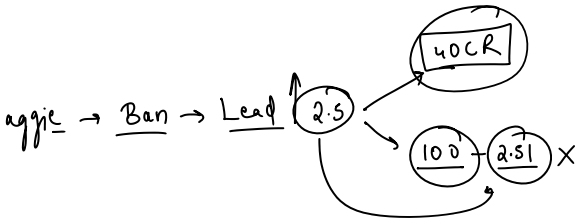
Hypothesis Testing

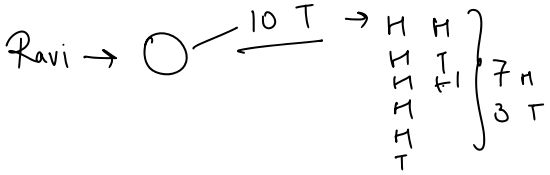
Wednesday, July 24, 2024 9:13 PM

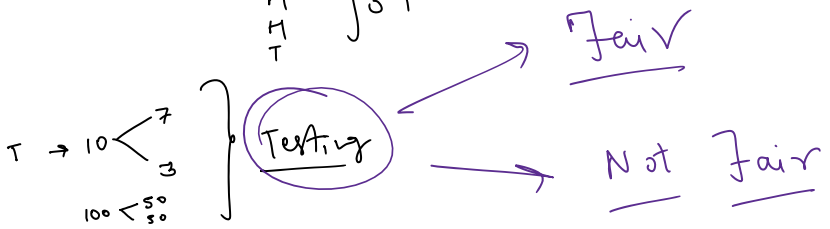
Content

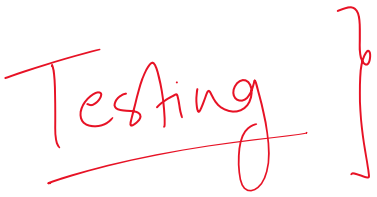
1. Module Overview [15 mins]
2. Motivation [4 mins]
3. Hypothesis Testing [10 mins]
   - Judge in court ex [5 mins]
   - ML Deployment ex [7-8 mins]
4. HT Terminologies
   - Deep dive of Judge in Court ex (H0 and Ha) [5-7 mins]
   - P-Value [5 mins]
   - **Break [5 min]**
   - Quiz 1 [5 mins]
   - Quiz 2 [3 mins]
   - Significance/Confidence Level [10 mins]
   - Quiz 3 [3 min]
5. Types of Errors [10 mins]
   - **Virus ex [3-4 mins]**
   - Quiz 4 [4 mins]
6. Tailed Tests [15 mins]
   - Quiz 5 [3 mins]
7. Deep dive in Coin Toss ex [8 mins]
8. Hypothesis Testing Framework [2-3 mins]
9. Practice Question [5 mins]

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>





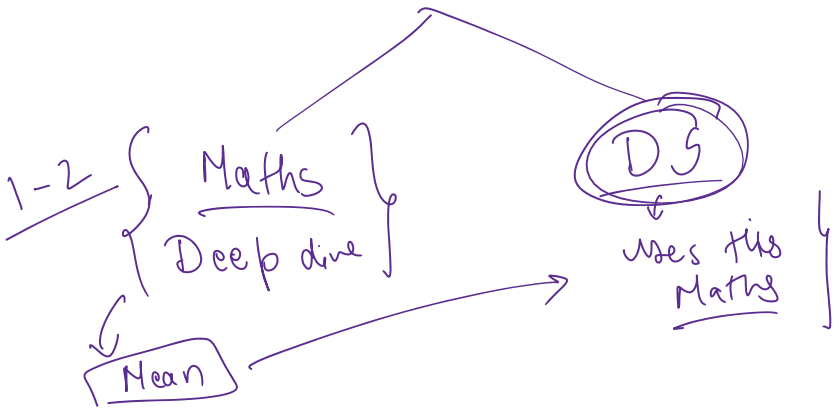




Company & Role	Responsibilities	Skills Required
Truecaller, Data Scientist	You will be responsible for collecting, organizing, analyzing, and interpreting Truecaller data and drawing insightful conclusions to enhance BU's road map, business plan, and customer experience. The individual will Possess a proactive mindset in tackling novel business problems, leveraging advanced analytical techniques and tools to extract valuable insights from large and complex datasets. You exhibit a strong inclination towards **data exploration and hypothesis testing**, with a	Able to work on machine learning algorithms and **statistical models** to uncover meaningful insights from large and complex datasets. Collecting, organizing, analyzing, and interpreting all data and drawing insightful conclusions which enable us to work more smartly. Providing regular, accurate, and comprehensive **statistical reports**. Providing objective insight and

