Module Overview

1. **Introduction to Hypothesis Testing**
   Understanding the basic concepts of hypothesis testing, including null hypothesis, alternate hypothesis, P-value, and significance level. etc.
2. **Types of Errors & Framework**
   Exploring the concepts of Type I and Type II errors, and understanding the framework for hypothesis testing.
3. **Numerical vs Categorical Tests**
   Differentiating between hypothesis tests for numerical vs categorical data, and their respective methodologies using different hypothesis tests like Z-tests, T-tests, ANOVA etc.
4. **Categorical vs Categorical Tests**
   Exploring hypothesis tests tailored for comparing categorical variables and analyzing their relationships using Chi- Square test.
5. **Numerical vs Numerical Tests**
   Understanding hypothesis tests designed for comparing numerical variables and assessing their correlations & its types.
6. **Advanced Hypothesis Tests**
   Delving into advanced hypothesis testing techniques such as the Kolmogorov-Smirnov test, A/B testing, and Two-way ANOVA.
7. **Exploratory Data Analysis (EDA) & Feature Engineering**
   Introduction to exploratory data analysis techniques and the process of feature engineering to enhance predictive modeling.
   Techniques like Handling Missing values, Outliers, Categorical & Numerical Data etc.
8. **Yulu Case study**
   You will be working on the real time dataset to extract valuable insights and provide actionable recommendations.
9. **Delhivery Case Study**
   You will be working on the real time dataset to conduct EDA & Feature Engineering to extract valuable insights and provide actionable recommendations.
10. **Module Test & Module Re-test**

Expectations from Learners

- ≥ 85 % PSP
- ≥ 70 % Module Test (90 mins)
- Business case study
- Mock interview

Content and Notes

- GitHub
- Dashboard

} Codes, Hand written Notes

- End of Module: Cheat Sheet
- Post-Lecture Content in Dashboard

Introduction to Hypothesis Testing

* Motivation: Imagine you are a DS at Amazon

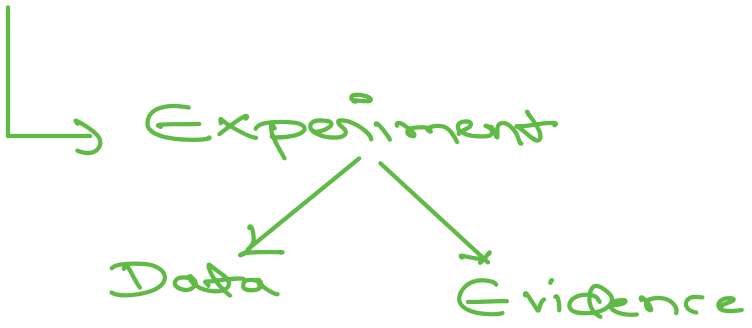
Let's say an executive suggests to remove Add to Cart button

Question: What will be impact of this?

Case 1: No impact on Sales

Case 2: Buying becomes simple process and increases Sales

Case 3: Sales Decrease



Hypothesis Testing

Consider scenario of Coin Toss in a Match

Default Assumption: Fair Coin (50%-50%) H T

Claim: The coin is Biased

What will be needed to say Coin is Biased ?

- Data
- Evidence

Let's consider following Exp's to collect Data

1)

2)

3)

Our Focus today will be on:

1) Assumptions

2) Data and Evidence to support/Negate the assumptions

Judge in Court Example

Consider a judge Hearing case of a Murder Suspect

Default Assumption: → (Reject this)

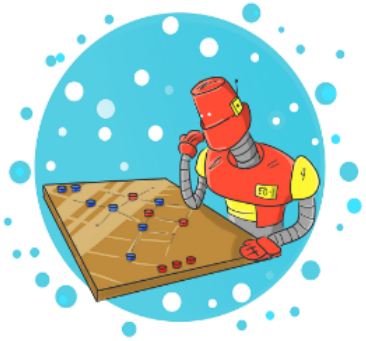
"Innocent Until proven Guilty"

What data must be collected to reject Default Assumption?

- Fingerprints on murder Scene
- Murder Weapon : Knife
- Forensic Analysis
- Eye-witness
- CCTV Footage

Etc.

