Agenda

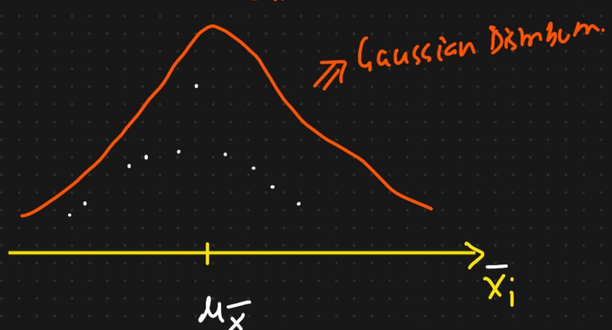
1. Central Limit Theorem ✓
2. Z-score And z stats ✓
3. Z-test, t-test } Solve problems Assignment

} Infurential stats
Hypothesis Testing.

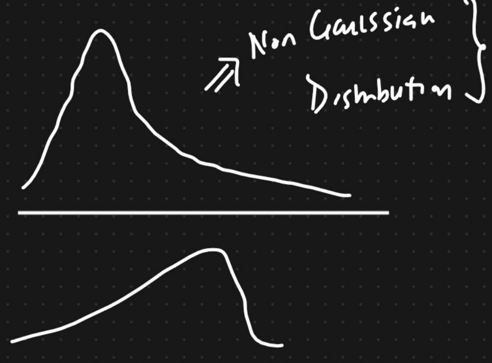
① Central Limit Theorem

① $X \approx N(\mu, \sigma)$

CLT



② $X \not\approx N(\mu, \sigma)$



1, 2, 3, 4, 5

n = sample size => any value

$S_1 = \{x_1\ x_2\ x_3 \ldots x_n\} = \bar{x}_1$

$S_2 = \{x_2, x_3, x_4, x_6 \ldots x_n\} = \bar{x}_2$

$S_3 = \{\quad\} = \bar{x}_3$

$S_4 = \{\quad\} = \bar{x}_4$

$S_5 = \{\quad\} =$

$\vdots$

$S_m = \{\quad\} \quad \bar{x}_m$

n = sample size

$\boxed{n \geq 30}$

$S_1 = \{\quad\} = \bar{x}_1$

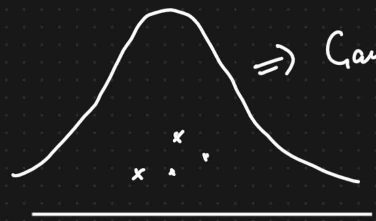
$S_2 \quad = \bar{x}_2$

$S_3 \quad = \bar{x}_3$

S_4

$\vdots$

$\bar{x}_m$

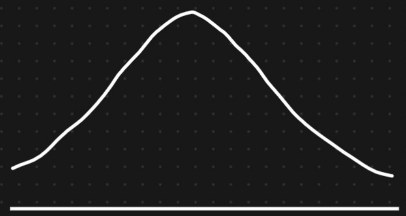


⇒ Gaussian/ Normal Distribution

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.

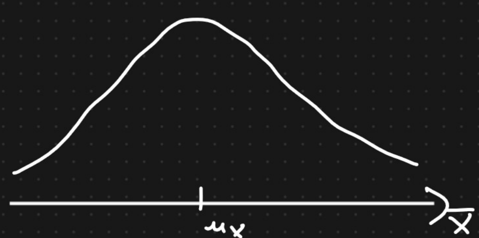


$X \sim N(\mu, \sigma)$

σ = population std

μ = population mean

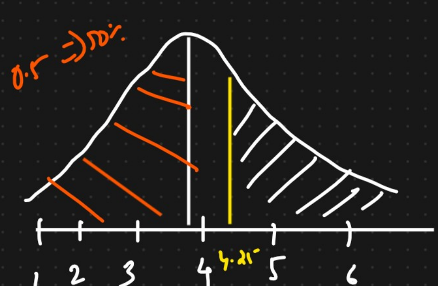
Sampling Distribution of the mean



$$\bar{X} \sim N\left(\mu, \underbrace{\frac{\sigma}{\sqrt{n}}}_{\text{Standard error}}\right)$$

