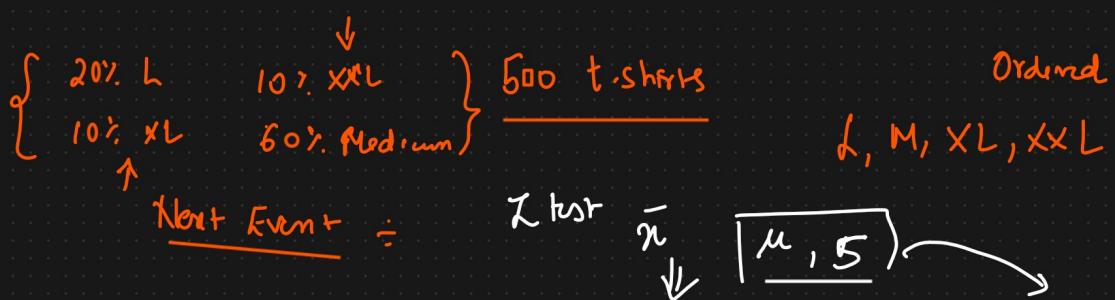


(2) Influential Statistics {Data Analyst, Data Scientist}

① 100K \Rightarrow T-shirt \Rightarrow No \Rightarrow Sample data \Rightarrow X_L, L, Small

② iNeuron \rightarrow Meetup \rightarrow Hitesh \Rightarrow 300-400 people \rightarrow T-shirts



③ ATM ④ Measure the size of entire sharks CI []

⑤ Amazon delivery {Percentile, Quantiles} \Rightarrow

(*) Hypothesis Testing

① A factory has a machine that fills 80ml of baby medicine in a batch. An employee believes the average amount of baby medicine is not 80ml. Using 40 Samples, he measures the average amount dispersed by the machine to be 78ml with a standard deviation of 2.5

(a) State Null and Alternate Hypothesis

(b) At a 95% CI, is there enough evidence to support machine is not working properly.

Ans) Step 1

$$n=40 \quad \bar{x}=78 \quad s=2.5$$

$H_0: \mu = 80$ {Null Hypothesis}

$H_1: \mu \neq 80$ {Alternate Hypothesis} } Why Z test?

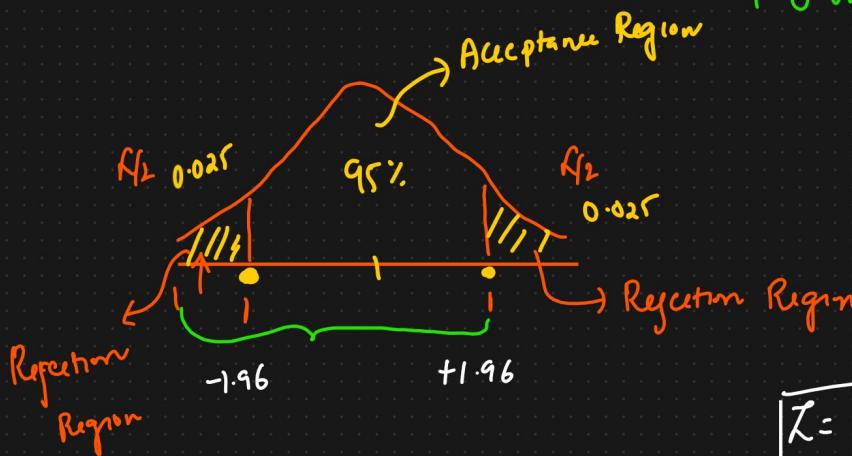
Step 2:

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

$n > 30$ or $n \leq 30$

① $n > 30$ ② population std or sample std

Step 3: Decision Boundary



$$1 - 0.025 = 0.9750$$

① Sample std =
② $n < 30$ =

$n=1$

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

$$Z = \frac{\bar{x} - \mu}{S / \sqrt{n}} \Rightarrow \text{Standard Error}$$

Sample Standard deviation

$$\text{deviation} = \frac{78 - 80}{2.5 / \sqrt{40}} = \frac{-2 \times \sqrt{40}}{2.5} = \frac{-2}{2.5} \times 6.32 = \underline{\underline{-5.05}}$$

⑤ State the Results

Decision Rule: If $Z = -5.05$ is less than -1.96 or greater than 1.96 , then reject the null hypothesis with 95% C.I

Reject H_0 Null hypothesis {There is some fault in the machine}

Q) In the population the average IQ is 100 with a standard deviation of 15. A team of scientists wants to test a new medication to see if it has a +ve or -ve effect, or no effect at all.