ML Model Example



Default Assumption

- Legacy is Better

Claim

- New is Better

RecSys Model: Existing

Legacy Model

Your team's Model:

New Model

↳ Test Data: offline

Eval ⇒ 70% more Sales

Note: The Burden of proof is on Rejecting the Default Assumption (or Claim)

Hypothesis Testing Terminologies

2 possible judgements:

1) Accused is Guilty
2) Accused is Innocent

Default Assumption: Innocent

↓

1) Null Hypothesis: (H_0)
   - Status Quo
   - Baseline Assumption
   - Default Assumption

Assumed to be True Until Enough Evidences are present

② Alternate Hypothesis (H_a)

① Any other Hypothesis apart from Null

② We prove this By Rejecting the Null Hypothesis

Null : Innocent (Judge)

Suppose we have the following data/evidence:-

P (Evi | Null Hypothesis) P↓

- **Evidence 1:** Suspect has a knife in their pocket
  - Not enough evidence to say he is guilty
  - Innocent people can also have a knife
- **Evidence 2:** Knife has blood stains on it
  - Not enough evidence.
  - Chef, working with meat / maybe cut himself
- **Evidence 3:** Blood matches victim
  - Perhaps this is too much to be coincidence
  - But still could be planting of evidence
- **Evidence 4:** Suspect fingerprints found on victim
  - Enough evidence
- **Evidence 5:** CCTV catches suspect at scene of crime / Eyewitness
  - Enough evidence

High P↓

Slightly Less High P↓

→ Moderate P↓

→ Low P↓

→ Very Low P↓

Ho

P-Value

If we observe all these evidences, then we can reject H_0, because an innocent person is highly unlikely to have this

A Low p-value indicates the observed Data is highly unlikely to occur under Null Hypothesis

P-Value = P (Data | Ho is True)

P-value ∝ P(Ho) ✗

P-value ∝ P (Ho) ✗

Claimed : < 8 gm (H_a)

Null : ≥ 8 gm (H_0)

$H_0 = 8$?

Two-tailed test

Confidence and Significance Level

P-Value : Probability of Observing the Data atleast as Extreme or More given that Null Hypothesis is True

Goal ⇒ Reject the H_0

↓

P_{low} is Need

Threshold (α) Significance Level

$\alpha = 0.05$ or 0.01 (default)

↓

Confidence level $(1-\alpha)$ (95%)

CL($\alpha = 0.05$) = 0.95

$P\text{-val} < \alpha$

↓

Reject H_0

P-value = 0.001

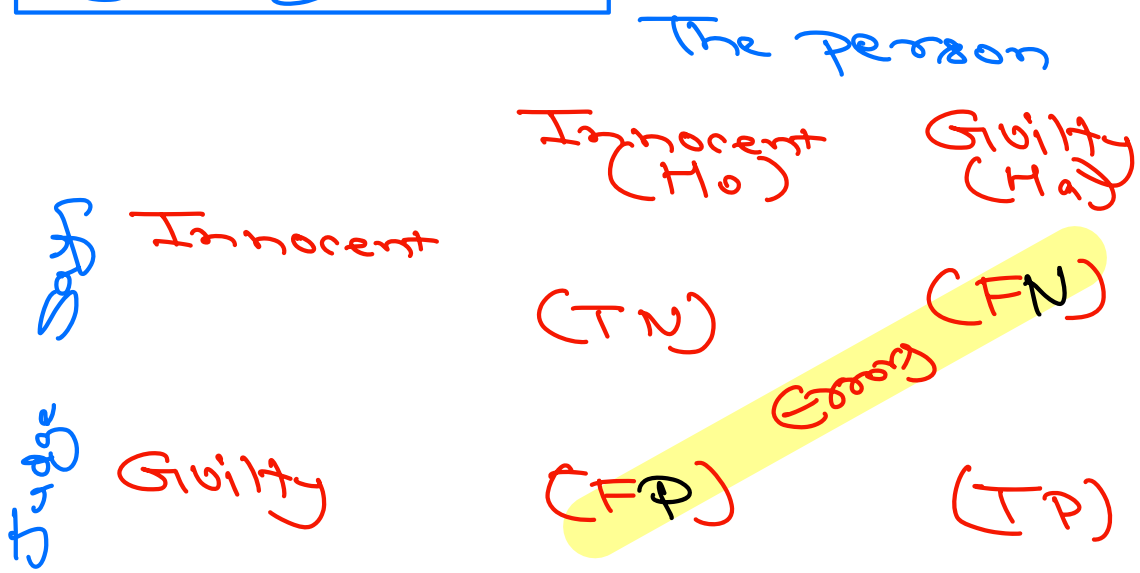
$\alpha = 0.05$

P_val < α : T

↓

Reject H_0

Types of Errors



H_0 → Innocent (Negative)

