Hypothesis Testing →

→ What is this?

→ Z Test, T Test, Chi-Square, Anova Test

→ Feature Engineering

⇒ Youtube data Scientist →

→ Every 15 min → 1 Ads

→ ⇓

2 Ads

Case 1 → Move out of youtube

Case 2 → You will buy Premium

⇒ Assumption → More Premium users will be there

↓

Collecting Data

↓

We can Confirm/deny our assumption

⇒ Flipkart DS →

⇒ Product Page → Add to Cart → Checkout → place order

⇒ Product Page → Buy Now → Checkout

⇒ ① Nothing Changes ←

② Checkout Conversion will increase ←

③ Everything will go down ←

⇒ When we go for a toss → we make certain assumption

⇒ H vs T → 50% - 50% Coin is fair

→ Is there a way we can proof that coin is fair?

1) Toss Coin	10 times	→ 6 Head	4 Tails	11:3
2) Toss —"—	100 times	→ 60 Head	40 Tails	12:6
3) —"—	1000 times	→ 600 Head	400 Tails	8:7
4) —"—	100000 times	→ 60000 H	40000 T	0:15

⇒ Because we got large data to support that
Coin is baised.

⇒ How much data do we need?

⇒ Accused in Presented in Court?

⇒ "Innocent Untill Proven Guilty" →
Guilty untill Proven Innocent

⇒ Police / lawyer → provide evidence / data
→ Fingerprints, blooded knife, CCTV

After enough data is presented → My assumption (Innocent) was wrong
↓
The accused is a criminal

⇒ Amazon / Netflix DS → (A)
→ A legacy model has been working for the website for last 1 year

→ Your team Created Model → (B) which you claim to be better than
(A).

⇒ Assumption → ① A is same as B (Inherent assumption)

② B is better than A (Claim)

↓

I will provide data to support my claim.

→ user feedback
→ performance
→ Edge Cases

Third Umpire →

⇒ usually checks the on field umpire decision

→ Assumption → On field umpire is correct

Claim → Umpire is wrong →
- Slow motion video
- Ultra edge
- Ball tracking.

⇒ Covid Virus →

Assumption → Does not have disease.

Claim → I have a disease
- Symptom
- Blood cell counts
-

Hypothesis Testing Terminologies →

① Null Hypothesis (H_0) → Inherent assumption

1) Accused is Innocent
2) No virus
3) Previous Model is similar

② Alternate Hypothesis (H_a) → Claim for which data is provided

1) Criminal
2) have disease
3) B is better than A

Q If null hypothesis is rejected (enough data to reject null hypothesis is provided). does it mean H_a is true →

⇒ Criminal Case →

H_0 → Innocent

H_a → Criminal

A) Accused have a knife in pocket
B) knife has blood
C) Blood matches the victim
D) Knife also have fingerprint
E) CCTV footage shows the crime

⇒ Probability of seeing such extreme data if the person is innocent is very low.

$P(\text{data} \mid H_0)$ is very low

$\boxed{\text{P-value}}$

⇒ when to reject your null hypothesis →

we reject H_0 when P-value is low

A juice brand claims that its new manufacturing process has reduced the sugar content in its juice boxes to 8 gms.
Now, the Food Safety and Standards Authority of India (FSSAI) wants to test the claim of the juice brand. Set up the null (H0) and alternate (H1) hypotheses for this claim, and choose the correct option:

A → Reduced to 8 gm

B → Not Reduce to 8 gm

For FSSAI

H_0 → Reduce to 8 gm

H_a → Not Reduce to 8 gm

If the question would have been →

Brand is asked to proof their Claim that they have reduced the Sugar to less than 8gms

$H_0 \rightarrow$ The sugar Content is still $\geq$ 8gm (Industry level)

$H_a \rightarrow$ ——— " ——— has decreased.

A coffee shop makes amazing coffee.
Once, a coffee maker salesperson challenges the shopkeeper that his machine makes better coffee.
Set up the null (H0) and alternate (H1) hypotheses for this claim, and choose the correct option:

A $\rightarrow$ Coffee shop makes good coffee

B $\rightarrow$ Machine makes better Coffee

$\hookrightarrow$ B will be providing data to support their Claim.

$H_0 \rightarrow$ Coffee shop makes better Coffee

$H_a \rightarrow$ Machine makes better Coffee

$\Rightarrow$ ~~Slightly diff^n example~~ $\rightarrow$

(1) $H_0 \rightarrow$ Innocent $\quad H_a \rightarrow$ Criminal

A 100 witness $\longrightarrow$ 50 says $\rightarrow$ Criminal
50 says $\rightarrow$ Not Criminal

B $\longrightarrow$ 70 says $\rightarrow$ Criminal
30 says $\rightarrow$ Not Criminal

C $\longrightarrow$ 95 $\longrightarrow$ 95% Criminal
5

Pvalue is 5%

"Innocent untill proven guilty beyond a reasonable doubt"

⇒ Confidence and Significance level →

Threshold value (α) → If p value is less than α we can reject your null hypothesis.