n = Sample size

① $X \sim N(4, 1)$



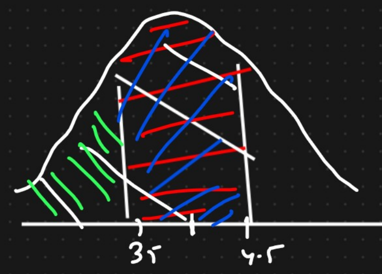
$x_i = 4.25$

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

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?



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

$$Z\text{-score} = \frac{3.5-4}{1} = -0.5 = 0.30854$$

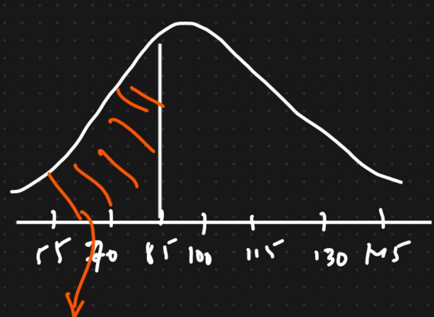
$= 0.3829$

$= 38.29\%$

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

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



$0.15866 = 15.86\%$

(A) IQ ≥ 85

$1 - 0.1586 = 84.13\%$

$\boxed{75 \geq IQ \leq 100}$ $\Rightarrow$ Internal Assignment

① Hypothesis Testing And Statistical Analysis

① Z test
② t test
③ CHI SQUARE
④ ANNOVA

① Z-test

① 1) The average heights of all residents in a city is 168cm, with a $\sigma = 3.9$. A doctor believes the mean to be diffrent. He measured the height of 36 individuals and found the average height to be 169.5 cm.

(a) State null and Alternate Hypothesis

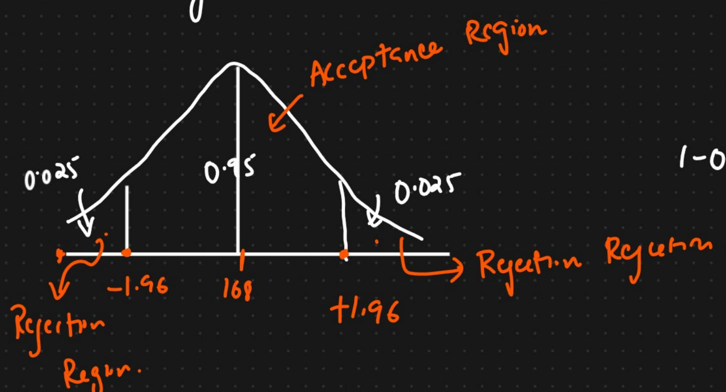
(b) At a 95% confidence level, is there enough evidence to reject the null hypothesis.

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

① Null Hypothesis H_0 : $\mu = 168cm$

Alternate Hypothesis H_1 : $\mu \neq 168cm$

② Decision Boundary band C.I.



2 Tailed Test

$1 - 0.025 = 0.9750$

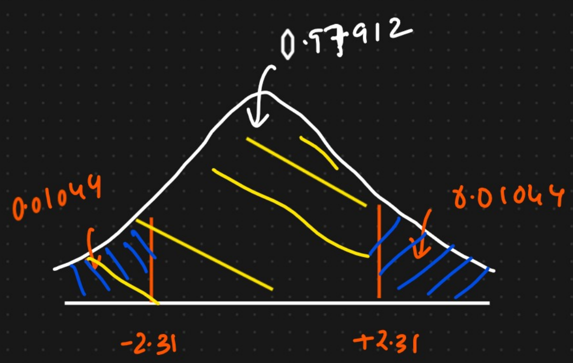
If Z test is less than −1.96 or greater than +1.96, Reject the Null Hypothesis

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

$\{\sigma/\sqrt{n}\}$ ⇓ {CLT}

Conclusion

$2.31 > 1.96$ Reject the Null Hypothesis.