Case 1: We decide that the accused is innocent, and he is actually innocent (i.e. **True Negative**)

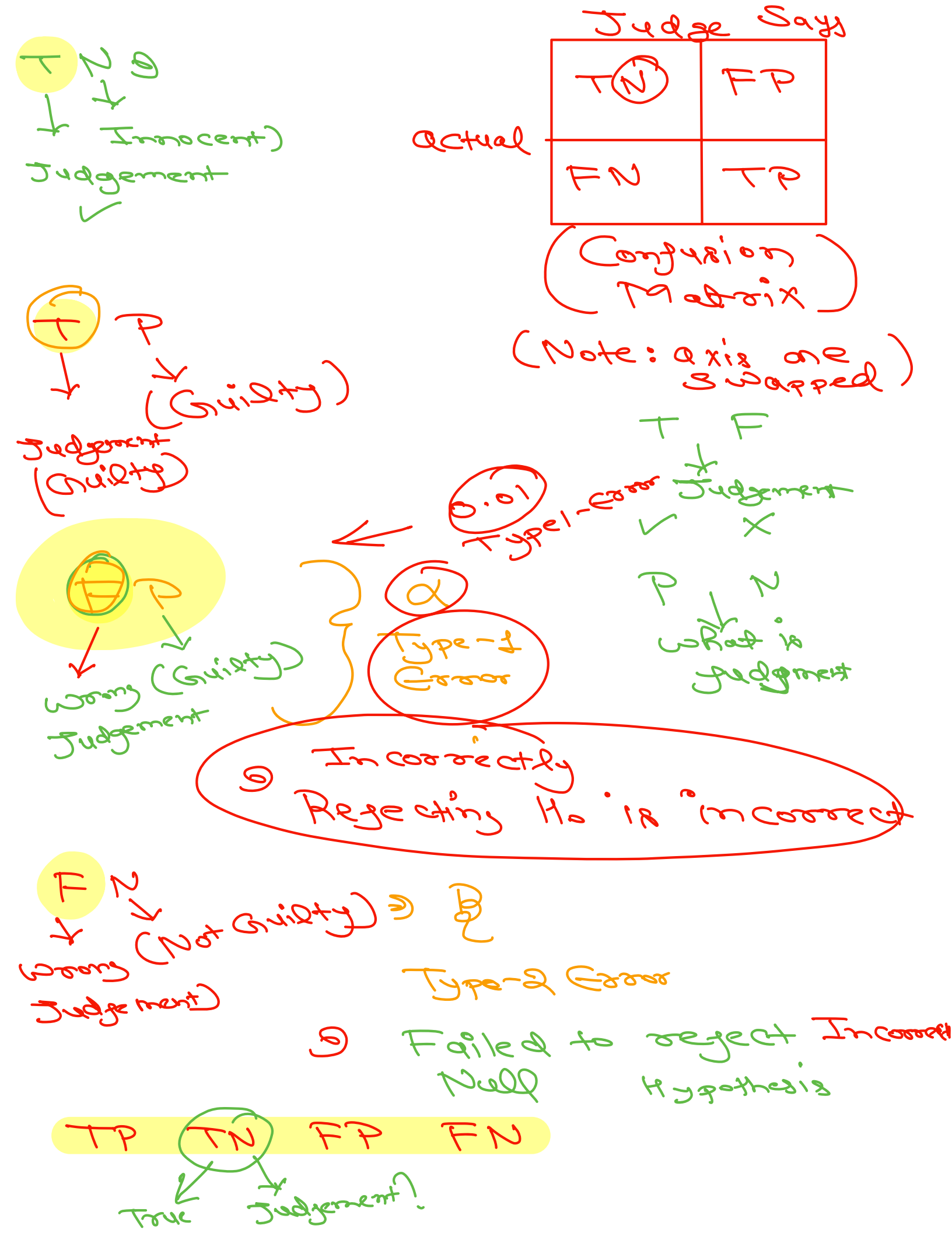
Case 2: We decide that the accused is guilty, but he is actually innocent (Type-1) FP