Company & Role	Responsibilities	Skills Required
Truecaller, Data Scientist	You will be responsible for collecting, organizing, analyzing, and interpreting Truecaller data and drawing insightful conclusions to enhance BU's road map, business plan, and customer experience. The individual will Possess a proactive mindset in tackling novel business problems, leveraging advanced analytical techniques and tools to extract valuable insights from large and complex datasets. You exhibit a strong inclination towards **data exploration and hypothesis testing**, with a focus on continuously improving the accuracy and efficiency of analytical models. You will work in close collaboration with the product, engineering, and platform team to identify and follow up on metrics, evaluate **A/B tests** and user behaviour, and identify product insights. The role requires working both independently and proactively with identifying improvement opportunities as well as being involved in separate projects.	Able to work on machine learning algorithms and **statistical models** to uncover meaningful insights from large and complex datasets. Collecting, organizing, analyzing, and interpreting all data and drawing insightful conclusions which enable us to work more smartly. Providing regular, accurate, and comprehensive **statistical reports**. Providing objective insight and analysis to influence decision-making. Key Skills: Product Engineering Machine Learning Data Modelling **Hypothesis Testing** SQL, Python

IBM, Data Analyst	Our Worldwide S&A team is building a new mission called Recovery 2.0 and is looking for talented data scientists to help solve real business problems and to deliver revenue for IBM. You will be responsible for designing and developing the MVP, pulling together various datasets from IBM's backend systems into a dataset that contains recovery opportunities that can then be executed by our S&A Sellers in their territories, quarterly. You'll be responsible for cleansing the data and store in a data mart. You'll be responsible for working with stakeholders to augment the data such that our teams can productively identify the opportunities that have the highest propensity for success.	Knowledge of Python, Jupyter Notebooks Mastery of Excel, SPSS, **Statistics** or other data analysis tools Strong communication skills, with the ability to effectively communicate complicated analyses to non-technical employees and leadership Strong business acumen with the ability to develop data-driven business recommendations
Infosys, Data Scientist	Enjoy working in complex business environments, refining user requirements, ability to analyze data to answer business queries; adapt in quick starting optimization options to solve business problems using the latest technologies in data mining, data visualization, statistical modelling, pattern recognition, and optimization across the chosen industry and functional area WITHIN the guidelines, policies and norms of Infosys. Define new business metrics as required and build measurement systems as needed Advise clients & internal teams on the right analytical models, analytic platforms and approaches to take in addressing complex, open-ended business problems. The job entails sitting as well as working at a computer for extended periods of time. Should be able to communicate by telephone, email or face-to-face.	Expertise in areas of quantitative discipline such as **Statistics**, Applied Math, Operations Research, Computer Science, Engineering or Physics Experience in **exploratory data analysis**, devising, deploying and servicing **statistical models** **A/B testing**, Web analytics, click stream data analysis, and other KPIs on marketing campaigns Fluent in using one or more analytical tools Python Advanced SQL. Knowledge of visualization tool Tableau.
Xpressbees, Data Analyst	You will be responsible for analyzing and interpreting large data sets, creating and maintaining reports, identifying trends, and providing insights to the organization. As a Data Analyst will work closely with the operations team to optimize supply chain efficiencies and advise on strategic decisions.	Strong analytical skills, with the ability to **interpret large datasets and provide insights** to the organization. Proficient in **data analytics and statistics**, with experience using tools such as Excel, SQL, and Python. Excellent communication skills, with the ability to present insights clearly and effectively to stakeholders. Experience in data modelling, with knowledge of machine learning algorithms and data visualization tools.



