1. Hypothesis testing
2. Ztest, ttest, chisqure, Anova(f-test)

$$Z\text{-Score} = \frac{D.P - \text{mean}}{Std}$$

C.I ⇒ Confidence interval

(α) Significance value ⇒ 1 - C.I.

1. Null hypothesis (Acceptance and rejection)
2. Alternate hypothesis
3. P_value

Sample | Population

Q. the avg height of all residents in a city is 168 cm with a $\sigma = 3.9$. [A doctor believes the mean to different. He measured the height of 36 People and found the avg height to be 169.5 cm.]

1) State Null and Alternative hypothesis.
2) At 95% CI, is there enough evidence to reject the Null hypothesis

CLT

$\mu = 168$ cm., $\sigma = 3.9$,
$n = 36$, $\bar{x} = 169.5$.

CI = 95%, $\alpha = 5\%$ $(1 - CI)$

graph

H_0 = Residents in a City is 168 cm. $\Rightarrow \mu = 168$ cm.

H_1 = Residents in a city is not 168 cm $\Rightarrow \mu \neq 168$ cm.

$\mu <$
$\mu >$

$H_0 \Rightarrow \mu = 168$ cm

$H_1 \Rightarrow \mu \neq 168$ cm. $(< 168, > 168)$

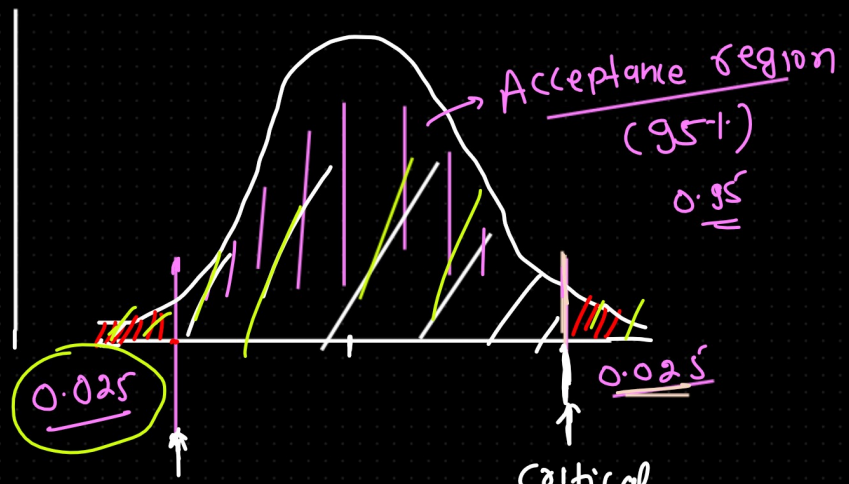
Assumption $\Rightarrow$ Data should be normally Distributed.

Complete area (100% ⇒ 1)

$\mu = 168$

$\mu \neq 168$ $(\mu < 168, \mu > 168)$

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



1 − 0.025

0.975

AUC

1 − 0.025 = 97.5

Rejection region = 5%

$\left(\frac{0.05}{2}\right)$

-1.96 ← → $+1.96$

2_score ⇒ $\frac{169.5 - 168}{3.9} =$

$\Rightarrow \frac{169.5 - 168}{3.9/\sqrt{36}}$ ← S.E ← CLT

$\Rightarrow \frac{1.5}{0.65}$

$\Rightarrow 2.31$

$2.31 > 1.96$

95-1.

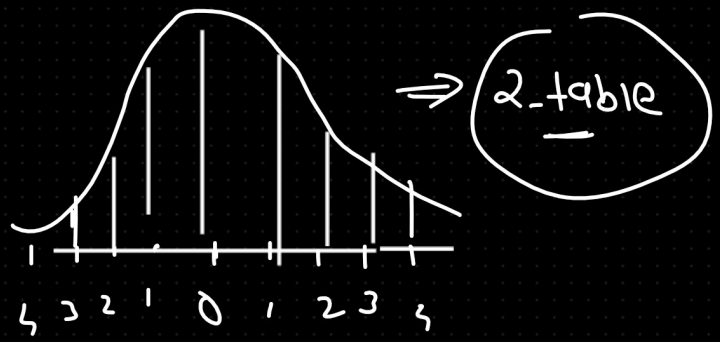
$\mu \neq 168$

(reject it)

