Inferential Statistics

— It is the technique where we use the data that we measure to form a Conclusion.

Sample data → Conclusions related to Population data.
measure

Hypothesis Testing

— Hypothesis testing is use for making a decision or a Conclusion about population data.

— Null Hypothesis (H_0) → Coin is fair

— Alternate Hypothesis (H_1, H_A) ⇒ Coin is not fair

ex = Whether the Coin is fair or not.

— Null Hypothesis say nothing new is Happening.

ex Prob statement :→ population mean is 360, but domain expert says it may be 350.

$\mu = 360$ ↘ 350

$H_0 = \mu = 360$ ⇒

$H_1 = \mu \neq 360$ ⇒

Hypothesis testing aim = we try to gather evidence to reject the null hypothesis

Steps involved in Hypothesis testing

Step ① formulate a null Hypothesis (H_0) & Alternate Hypothesis (H_1)

Step ② Select Significance level (This a probability of rejecting null Hypothesis)

Step ③ Check assumptions (ex. distribution of (sd))

Step ④ Decide which test is appropriate (z-test, t-test, Chi-square test, Annova)

Step ⑤ → state the relevent test statistics →

Step ⑥ → Conduct a test →

Step ⑦ → Reject or not reject null Hypothesis

Step ⑧ - Interprete the result.

ex 1

Z test

- Company → Productivity test → training → Productivity test
  Decision → after taking productivity increase / decrease.

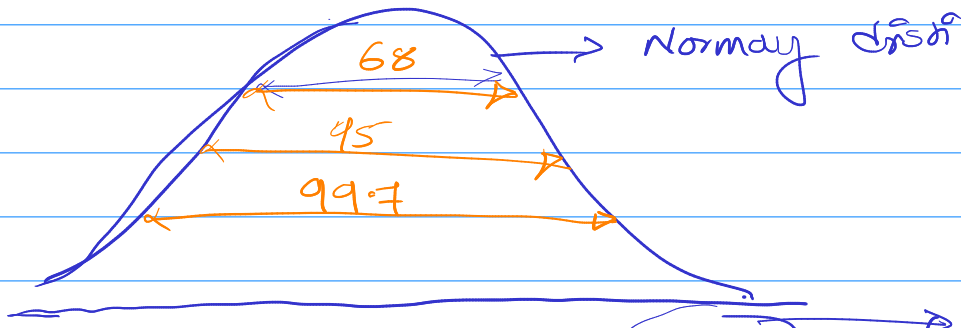
⇒ The average (mean) productivity was 50 unit per day with known population standard deviation of 5 unit

- After training productivity of random sample of 30 employee. The sample has an average productivity of 53 unit per day

→ after training productivity increase / decrease

$\mu = 50$ $\sigma = 5$ $n = 30$ $\bar{x} = 53$

$H_0 = \mu = 50$ $H_1 = \mu \neq 50$ $\mu > 50$ $\bar{x} = 53$

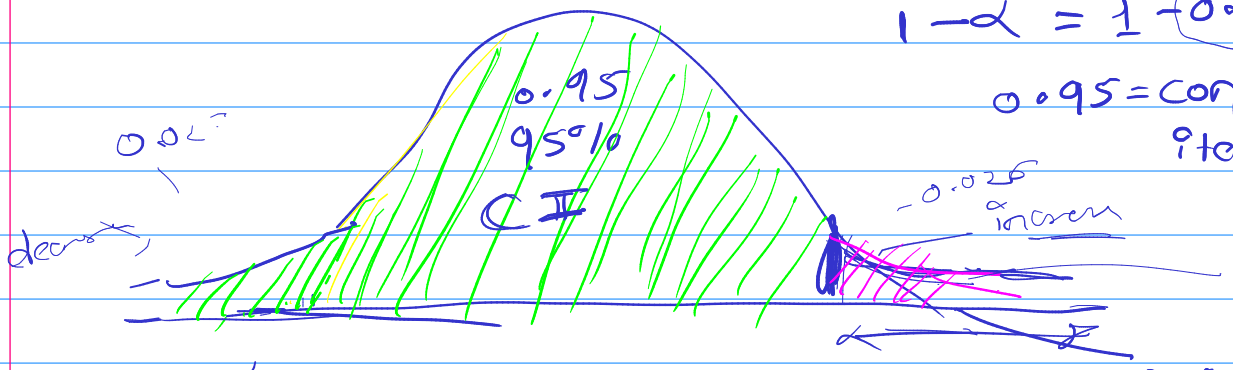


→ Normaly distributed

α → significance level ($\alpha = 0.05$)