↓

Significance level

↓

If not mentioned → $\alpha = 0.05$

Confidence level → $1 - \alpha = 1 - 0.05 = 0.95$

Repeat

→ If p-value is low → Reject H_0

→ threshold value → (0.05) (Probability (data | H_0) < 5%)

↓

significance level (α) → (0.05)

→ Confidence level → $1 - \alpha =$ 0.95

95% sure that H_0 is not true → Reject H_0

~~H_0 is true~~

Quick Qu: When do we reject Null Hypothesis →

$\boxed{\text{P-value is less than significance level } (\alpha)}$

Doubt ⇒ p_value $P(data \mid H_0)$ is low → reject H_0

⇒ ~~p_value $P(data \mid H_a)$ is high → accept H_a~~

Local people claim that the river water quality is good.
An environmental agency wants to test their claim. Upon testing, they obtained a p-value of 0.001.
The significance level (α) is set at 0.05. What is the appropriate conclusion?

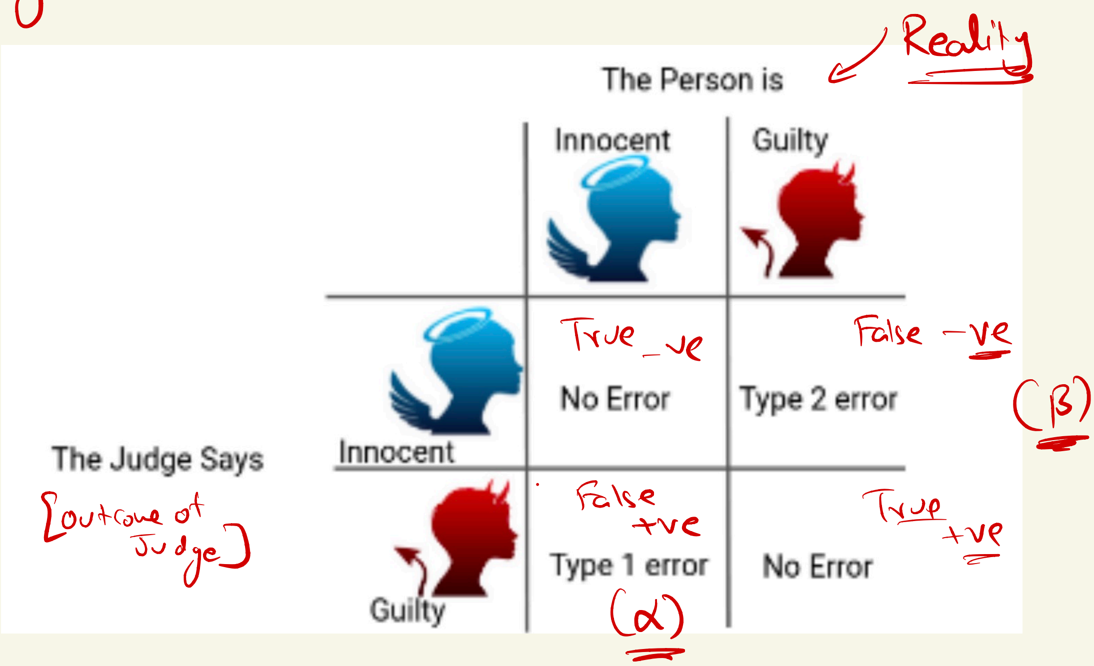
→ $p\_value = 0.001$
$\alpha = 0.05$

$p < \alpha$ → reject your H_0

⇒ Criminal

$p\_value \rightarrow 0.04$ $\alpha = 0.05$

→ My test → Yes, he is a criminal



Test → H_0 →

→ H_a →

Covid +ve →

+ve
-ve
} → Test Result (+ve → reject H_0 ; -ve → accept H_0)

False and True } → Comparing reality →

⇒ Covid Test

H_0 → No disease

H_a → Have disease

Reality

Test Result \ Reality	-ve	+ve
-ve	True -ve	False -ve
+ve	False +ve	True +ve

+ve test Result → I have disease

A security system is designed to detect intruders in a restricted area. The null hypothesis (H0) is that there is no intruder, and the alternate hypothesis (H1) is that there is an intruder. What is a Type II error in this context?