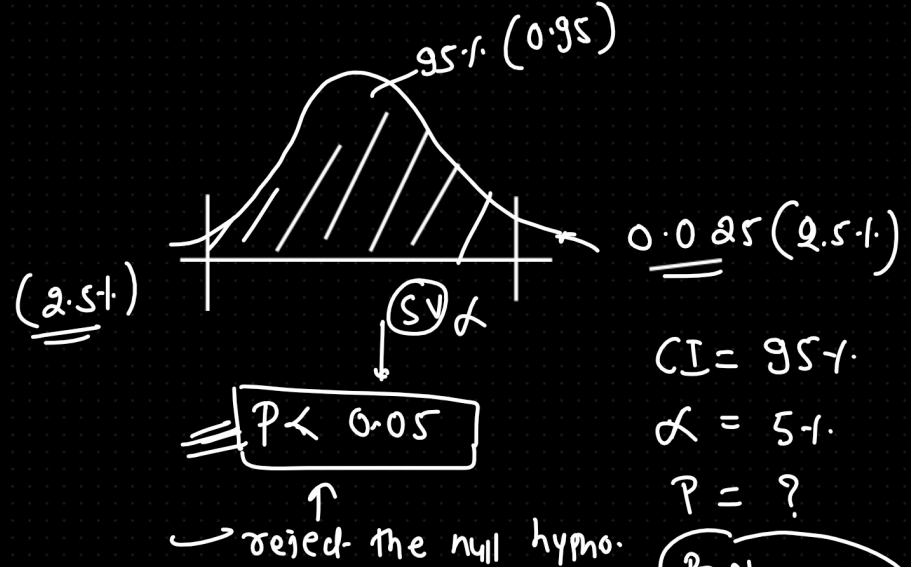
$\mu_0 >$ $\mu_0 <$

$\mu_0 > 168$

{ Reject the null hypothesis
  Accept the alternate hypothesis



P-value ⇒



CI = 97% $\quad \alpha = 3\%$

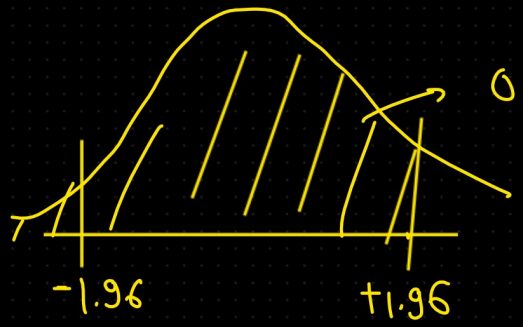
$P \le \alpha$ ↓

$P \le 0.03$

CI = 99%

$\alpha = 1\%$

$\Rightarrow P \le 0.01$



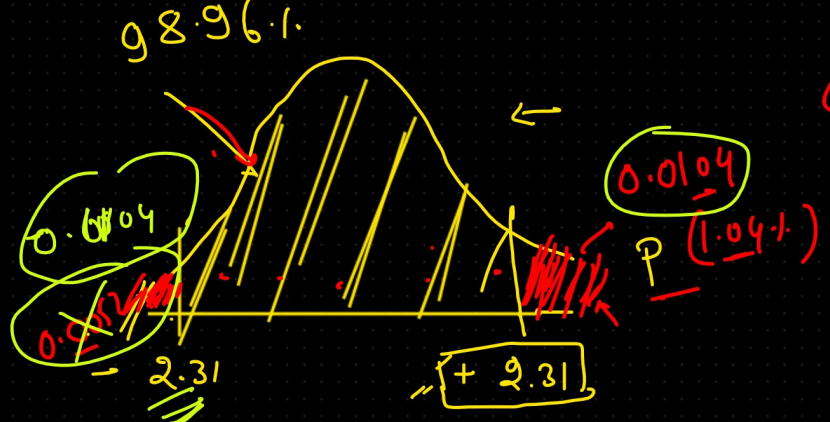
my_2_score = 2.31

CI = 95 (0.95)