$1 - 0.98956 = 0.01044$

C.I = 99% Assignment

① p value = 0.01044 + 0.01044

= 0.02088

$P < 0.05$ ⇒ Reject the Null Hypothesis.

② A factory manufactures bulbs with a 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 find the average time to be 4.8 years.

(a) State null and alternate hypothesis

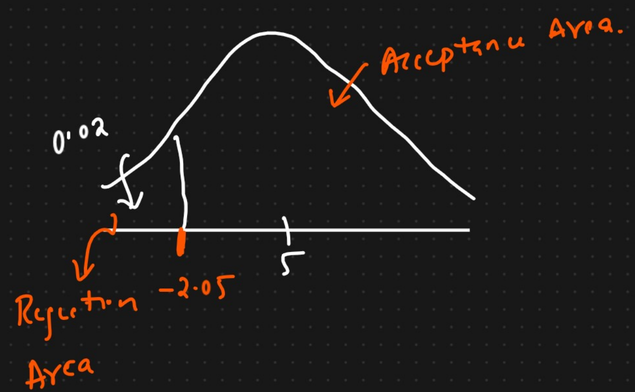
(b) At a 2% significance level, is there enough evidence to support the idea that the warranty should be revised?

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

① H_0 $\mu = 5$

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

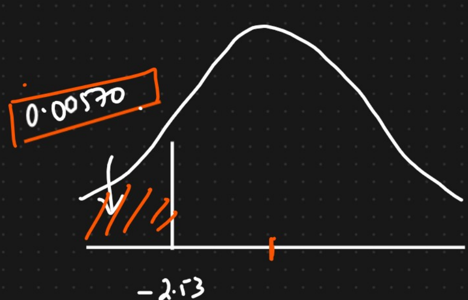
② Decision Boundary



Z test < −2.05 Reject the Null hypothesis.

(*) $$Z.test = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{4.8 - 5}{0.50 / \sqrt{40}} = \boxed{-2.53}$$

$-2.53 < -2.05$ ⇒ True ⇒ Reject the Null hypothesis.



Pvalue

$\Downarrow$

α

$\boxed{0.00570 < 0.02}$ 95%

$\Downarrow$

Reject the Null Hypothesis.

② T Test

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? C.I = 95% $\alpha = 0.05$

Ans) ① $H_0 = \mu = 100$

$H_1 = \mu \neq 100$

$n = 30 \quad \mu = 100 \quad \bar{x} = 140 \quad s = 20 \quad C.I = 95$

$\alpha = 0.05$

② $\alpha = 0.05$

Degree of freedom

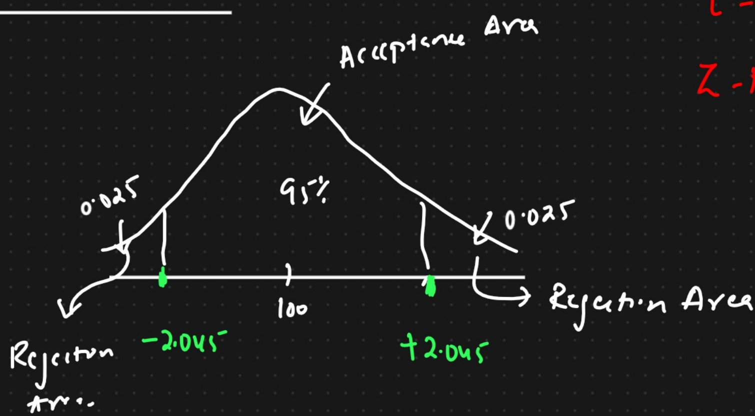
3 people $\qquad 3 - 1 = 2$

$dof = n - 1 = 30 - 1 = 29$

③ Decision Rule

t - test ⇒ Sample std.

Z - test ⇒ Population std



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 +ve effect on intelligence.

BANK A

A

B

95% 99%

Hypothesis Testing

$\boxed{2000}$

$\boxed{200 - 300} \Rightarrow$