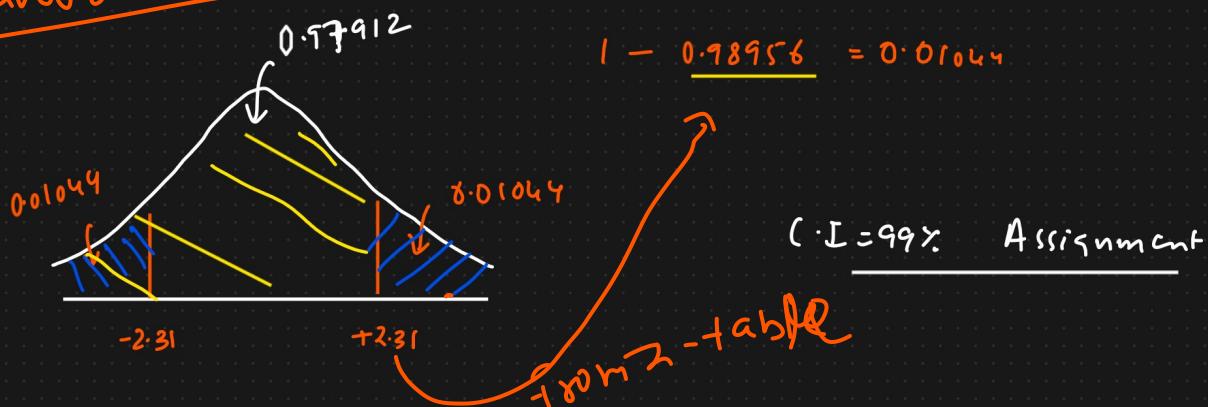
If Z-test is less than -1.96 or greater than +1.96, Reject the Null Hypothesis

$$\textcircled{4} \quad Z\text{-test} = \frac{\bar{x} - \mu}{\left\{ \sigma / \sqrt{n} \right\}} = \frac{169.5 - 168}{3.9 / \sqrt{36}} = \boxed{2.31}$$

\downarrow
 $\left\{ \text{CLT} \right\}$

Conclusion

$|2.31| > 1.96$ Reject the Null Hypothesis. \rightarrow Do Q&P is current method



$$\textcircled{1} \quad P\text{-Value} = 0.01044 + 0.01044 \\ = 0.02088$$

$P < 0.05 \Rightarrow$ Reject the Null Hypothesis.

(2) A factory manufactures bulbs with an average warranty of 5 years

with standard deviation of 0.50. A worker believes that the bulb will malfunction in less than 5 years. He tests a sample of 40 bulbs and finds the average time to be 4.8 years.

(a) State null and alternate hypothesis

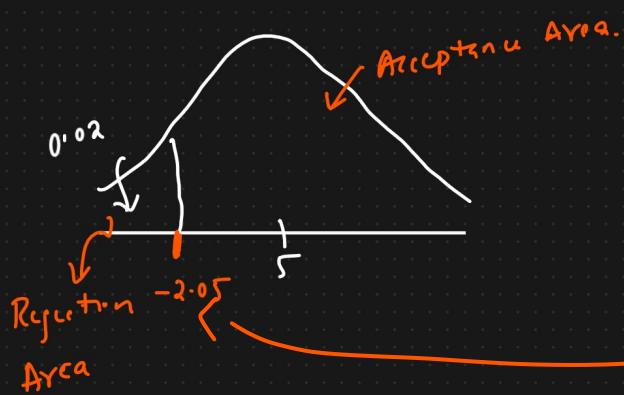
(b) At a 2% significance level, is there enough evidence to support the idea that the warranty should be revised?

$$\text{Ans) } \mu = 5 \quad \sigma = 0.50 \quad n = 40 \quad \bar{x} = 4.8 \text{ years.} \quad C.I = 0.98 \quad d = 0.02$$

① $H_0: \mu = 5$

$H_1: \mu < 5$ {1 Tail Test}.