Mean

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**

YT → Ad → IS MIN

2 ad → IS MIN → Ad blockers

Assumpti

People Irritate

Assumption

Premium ↑

Assumpti

Hypothesis

Test

- Consequences
- True / Not

Coin → 10 → 7 Heads

3 Tail

15 fair 20 unfair

Proof

→ 20 Heads

Proof → 100 → 70 Heads, 30 tail — 18 fair, 23 unfair

Rare - fair → 1000 (70%) → 700 Heads, 300 Tails — 21 fair, 25 unfair

100 L → 70 L H, 30 L T — 10 fair, 30 unfair

50 - 50 → fair — unfair

5 Judge in court

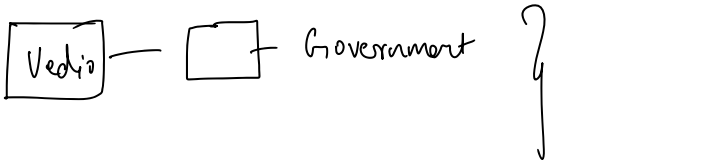
" Innocent until proven guilty "

· default hypothesis " Innocent "

assumption

YT → Recommendation → Government ?

Y1 → Recommendation



Suhag → I have a better one

default hypothesis " old one is still better "

Cricket

Third umpire

default Assump : " field umpire is correct "

Virus

default " No Virus "

Null hypothesis

: default : No change

- A Null Hypothesis is like a baseline assumption which we can attempt to disprove

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

: default : No Change

- A Null Hypothesis is like a baseline assumption which we can attempt to disprove

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

Alternate Hypothesis

Maggie : → " No Lead " → Null H_0 Hypo

→ " Lead " Alternate H_1 / H_a

→ "old Model is better" H_0

" New is better" H_1

→ " field umpire is correct" H_0

" Third umpire — H_1

How does Hypothesis Testing help the judge?

- In order to test the null hypothesis, the judge will seek evidence.
- If there is enough evidence, data to conclude that the default assumption was incorrect, the

→ innocent

→ Burden

the judge will seek evidence.

- If there is enough evidence, data to conclude that the default assumption was incorrect, the judge will **reject the null hypothesis**.
  - Then that means there is an Alternate hypothesis that is true
- If evidence is NOT enough, the judge will stick to it.

This is **Hypothesis Testing**

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

Burden proof evidence

Reject the null hypothesis

< 5000 — Null

> 5000 — Alternative

Judge

H_0 → The person is innocent

Data

- Knife in his pocket : Innocent can have knife
- Knife has blood stains : May be his cook / chef

Rare Prob

So Blood it Matches of victim

Prob low { Blood it Matches of Victim : This is too much

Prob low { # fingerprint of the victim :

Rare ↓ (low) { P (Data | ~~innocent~~ given) }

{ Rare ↓ P (Data | given innocent) × }

P value

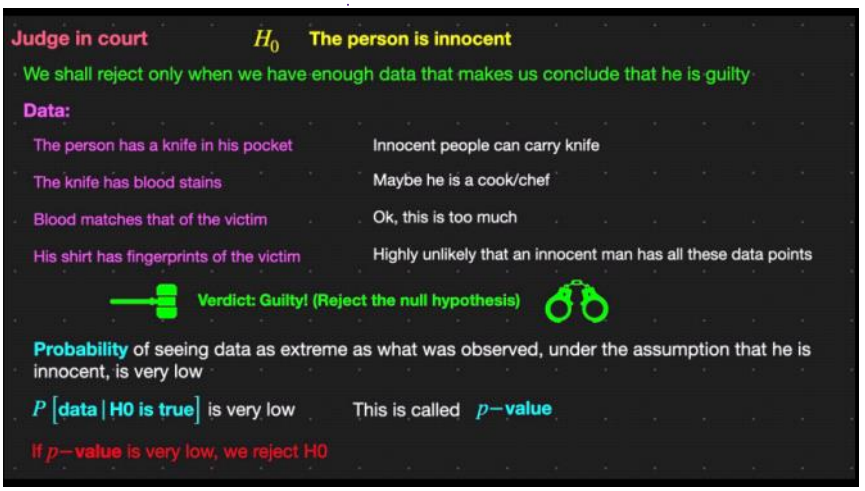
Prob — low
Rare
innocent → ~~Not~~ innocent

