DAV-3

HYPOTHESIS TESTING

(Class starts @ 9:08 PM)

Lecture 1: Introduction to Hypothesis Testing

#Agenda

1. Module Overview
2. Problem Statement
3. Hypothesis Testing
4. Null & Alternate Hypothesis
5. p-value
6. Significance level
7. Types of Error
8. Failed Test
9. HT framework

Class starts at 9:10 PM

70% of this module deals with hypothesis testing

30% deals with feature engineering ML point of view

Module Overview

Real World Problem → How HT is used to solve those problems?

HT

For Example,

Website

Logo

I want to change logo go from blue to orange

② Power of EDA

↳ Feature Engg

$x_1 \rightarrow x_1'$ [derived column]

$x_2 \rightarrow x_2'$

$x_3 \rightarrow x_3'$

↳ Features

Transformation / Engineering

"Feature Engg"

x_1, x_2 and x_3

↳ Feature / Column

$x_1' \rightarrow x_1$

sales

↳ revenue

↳ profit

no. of siblings, dependents

↳ family size

③ Types of Hypothesis Test

- → Z-test
- → T-test
- → Chi-squared Test
- → ANOVA
- → Correlation Analysis

Problem Statement (Youtube)

1 ad ⟶ 15 mins

2 ads ⟶ 15 mins

this already happening

will be good or not

Sol

A/B (HT)

H0

Case I: No Impact

H0

H1

Case II: Impact

effect can be +ve

pos impact: users switching to premium

neg impact: Switch to diff platform

Step 1: Coming up with Ho & Ha

Q) What is Hypothesis Testing?

IPL ⟶ CSK Vs MI

{Toss}

Step1
Hypothesis → Coin is fair
P(H) = P(T)
⇓
assumption
P(H) = 0.5

Ha: Coin is not fair

Exp 1 Step 2
{ n = 100, H = 90 } seem biased

Exp 2
{ n = 1000, H = 900 } seem biased

Ho :- Null hypothesis
↑
Default Assumption
↳ Coin is biased

↳ Ho: Coin is fair
Ha: Coin is biased

Null hypothesis is always no effect assumption

✓ Step 1 : Assumption (Hypothesis)

✓ Step 2 : Data / Evidence

H_0 : Coin is fair

H_a : Coin is not fair

100 , H = 50 T = 50

Judge in Court

Suspect at hearing

↳ charged for murder

"Innocent untill proven guilty"

1. Assumption: Suspect is not guilty
   ↳ Default Assumption
2. Evidence:
   - fingerprint ✓
   - CCTV ✓
   - Murder weapon ✓

H_0: Suspect is innocent

H_a: Suspect is guilty

ML Deployment Example

(Rec Sys)

ML 1 : Legacy Model (10yrs back) ML 1

ML 2 : New version of ML 1

(Rec Sys)

HT

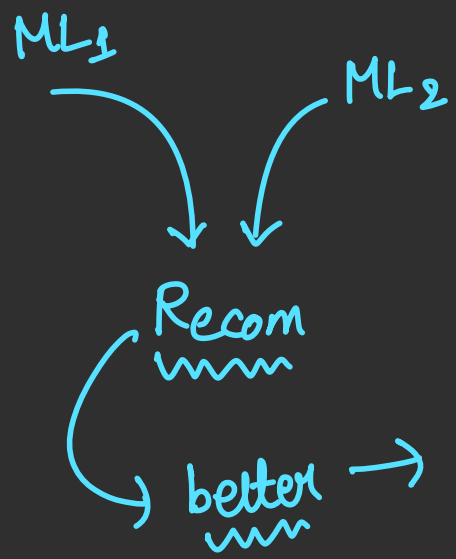
1. Assumption : $ML_1 == ML_2$

2. Evidence : metrics → change in accuracy, Increase in sales, Less memory consmp.

Some sample of users

ML1 → 68%

ML2 → 68% ✗


ML1
ML2
Recom
better

Terminologies

1. Null Hypothesis (H_0)
   - → Default assumption
   1. Coin is fair (no effect)
   2. Suspect is innocent

2. Alternate Hypothesis (H_a)
   - → Opp of Null Hypothesis
   1. Coin is unfair
   2. Suspect is guilty

1. Assumption
   - → H_0 and H_a
     - H_0: Coin is fair
     - H_a: Coin is biased

(Default Assumption) → H_0 → Null Hypothesis

↳ It's not gonna rain

H_0: Coin is fair

H_a: Coin is unfair

Case I

H_0: Coin is unfair

H_a: Coin is fair

Case II

Step 1 Frame H_0 & H_a

Step 2 Collect Data

Step 3 Calculate p value

(Why ?) → p value → H_0 is true / H_a is true

"pvalue calculation"

Q) Why H_0 is the default Assumption? [VVI]

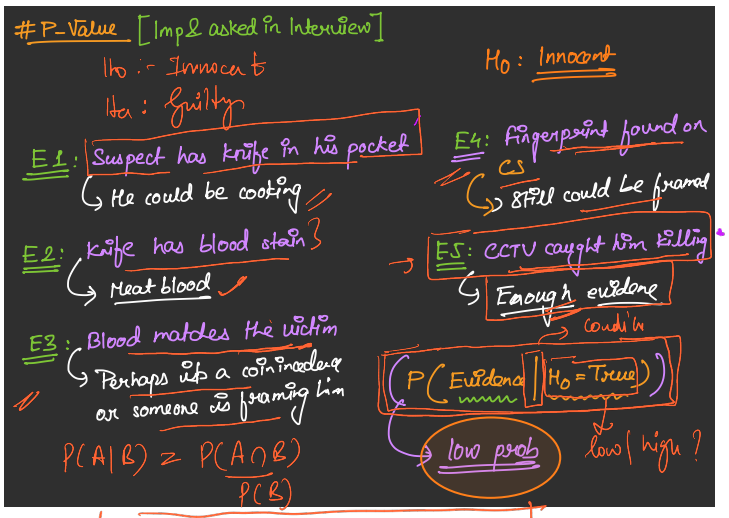
Coin Fair Example

H_0 : Coin is fair / Coin is unfair

↳ for which you can calculate the probability

✓ P(H) = 0.5 (given that coin is fair)