$H_0 \rightarrow$ No Intruder

$H_a \rightarrow$ Intruder

Type II $\rightarrow$ False -ve

Test Result $\rightarrow$ -ve

In reality $\rightarrow$ +ve

+ve $\rightarrow$ Intruder in Reality

-ve $\rightarrow$ H_0 is true as per test $\rightarrow$ No Intruder as per test

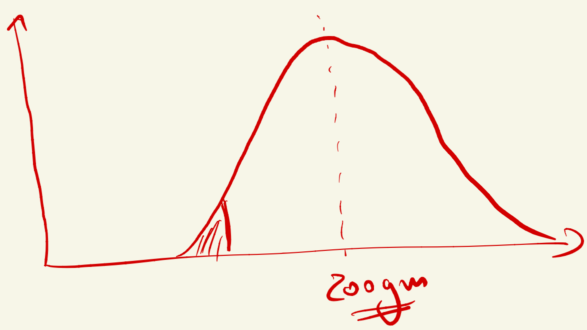
$\Rightarrow$ $\boxed{\text{My test failed to identify the Intruder}}$

$\Rightarrow$ Q A Burger Company is known to have ~~that~~ ~~every~~ burger is of weight 200gm!

$\rightarrow$ A customer claims that burger are lighter $\rightarrow$

$H_0 \rightarrow$ Burger = 200gm $\Rightarrow$ $\mu = 200gm$

$H_a \rightarrow$ Burger < 200gm $\Rightarrow$ $\mu < 200gm$



$\rightarrow$ Left tailed test

Q Legacy model → 90% accuracy

New ML model → 95% accuracy (claim)

H_0 → avg-perm = 90%

H_a → H_a > 90%

Right tailed test

(mean H_a > mean of H_0)

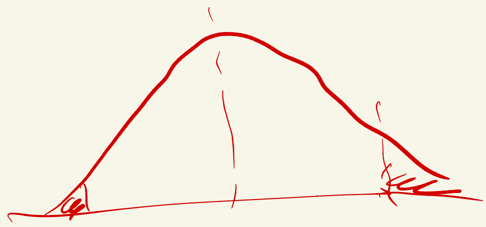
⇒

Q Avg height in India = 68 inch

→ Claim → It is not ?

$H_0 \rightarrow \mu = 68$ inch

$H_a \rightarrow \mu \neq 68$ inch



Two-Tailed test

A restaurant says that its new menu item has a mean waiting time of less than equal to 10 minutes.
You claim that the mean wait time is greater than 10 minutes.
What should be the null and alternate hypotheses to test their claim?

$H_0 \rightarrow \mu \leq 10$ min

$H_a \rightarrow \mu > 10$ min

Q. Coin Toss

$H_0 \rightarrow$ Fair Coin

$H_a \rightarrow$ unfair

$\rightarrow$ Toss the Coin 10 times $\rightarrow$ 7 Head / 3 Tails

$\alpha = 0.05$

Q. Is my Hypothesis true $\rightarrow$

p_value $\Rightarrow P(\text{data} \mid H_0) \rightarrow$

$\rightarrow P(\text{Get 7 Heads or more} \mid \text{Coin is fair}) \rightarrow$

$\rightarrow$ Prob of getting 7 or more Heads if coin is fair?

$\Rightarrow$ Prob of $\left.\begin{array}{l} H=7 \\ H=8 \\ H=9 \\ H=10 \end{array}\right\} \rightarrow 1 - \text{cdf}(6)$

H/T $\rightarrow$ 2 possibili $\rightarrow$ binomial.

Fix number of events

p_value = 1 − binom. cdf (k=6, n=10, p=0.5)

$\Rightarrow$ 0.17 $\rightarrow$ p_value $> \alpha$

$\Rightarrow$ H_0 is still true

$\Rightarrow$ 100 tosses $\rightarrow$ 70 Head

p-value = 1 − binom.cdf(n=100, p=0.5, K=69)

$\approx 4 \times 10^{-5}$

p-value << α $\rightarrow$ reject H_0
coin is unfair.

Conclusion

$\rightarrow$ First create Assumption (H_0 & H_a)

↓

Collect data (eg evidence 70 H out of 100 tosses)

↓

Calculate p-value ( )

↓

Conclude

$\rightarrow$ distribution

$\rightarrow$ which tail testistis

Doubts

Types of Error →

↳ If test Result ≠ Reality

Type 1 Error → False +ve

Type 2 Error → False −ve

⇒ +ve/−ve is defined by test result.

What is +ve test →

H_0 was rejected →

1) Accused is Criminal
2) Person has a virus

Test Results ↑	+ve	−ve ← Reality
+ve	True +ve	False +ve ← Type 1 Error
−ve	False −ve	True −ve

↳ Type 2 Error

Doubts

→ Assumption → H_0
H_a

→ Test | Experiment data →
(7 out of 10 tosses are Head)

experiment

⇒ How unlikely is this experiment data for a H_0 being true.

⇒ If H_0 is true (Coin is fair) what are the chances that I will see 7 or more Heads.

$$\Rightarrow \quad P(x=7,8,9,10 \mid H_0)$$

$$\Rightarrow \quad 1 - cdf(k=6 \quad n=10, \; p=0.5)$$

$$= 0.17$$

⇒ 70 out of 100 →

$$1 - cdf(k=69, \; n=100, \; p=0.5)$$ enough evidence to

$$\Rightarrow \quad 3.92 \times 10^{-5}$$ → reject H_0.