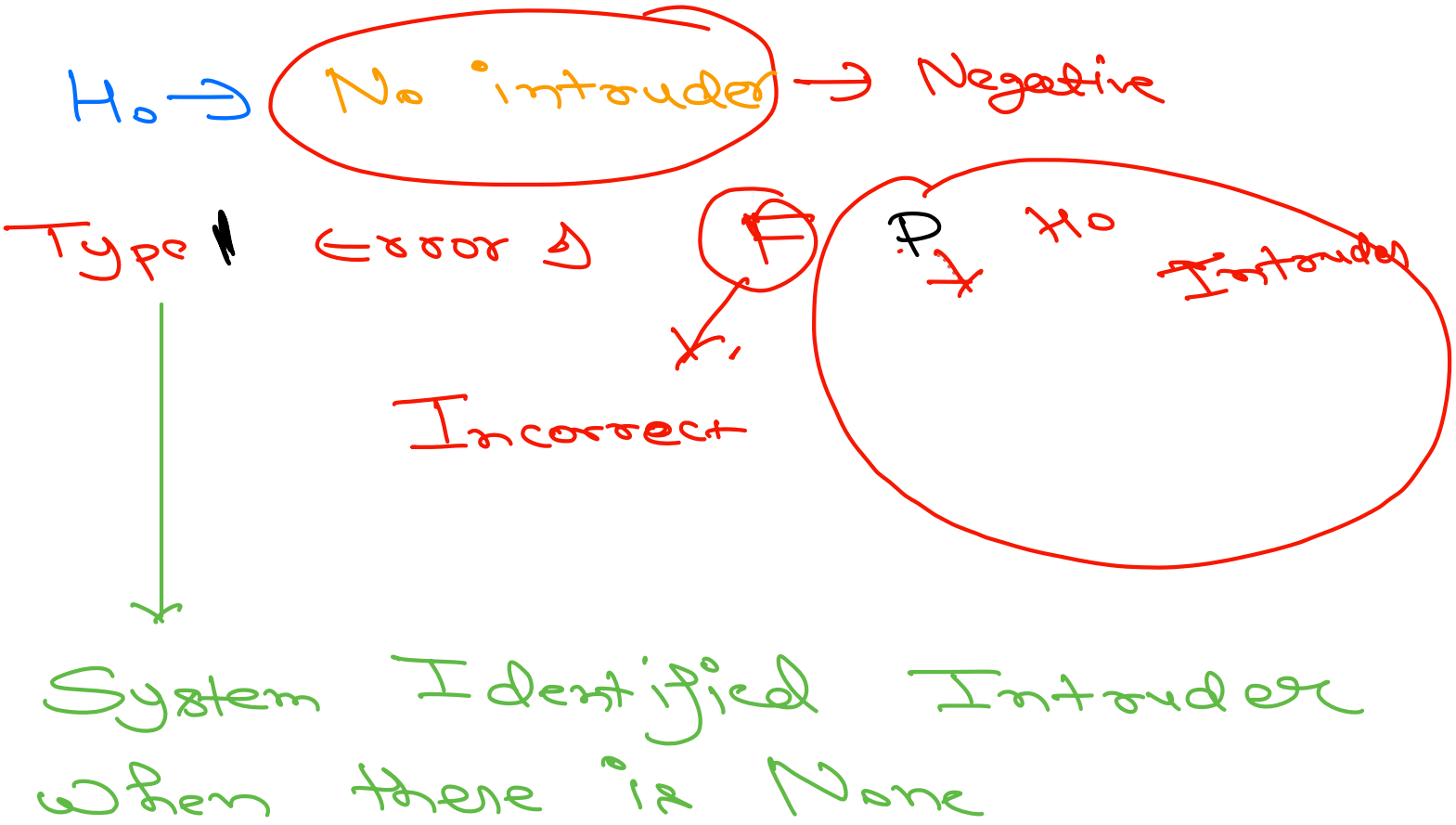
Case 3: We decide that the accused is innocent, but he is actually guilty (Type-2) FN