✓ P(H) = { Can't (given that coin is unfair)
         { Calculate



$P(\text{Evidence} \mid H_0 = \text{True})$ → low or high ?

$$P(\text{Evidence} \mid H_0 = \text{True}) = \frac{P(\text{Evidence} \cap H_0)}{P(H_0 \text{ is true})}$$

(P → low; $P(H_0 \text{ is true}) = 1$)

$P(E_1, E_2, E_3, E_4, E_5 \mid H_0 \text{ is True})$

P value is the probability of obtaining the observed results given that your null hypothesis is true

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(\text{Evidence} \mid H_0 = \text{True}) = \frac{P(\text{Evidence} \cap H_0 = \text{True})}{P(H_0 = \text{True})}$$

He is innocent but has been seen committing a murder

p-value ⟶ $P(\text{Data} \mid H_0 = \text{True})$

p-value

H_0: Suspect is innocent

(True)

$H_0 \rightarrow$ True / False

$CI = 1 - \alpha$

$= 1 - 5\%$

$= 95\%$

5%

pvalue

Medical domain

$\alpha = 1\%$

$CI = 99\%$

low prob → How to decide? [threshold]

(Assumption was wrong)

(Ha is True)

→ $\alpha = 5\%$

error — Chance of committing an error

95% I will be correct

$CI = 1 - \alpha$

- ✓ Step 1: Come up with Ho & Ha
- ✓ Step 2: Collect Data & evidence
- ✓ Step 3:- Calculate p value
- ✓ Step 4: ⓐ Take significance level

p-value < significance level

- ✓ Step 5: reject null hypothesis

fail to reject null hypothesis

I want to predict whether coin is fair or not

Flipping coin (100 times) $\alpha = 5\%$ 0.05

1) H_0:- Coin is fair

H_a:- Coin is unfair

2) Collect Data/Evidence

50 Heads ↳ 80 Heads

$\frac{50}{100} = 0.5$

3) p value → 0.002

0.7 > α

0.5 > α

4) p value < Significance level

=

Quiz

$$\begin{cases} H_0: & SC \text{ is } \geq 8g \\ H_a: & SC < 8g \end{cases}$$

Claim ⟶ New Manufacting proces helps produce SC less than 8g

Significanc & Confidence level

p-value ⇒ [0, 1]

Threshold → "Significance level" (α)

1000 W

997 g ✓

03 → I

Court

100 W

→ 50% Suspect is guilty

→ 50% Susp. is innocent

("Can't Say")

100 W

→ 75 → guilty

→ 25 → innocent

[Can't Say we still need more evidence]

1000 W

1) 997 → guilty

2) 03 → Innocent

$$\begin{cases} CL = 99.7 \\ \alpha = 0.3 \end{cases}$$

We want to keep the threshold in such a way that chances of error is less

p-value < threshold

↳ significance level

CL = 95% =

CL = 1 − α (sig level)

95% ⇒ 1 − 5%

↳ sig level

pvalue → 0.06

α → 0.05

{ if p-value < α : reject Ho
else accept Ho

0.06 < 0.05 ✗

pvalue $\longrightarrow [0,1]$

$\alpha \longrightarrow [0,1]$

(pvalue with alpha)

$CL = 1 - \alpha$

$\alpha = \underbrace{1 - CL}$

$1 - 1 = 0$

$\alpha = 1 - 0 = 1$

#Quiz

pvalue ⇒ 0.001

α = 0.05

pvalue < α :

↳ Reject Ho / Accept Ha

↳ else : Accept Ho / Reject Ha

01 → FP → retest → 00 → TN → FP

10 → happily → FN (Life gone) → FN 10 → it's not

Types of Error

Detecting Cancer 1

Able to detect Cancer 1

Judge in the court (Predicted)

Judge says

Actual person	I.	G.
I	(TP) No Error	Type 2 Error (FN)
G	Type I Error (FP)	No Error (TN)

↳ α (Significance level)

Confusion Matrix

FN → Important FP ?

Hai:- Person is Innocent → Positive 1

Hai: Person is Guilty → Negative 0

Actual - 11

TP → 11

10 → FN

01 → FP

Confusion matrix 00 → TN

10 → FN

FP → Supreme Court

10 → FN

01 → FP

00 → TN

Actual 1 0 Predicted → FN ✓ 10

FP Actual 0 Predicted 1 → 01 →

Person is Innocent → Positive

Person is Guilty → Negative

Actual person \ Judge says	I	G
I	True Positive	False Negative → Type II error
G	False Positive → Type I error (α)	True Negative

α = sig level

predicted (Judge says)

Actual	I	G
I		
G		

- I → Negative
- G → Positive

Quiz

What is Type 2 Error

✓ H_0: There is no Intruder

✓ H_a: There is a Intruder

NI → Negative Statement

I → Positive Statement