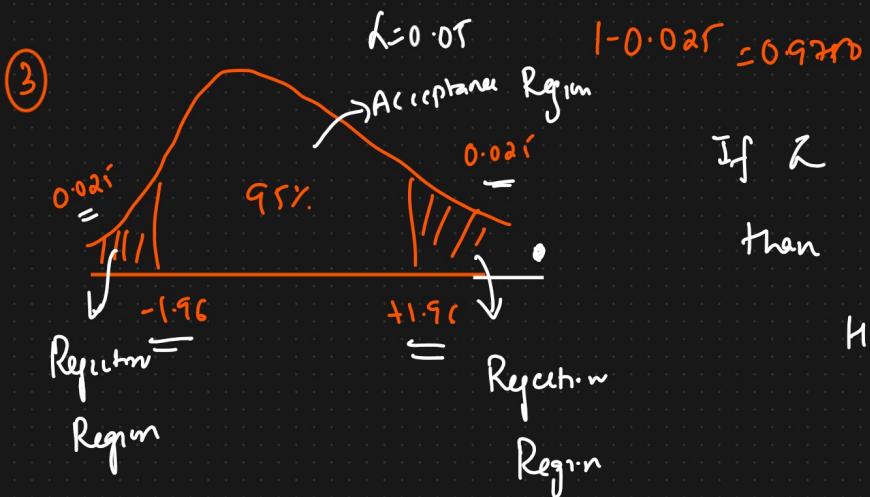
A sample of 30 participants who have taken the medication has a mean of 140. Did the medication affect Intelligence? $\left\{ \begin{array}{c} 95\% \\ \hline \downarrow \\ C.I. \end{array} \right\}$

Ans) $\sigma = 15 \quad n = 30 \quad \bar{x} = 140$

① $H_0 : \mu = 100$

$H_1 : \mu \neq 100$

② $\alpha = 0.05 \quad C.I = 95\%$



④ $Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{140 - 100}{15 / \sqrt{30}} = 14.60 \quad \text{---}$

$14.60 > 1.96$ Reject the Null Hypothesis

(*) A Complain was registered, the boys in the Municipal Primary School are underfed. Average weight of boys of age 10 is 32 kgs. With $s.d = 9 \text{ kgs}$. A sample of 25 boys was selected from the municipal school and the average weight was found to be 29.5 kgs? With $C.I = 95\%$. Check whether it's True or False?

$$\text{Ans}) \quad \mu = 32 \text{ kgs} \quad \sigma = 9 \text{ kg} \quad n = 25 \quad \bar{x} = 29.5 \quad \alpha = 0.05 \\ \text{1) } H_0 : \mu = 32 \quad \text{2) } \alpha = 0.05 \quad \cdot \quad 1 - 0.95 = 0.05$$

$$H_1 : \mu < 32$$

$$\text{③} \quad \begin{array}{c} \text{Rejection} = 1.96 \\ \downarrow \\ 0.025 \end{array}$$