common
Mirzapur ← P (GUN | innocent)
↓
90%

P < less → (~~Rare~~ | ~~Null~~) Reject

P >> Common — Null

• Class → Notes



μ < 8

old is b

8 <

Claim

Juice New < 8 gram

The juice brand claims that the mean sugar content in its juice boxes, produced using a new manufacturing process, is less than 8 grams.
The Food Safety and Standards Authority of India (FSSAI) wants to test the null hypothesis in this scenario.

Set up the null (H0) and alternate (H1) hypotheses for this claim, and choose the correct option:
33 users have participated
A

H0: The sugar content in the juice boxes produced using the new manufacturing process is not equal to 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is equal to 8 grams.
9%
B

H0: The sugar content in the juice boxes produced using the new manufacturing process is greater than 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is less than or equal to 8 grams.
12%
C

H0: The sugar content in the juice boxes produced using the new manufacturing process is less than 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is not equal to 8 grams.
39%
D

H0: The sugar content in the juice boxes produced using the new manufacturing process is greater than or equal to 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is less than 8 grams.
39%

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

$h_0 \geq$ 8 gram

h_1 (ha) < 8 gram

H0 — Null Hypothesis old is better

Claim → Alternate sales → My...

H0

Claim → Alternate

Sales → My Machine is better

- Null Hypothesis ($H0$$H0$): This represents the status quo or no effect,
- Alternate Hypothesis ($H1$$H1$): This represents the effect or change.
- Default assumption is Coffee shop makes amazing coffee so this will be a Null Hypothesis (H0)
- Sale's person is claiming that his machine makes better coffee so this will be Alternate Hypothesis (Ha)

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

H_0 Null — Innocent

H_1 Alternate — Not Innocent

Data to reject Null hypothesis

P (Data | Innocent

low ↓ P (Data | H_0) — Rare ↓

Reject the Null

Threshold → 5 % or 1 %. Medical

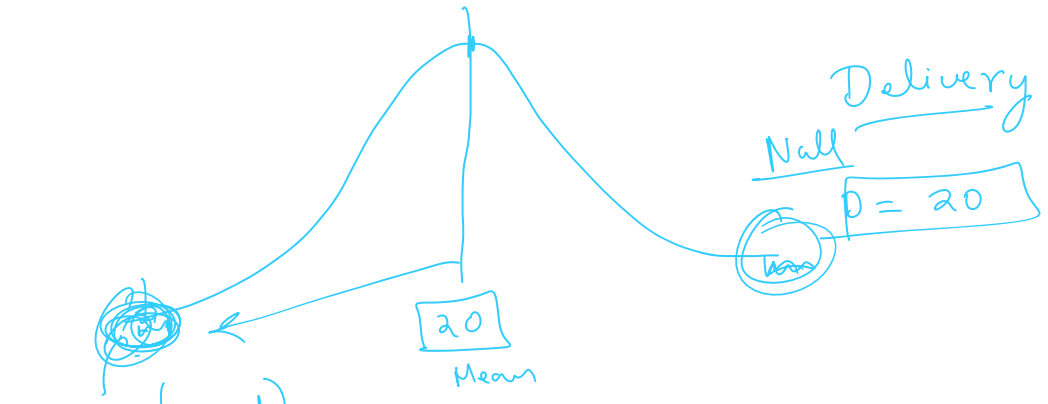
P. low (knife blood stains | Innocent)

P < 5 %.