$\alpha = 0.05$

$\alpha = 0.025 \Leftarrow P$

$\Rightarrow P < \alpha (0.05)$

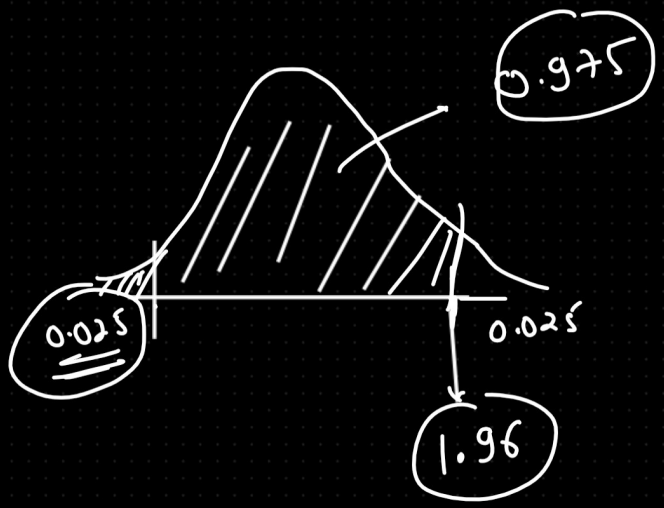


0.9896 ⇒ 98.96%

$1 - 98.96 = 0.0104 \times 2$

$P = 0.0208$

reject the null hypothesis


0.975
0.025
0.025
1.96

Question A factory manufactors bulbs with a avg warranty of 5 year with the std of 0.50. A worker believes that the bulbs will manufactor in less than 5 year. He tested a sample of 40 bubls and find the avg. time to be 4.8 year.

a) state the Null and alternate hypothesis

b) At 2% Significance level, is enough evidence to support the idea that the warranty should be revised.

$H_0 \Rightarrow \mu = 5$

$H_0 \Rightarrow \mu < 5$ (one tail test)

CL = 98%

α = 2%

$P \leq \alpha$

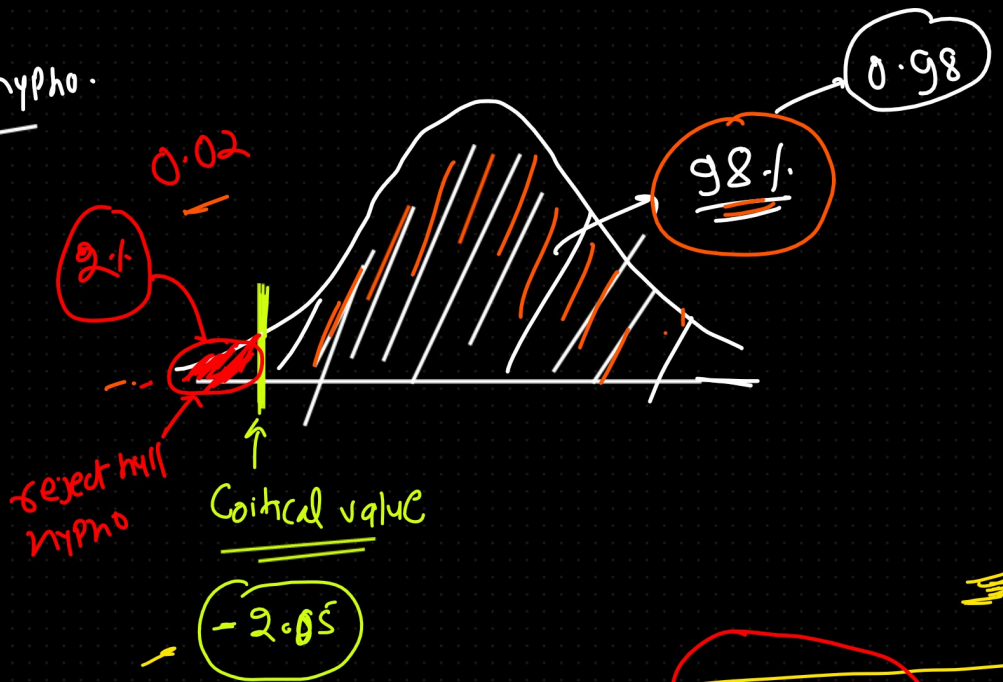
↑ reject the Null hypho.