$$85\% \quad 15\%$$

Acceptance Area

$$\text{python} \quad \Rightarrow \quad \{Z \text{ test}, p\text{-value}\}$$

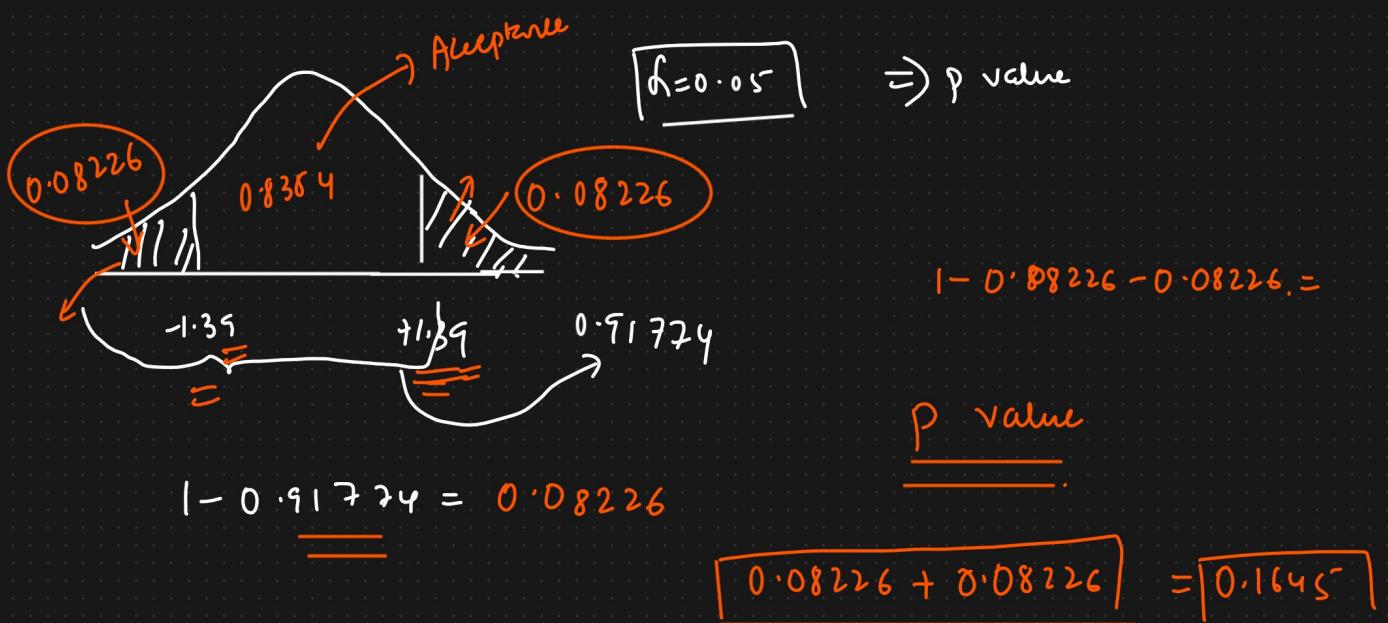
$$\text{④} \quad Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \\ = \frac{29.5 - 32}{9 / \sqrt{25}} = -1.39$$

Conclusion : $-1.39 > -1.96$ therefore we accept the Null Hypothesis

So, the boys are not underfed.

Significance value \Rightarrow

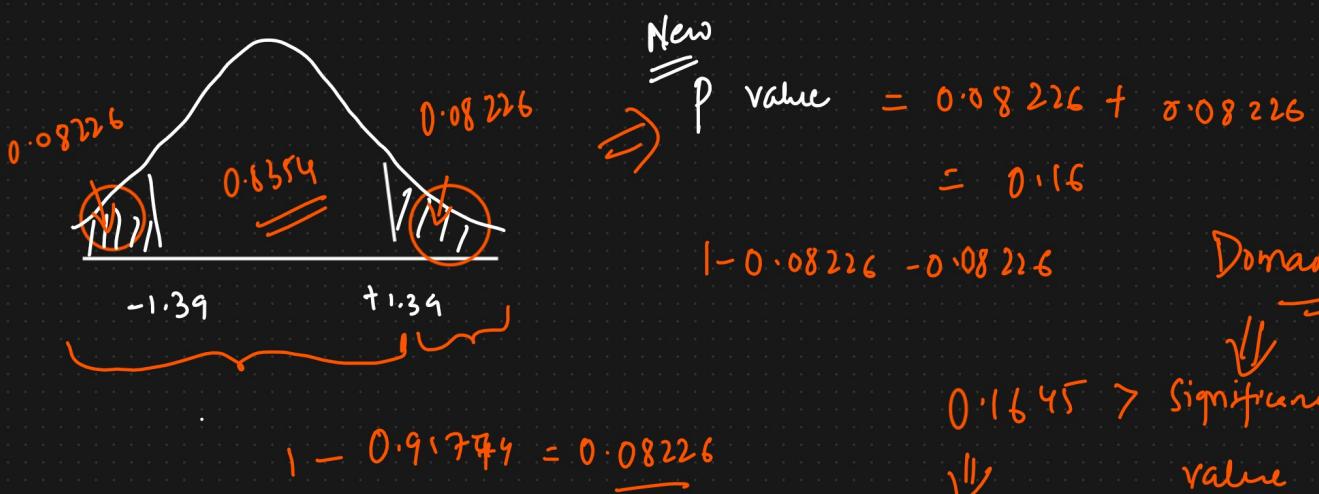
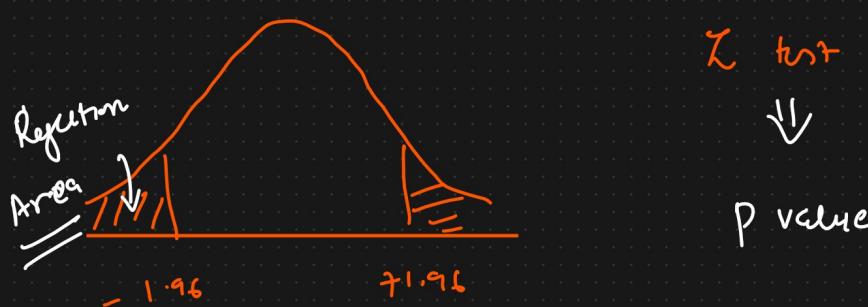
P value