② Decision Boundary



Z_{test}

Look for 0.02 in
Z-Score and
find corresponding

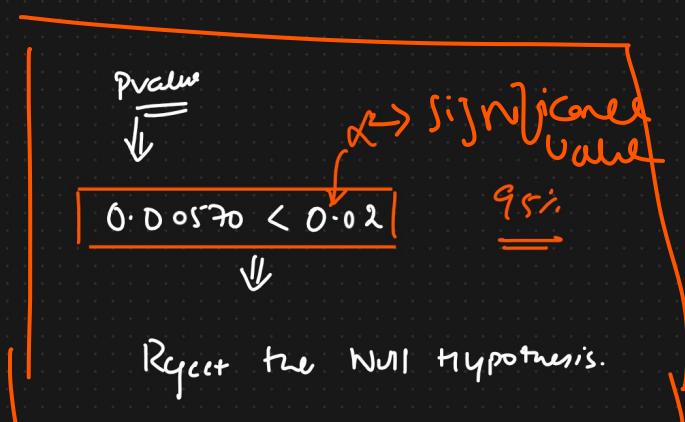
$Z_{\text{test}} < -2.05$ Reject the Null hypothesis.

$$④ Z_{\text{test}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{4.8 - 5}{0.50 / \sqrt{40}} = [-2.53]$$

$Z_{\text{test}} < -2.5$ This value
is just

Reject
Null hypothesis

$-2.53 < -2.05 \Rightarrow \text{True} \Rightarrow \text{Reject the Null hypothesis.}$



② T Test

DATA ANALYST

① In the population the average IQ is 100. A team of researchers want to test a new medication to see if it has either a positive or negative effect on intelligence, or no effect at all. A sample of 30 participants who have taken the medication has a mean of 140 with a standard deviation of 20. Did the medication affect intelligence? CI = 95% $\alpha = 0.05$

$$\text{H}_0: \mu = 100 \quad n=30 \quad \bar{x}=140 \quad s=20 \quad CI=95\% \\ \text{H}_1: \mu \neq 100 \quad \alpha = 0.05$$

$$② \quad \alpha = 0.05$$

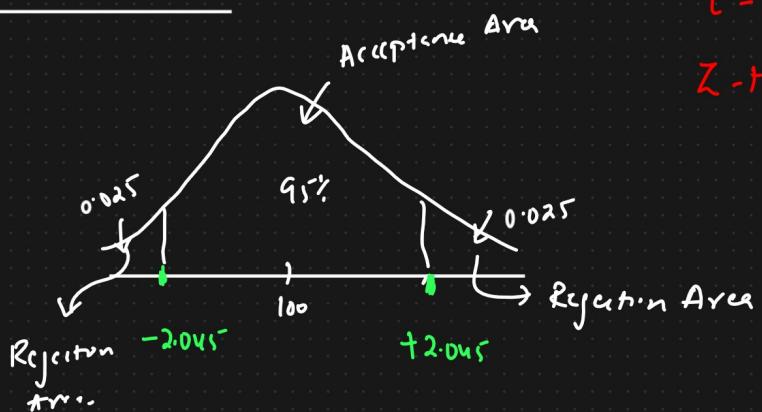
Degree of freedom

$$3 \text{ people} \quad . \quad 3-1 = 2$$

$$dof = n-1 = 30-1 = 29, \dots$$



③ Decision Rule



$t_{\text{test}} \Rightarrow$ sample std.

$Z_{\text{test}} \Rightarrow$ population std

use T -Table

If t_{test} is less than -2.045 and greater than $+2.045$, Reject the

Null Hypothesis.

④ T Test Statistics

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{140 - 100}{20/\sqrt{30}} = \frac{40}{3.65} = 10.96$$

$t > 2.045$ Reject the Null Hypothesis.

Conclusion : Medication has a true effect on intelligence.



$$\boxed{2000}$$

$$\boxed{200 - 300} \Rightarrow$$

Z-test

Population σ is given

T-test

Sample σ is given