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

$$= \frac{4.8 - 5}{0.50/\sqrt{40}}$$

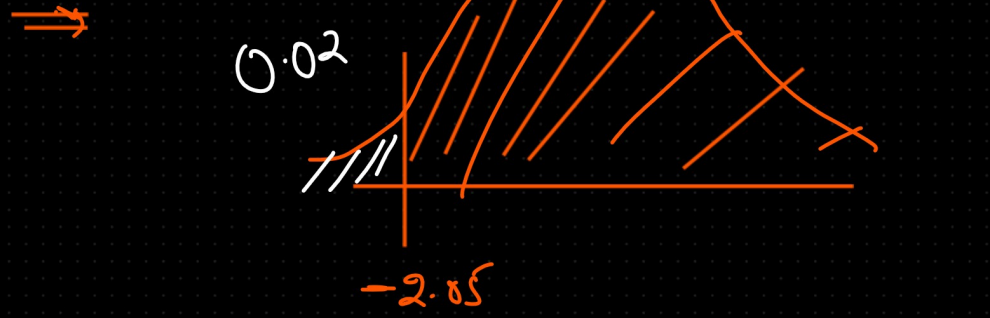
$$= \frac{-0.2}{0.50/6.32}$$

$$= \frac{-0.2}{0.07}$$



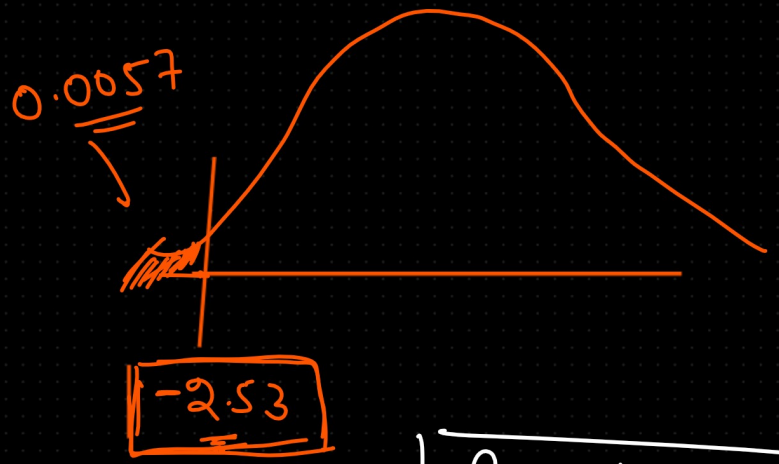
$Z\,score = -2.528$

$-2.528 < -2.05$



$P \leq 0.02$

reject



$P = 0.0057$

$\alpha = 0.02$

$P \leq \alpha$

$P < 0.02$

Reject the null hypothesis

Yes

t-test ⇒ [Population std dev ✗ Sample size ≤ 30 ✓]

Ques In the Population the avg IQ is 100. A team of researchers want to test a new medication to see if it has a either +ve or -ve 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 std of (20). Did the medication affect intelligence?

→ Null hypothesis $H_0 \Rightarrow \mu = 100$

alternate hypo $H_a \Rightarrow \mu < 100$, $\mu > 100$ (two tailed test)

CI = 95%
α = 5%

$$\begin{cases} n = 30 \\ \bar{X} = 140 \\ s = 20 \end{cases} \Leftarrow \text{Sample}$$

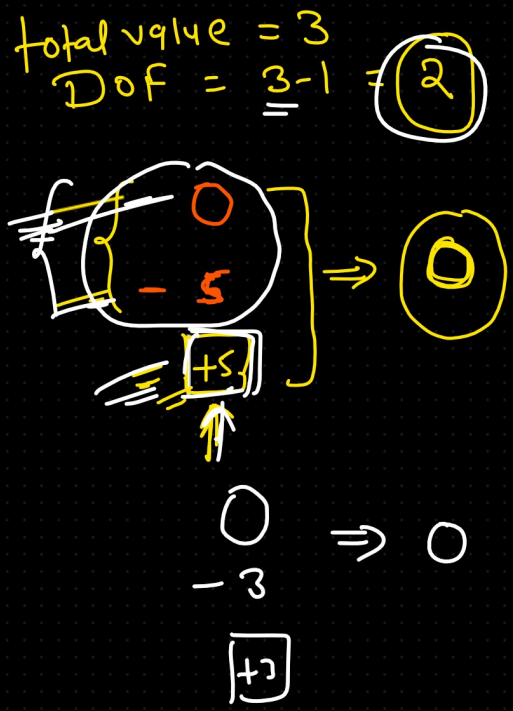
(1) Null hypothesis $\mu = 100$

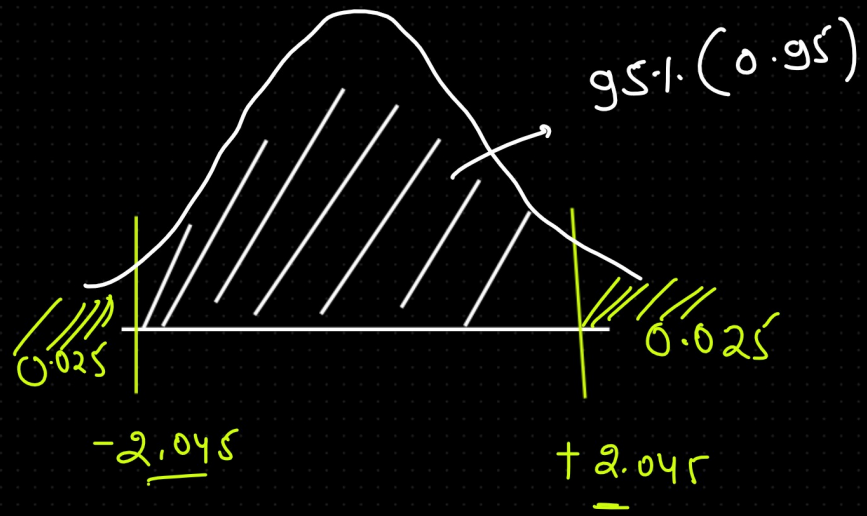
Alternate hypothesis $\mu \neq 100$ (2 tailed test)