Significance value

$0.1645 > 0.05$

$P \geq 0.05 \Rightarrow \text{Accept the Null Hypothesis}$



Domain \Downarrow

$0.1645 > \text{Significance value}$

\Downarrow

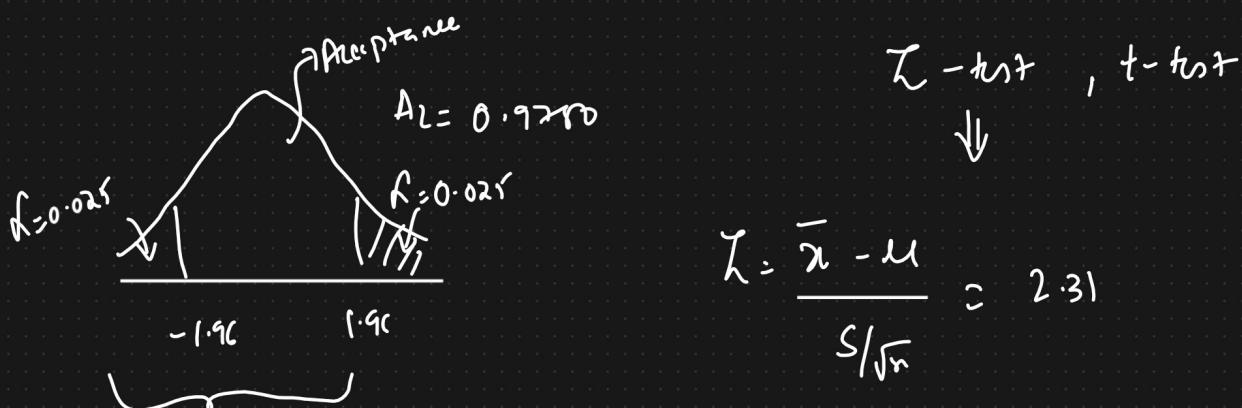
$\text{Accept the Null Hypothesis.}$

④ The average weight of all residents in town XYZ is 168 lbs. A nutritionist believes the true mean to be different. She measured the weight of 36 individuals and found the mean to be 169.5 lbs with a standard deviation of 3.9.

(a) At 95% C.I is there enough evidence to discard the Null Hypothesis??

$$\text{Ans}) \quad H_0 : \mu = 168 \quad n = 36 \quad \bar{x} = 169.5 \quad s = 3.9$$

$$H_1 : \mu \neq 168 \quad \bar{x} = \frac{\sum x}{n} \quad \alpha = 0.95 \quad \beta = 1 - C.I = 0.05$$



$2.31 > 1.96$ Reject the Null Hypothesis

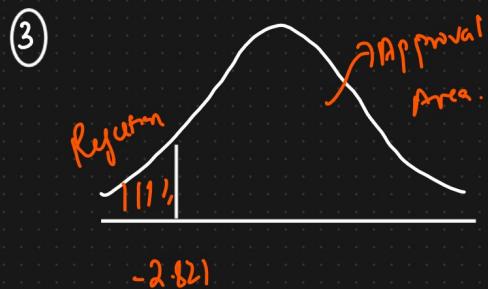
⑤ A company manufactures bike batteries with an average life span of 2 or more years. An engineer believes this value to be low. Using 10 samples, he measures the average life span to be 1.8 years with a standard deviation of 0.15.

a) State the Null and Alternative Hypothesis

b) At a 99% C.I, is there enough evidence to discard the H_0 ?