(old is better)

Reject the Null hypothesis

P(low < 1%)



P(Rare | good)

Prob

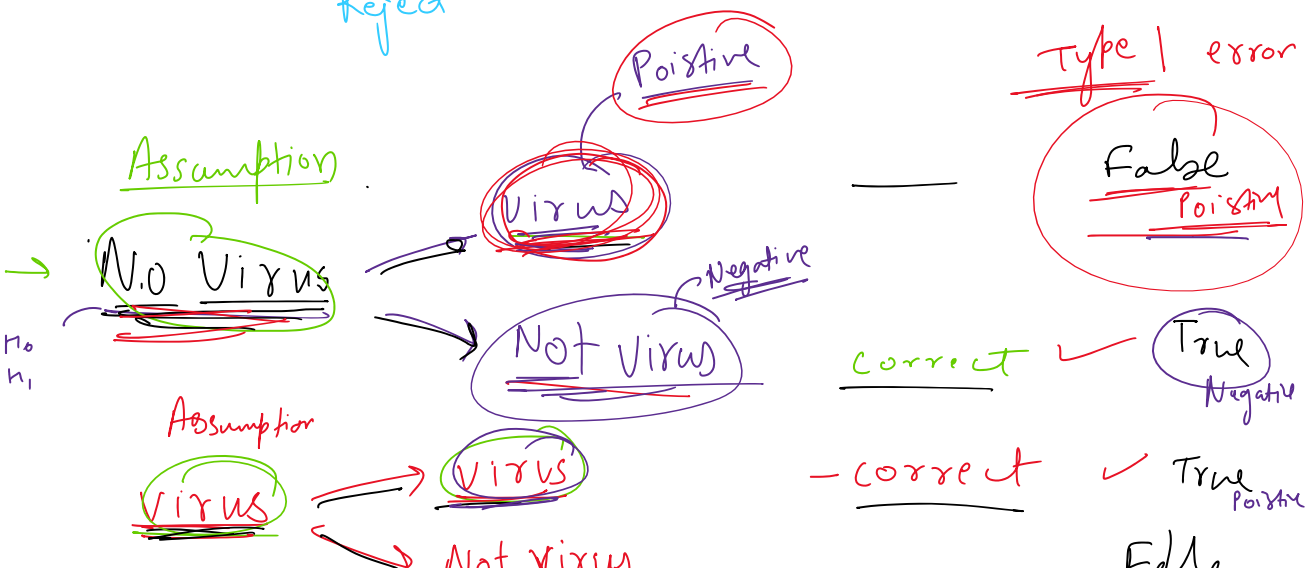
0.05 ← 5%.

0.0001 < 0.05 } Reject Null.

5%.

low [1 < 5] 100

Reject



Virus

Not Virus

False Negative

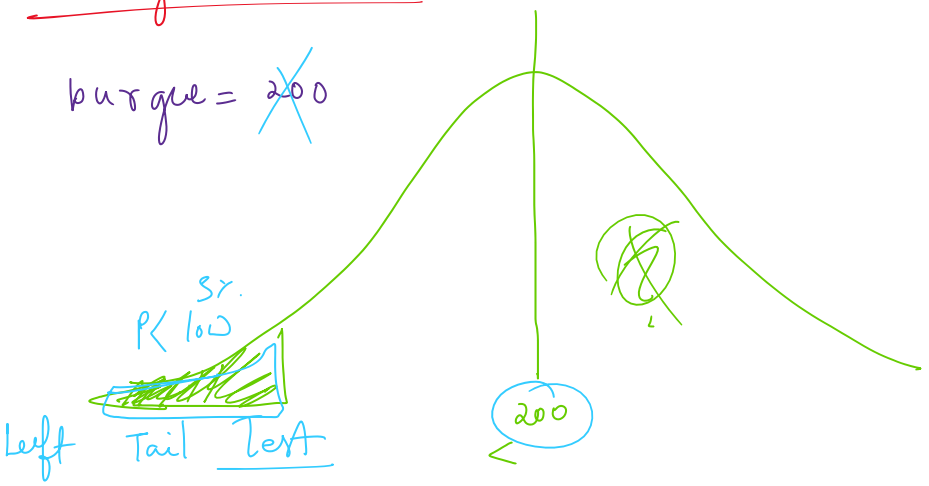
Type 2 error

MCD

burger

H_1 = burger < 200

H_0 = burger = 200



% T — 90 %.

Yogesh > 90 — claim.