(2) DOF $= n = 30$

$n - 1 = 29$

(3) Decision boundry → t-table

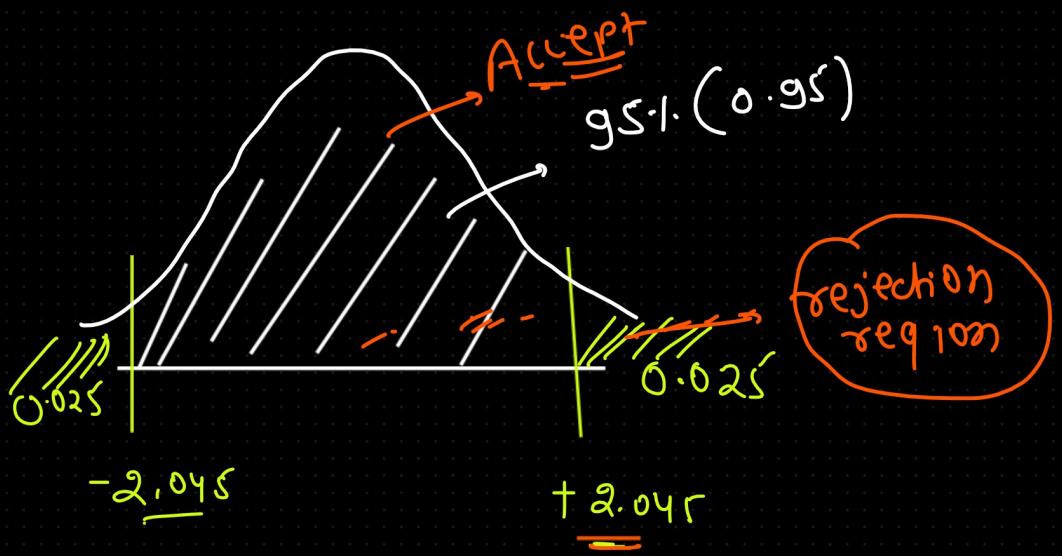


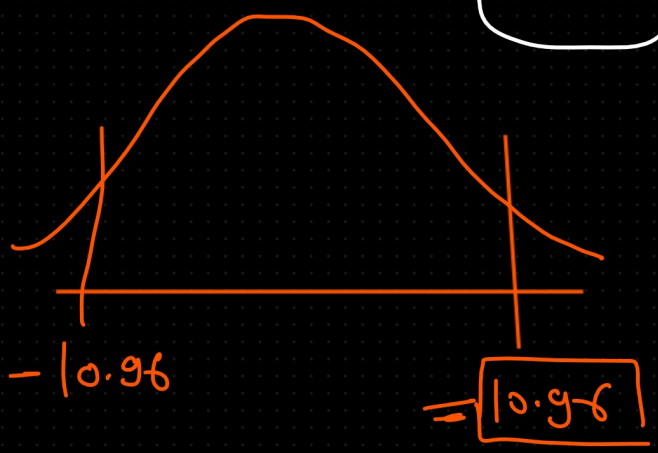


(4) Calculate t-stats

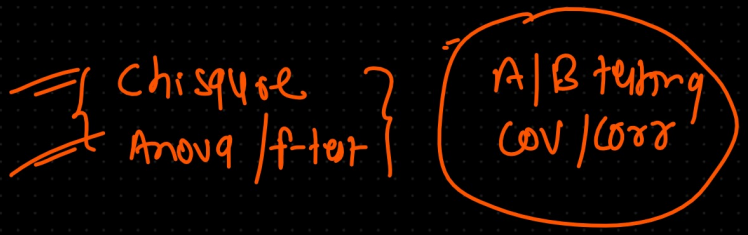
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{140 - 100}{20/\sqrt{30}} = \frac{40}{20} \times 5.47$$

$$= 10.96$$





= P-value =


Chisquare
Anova /f-test
A/B testing
Cov/Corr