Ans) $H_0 : \mu \geq 2$ $n=10$ $\bar{x}=1.8$ $S=0.15$ $\{$ of sample
 $H_1 : \mu < 2$ ≤ 3.0 $t-tst??$ Std is
 $\{$ given }

② $\alpha = 0.01$ $\alpha = 1 - C.I = 1 - 0.99 = 0.01$



Degrees of freedom: $n-1$

$= 9$

④ Calculate t-test Statistic:

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}} = \frac{1.8 - 2}{0.15/\sqrt{10}} = \frac{-0.2}{0.15/\sqrt{10}} = \frac{-0.2}{0.15/\sqrt{10}} = -4.216$$

⑤ Conclusion

$-4.216 < -2.821$ Reject the Null Hypothesis. }
 \Downarrow

Z test with proportions

⑥ A tech company believes that the percentage of residents in town XYZ that owns a cell phone is 70%. A marketing manager believes that this value to be different. He conducts a survey of 200 individuals and found that 130 responded yes to

Owning a cell phone

(a) State the Null and Alternative Hypothesis?

(b) At a 95% C.I., is there enough evidence to reject the Null Hypothesis?

Ans) $H_0: p_0 = 0.70$.

$H_1: p_0 \neq 0.70$

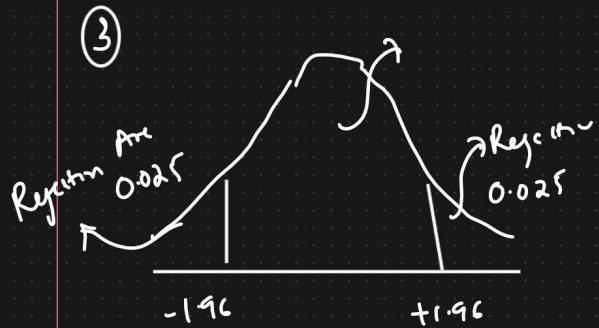
$$n = 200 \quad X = 130 \\ \hat{P} = \frac{X}{n} = \frac{130}{200} = \frac{13}{20} = 0.65$$

$$q_0 = 1 - p_0$$

② $\alpha = 0.05 \quad C.I. = 95\%$

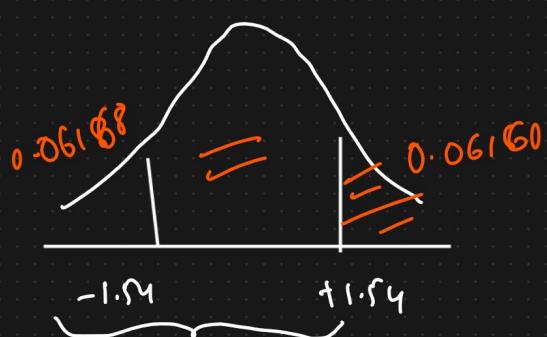
$$Z_{test} = \frac{\hat{P} - P_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

$$= \frac{0.65 - 0.70}{\sqrt{\frac{0.7 \times 0.3}{200}}} \approx -1.54$$



At 95% C.I. there is

$-1.54 > -1.96$, so we accept
the Null Hypothesis



p-value
 $2 \times 0.06168 > 0.05$

Accept Null Hypothesis

$$1 - 0.93822 = 0.06168$$

④ A car company believes that the percentage of residents in City ABC that owns a vehicle is 60% or less. A sales manager disagrees with this. He conducts a hypothesis testing surveying 250 residents and found that 170 responded yes to owning a vehicle.

(a) State the Null & Alternate Hypothesis

(b) At 10% significance level, is there enough evidence to support the idea that vehicle ownership in City ABC is 60% or less?

$$P\text{ value} = .014$$