$1 - \alpha = 1 - 0.05$

0.95 = confidence interval



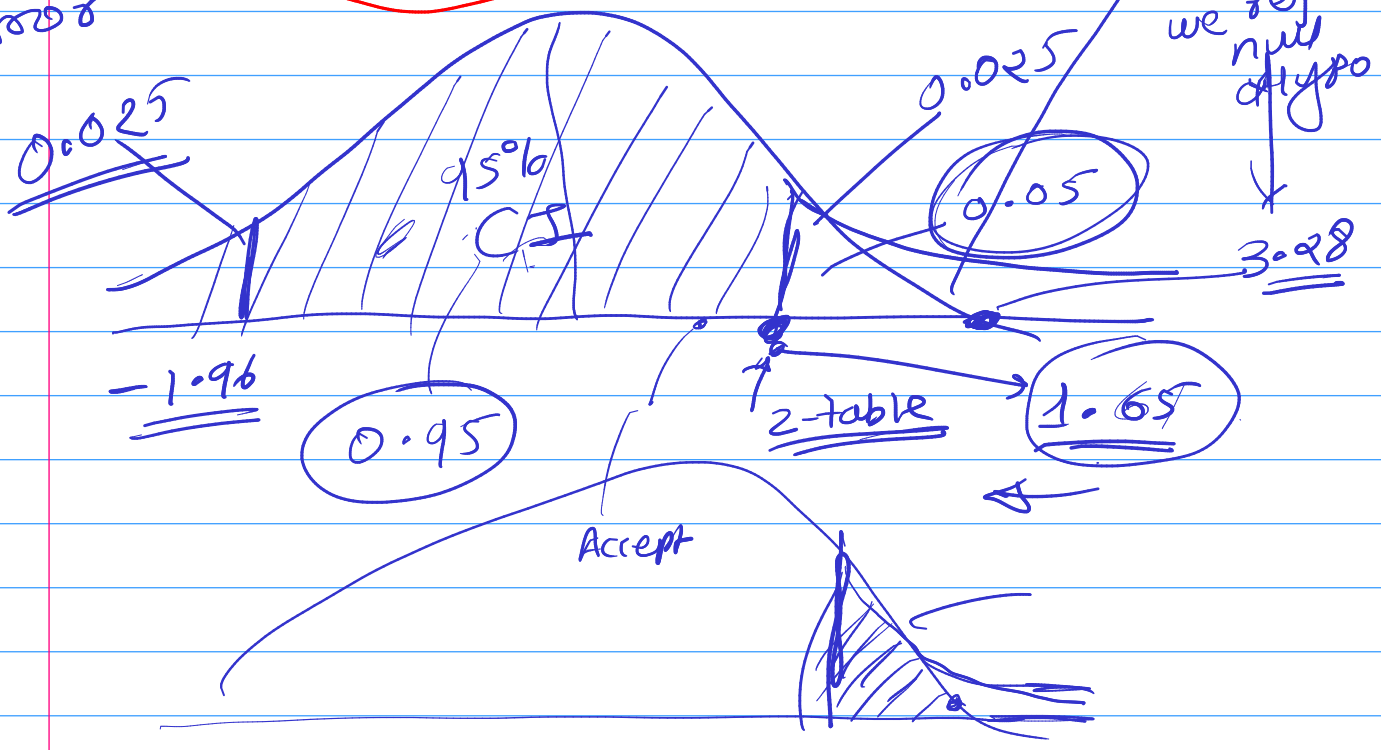
1 tail / 2 tail test = assumption —

1 tail test

Z-test

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{53-50}{5/\sqrt{30}} = 3.28$$

Standard error → $\sigma/\sqrt{n}$



① CI value = Ho Accept

② CI outside / SI - Ho Reject.