H_1 Yogesh > 90

H_0 old = 90

P (Data | H_0)

H0

P (Data | H0)

P low

Rare

P < 0.05

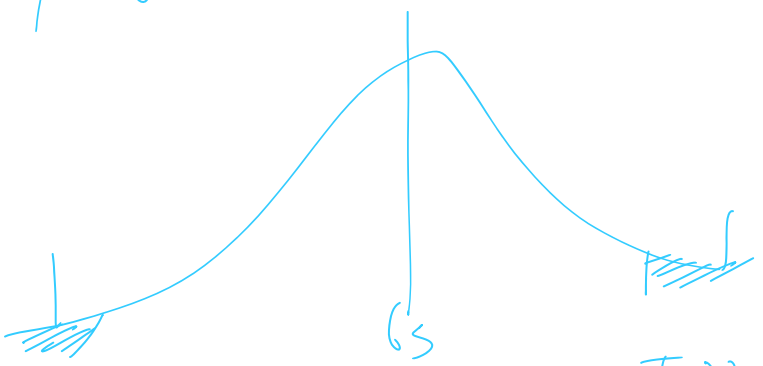
Right Tail

P (| H0)

low

Avg = 65 H0

Avg ≠ 65 H1



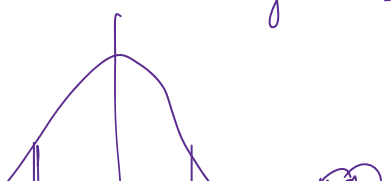
Two Tail Test

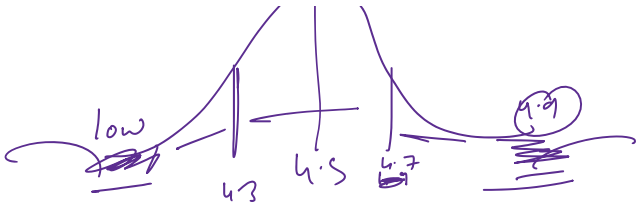
Avg = 4.5 ← H0

Avg ≠ 4.5

Avg < 4.5

Avg > 4.5



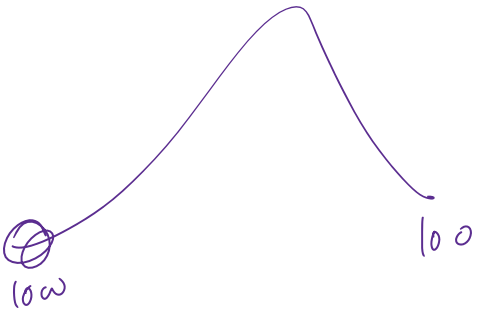


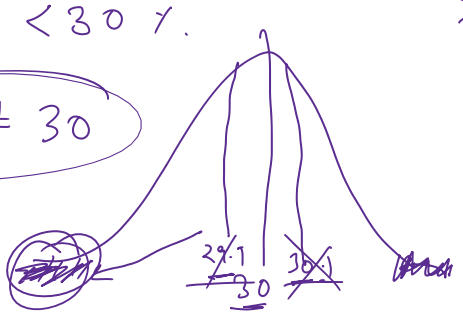
low

TAX → H_0 ~~30%~~ TAX

H_1 → < 30 %. >30%.