Case 4: We decide that the accused is guilty, and he is actually guilty (i.e. **True Positive**)

1. Type - 1
2. α
3. FP
4. ✗ Rejection of Correct H_0

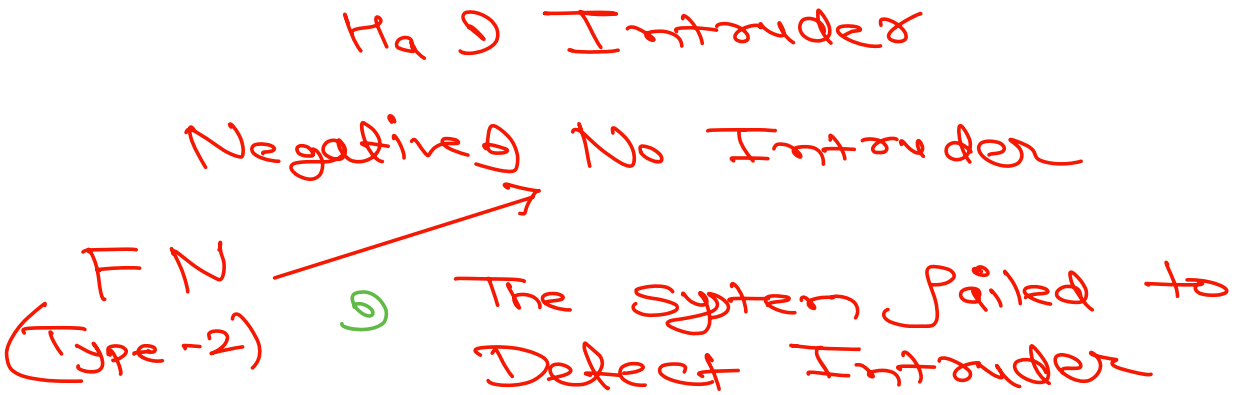
1. Type - 2
2. β
3. FN
4. Failure of Rejection of Incorrect H_0


T N
Innocent
Judgement
Judge Says
Actual
T N
FP
FN
TP
Confusion Matrix
T P
Guilty
Judgement
Guilty
(Note: axis are swapped)
T F
Judgement
P N
What is Judgement
0.01
Type-1 Error
Wrong Guilty Judgement
Type-1 Error
Incorrectly Rejecting Ho is incorrect
F N
Not Guilty
Wrong Judgement
Type-2 Error
Failed to reject Null Hypothesis
Incorrect
TP TN FP FN
True Judgement!

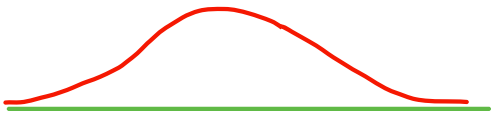
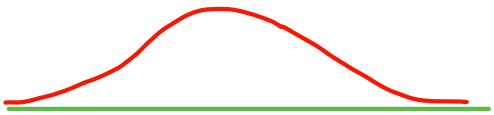
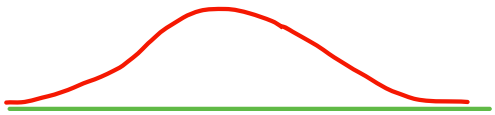


Detecting / an intruder when there isn't one.
Failing to detect an intruder when there is one.
Accepting that there is no intruder when there isn't one.
Rejecting the idea of an intruder when there is one

A security system is designed to detect uintruders 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?



Tailed Testy

Burger Example

> Suppose there is a burger place that claims that all their burgers are **200 grams**. A customer who consumed their burger is still hungry after eating, and wants to prove that their burgers are lighter, and not as much as promised.

Consider the example of the Legacy Model, which had an accuracy of 90%. You want to claim that your new model is better.

Suppose you are looking at the height of people in India. It is believed that the average height of Indians is 65 inches. You want to find out if that holds true for the people of your state. Are they taller or shorter?

Let's Revisit Coin Toss Example

Conclusion

- H_0 → Null Hypothesis
- H_a → Alternate Hypothesis
- α → Significance level
- P-value
- Errors → Type-I ← Symbol: FP
  what does it mean from H_0
  Type2 Error (FN)

Hypothesis Testing Framework