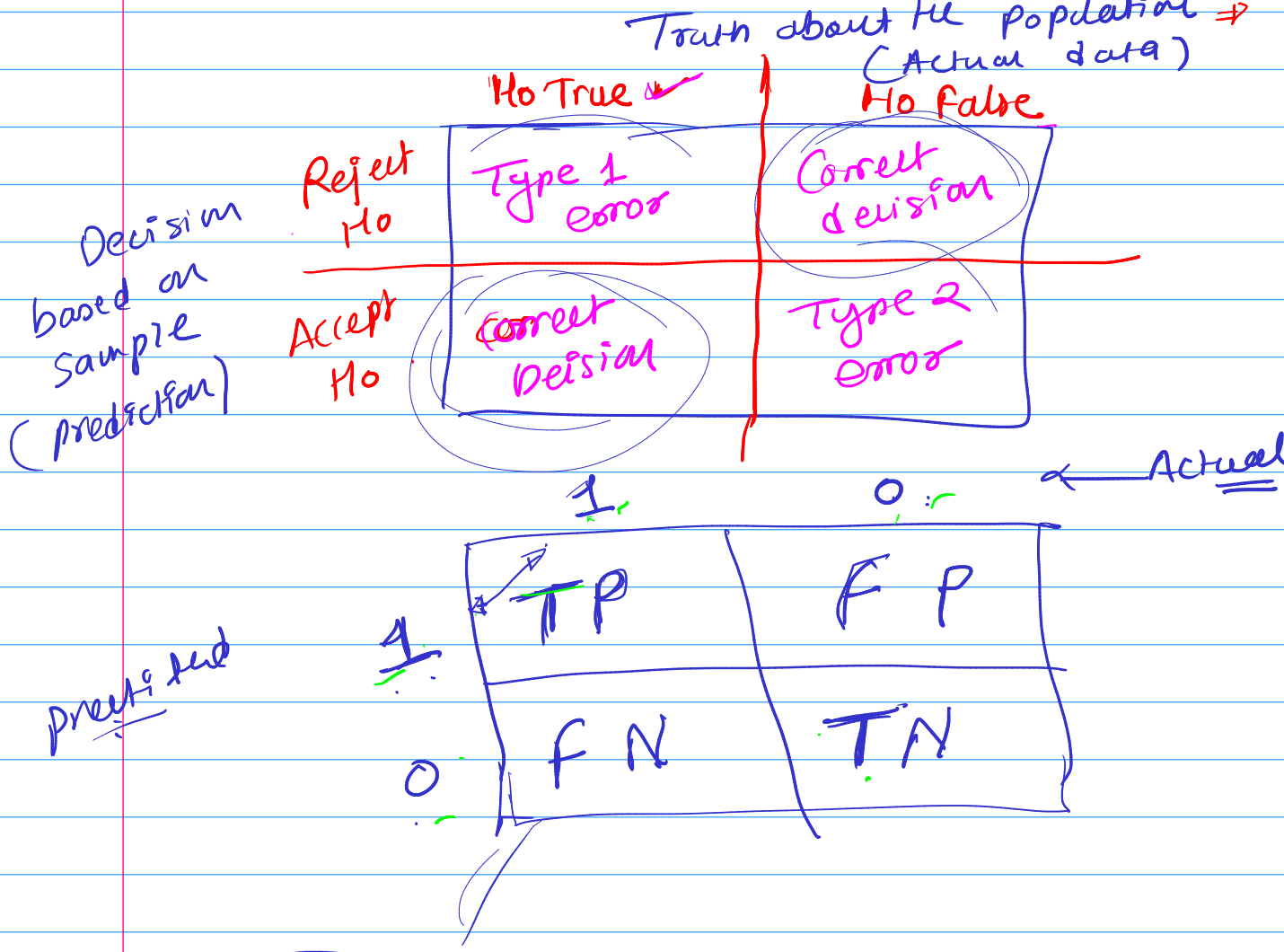
③ P value ⇒ Probability value where we can reject the null Hypothesis

T-test → Sample size less than 30

Chi-square test = Categorical data (Nominal / ordinal data)

Annova = multi categorical & numerical.

Confusion Matrix (Type 1 & Type 2 error)



estimate →