$H_1 \neq 30$





65 X — 1000

64.9 65.1

Arch

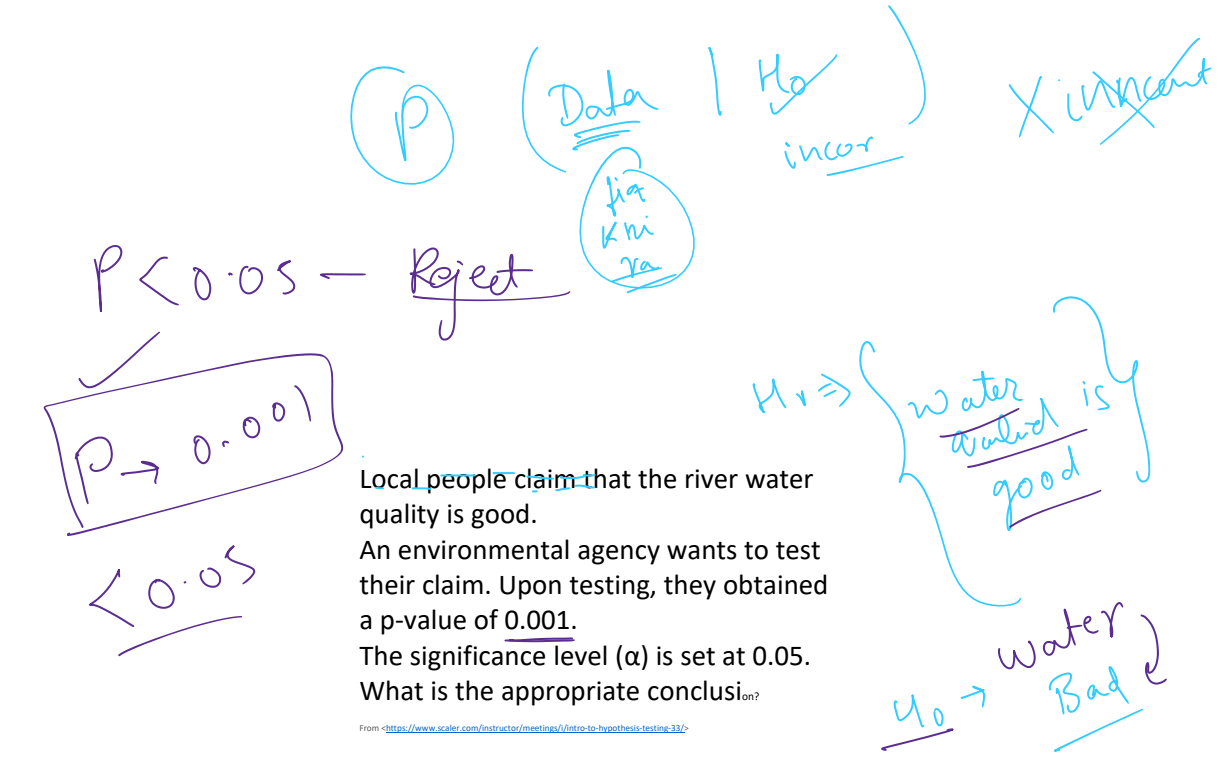
H_0 = Score ≤ 65

H_1 = Score > 65 Right

H_1 > Right

H_1 < left

≠ two

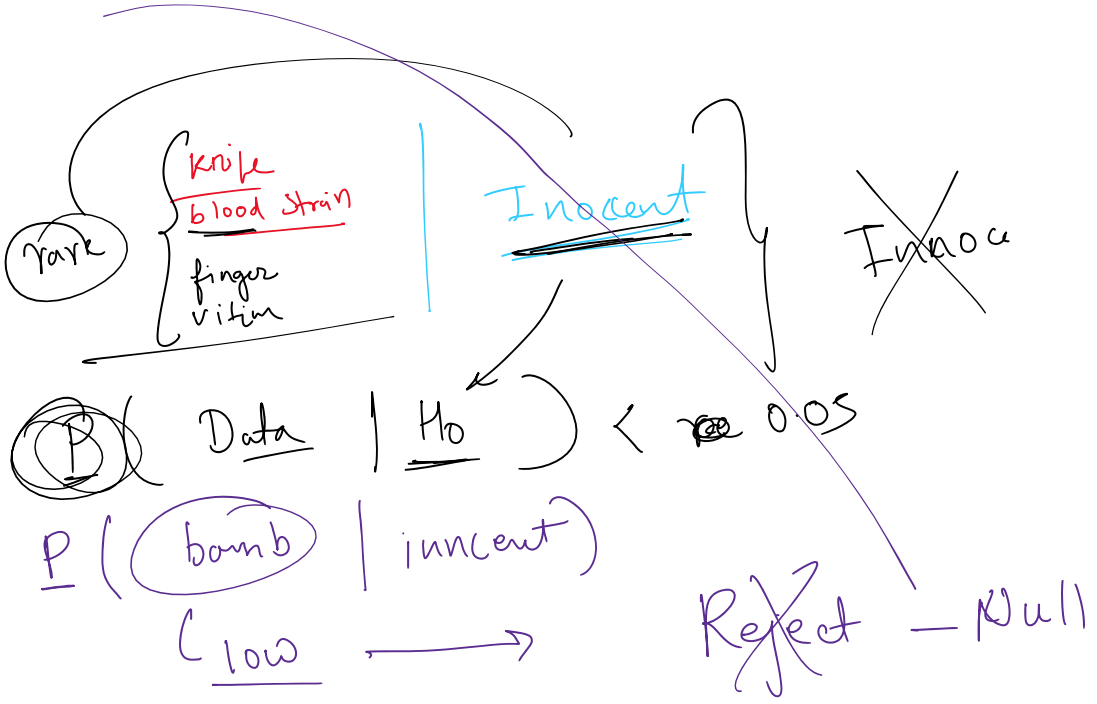


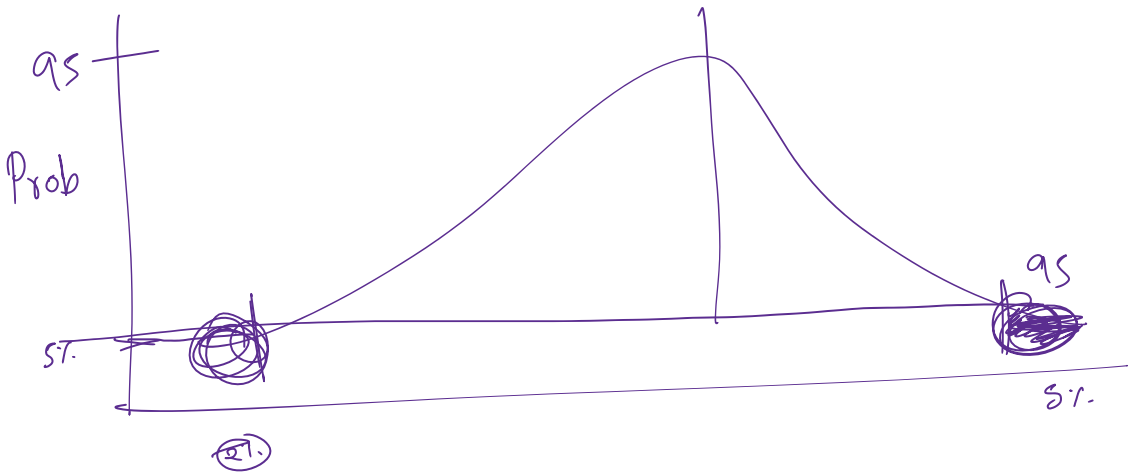
Local people claim that the river water quality is good.
An environmental agency wants to test their claim. Upon testing, they obtained a p-value of 0.001.
The significance level (α) is set at 0.05.
What is the appropriate conclusion?

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

A
Reject the null hypothesis
B
Fail to reject the null hypothesis
C
Increase the significance level
D
Conduct additional tests

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>





P (Phone Book | innocent)

95%.

Cannot reject Null

Juice Claim → J < 8 grams — H_1

$H_0 \geq 8$

The juice brand claims that the mean sugar content in its juice boxes, produced using a new manufacturing process, is less than 8 grams.
The Food Safety and Standards Authority of India ~~(FSSAI)~~ wants ~~to test the null hypo~~thesis in this scenario.

Set up the null (H0) and alternate (H1) hypotheses for this claim, and choose the correct option:

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

.
A
H0: The sugar content in the juice boxes produced using the new manufacturing process is not equal to 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is equal to 8 grams.
B
H0: The sugar content in the juice boxes produced using the new manufacturing process is greater than 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is less than or equal to 8 grams.
C
H0: The sugar content in the juice boxes produced using the new manufacturing process is less than 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is not equal to 8 grams.
D
H0: The sugar content in the juice boxes produced using the new manufacturing process is greater than or equal to 8 grams. H1: The sugar content in the juice boxes produced using the new manufacturing process is less than 8 grams.

From <https://www.scaler.com/instructor/meetings/i/intro-to-hypothesis-testing-33/>

YT → Claim

Better than you

H_a → CAR → Avg 200 KM / Litre

Avg $\leq$ 200