

# **SNS COLLEGE OF TECHNOLOGY**

Kurumbapalayam (Po), Coimbatore – 641 035

**An Autonomous Institution**

Accredited by NAAC – UGC with ‘A++’ Grade

Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai



## **DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

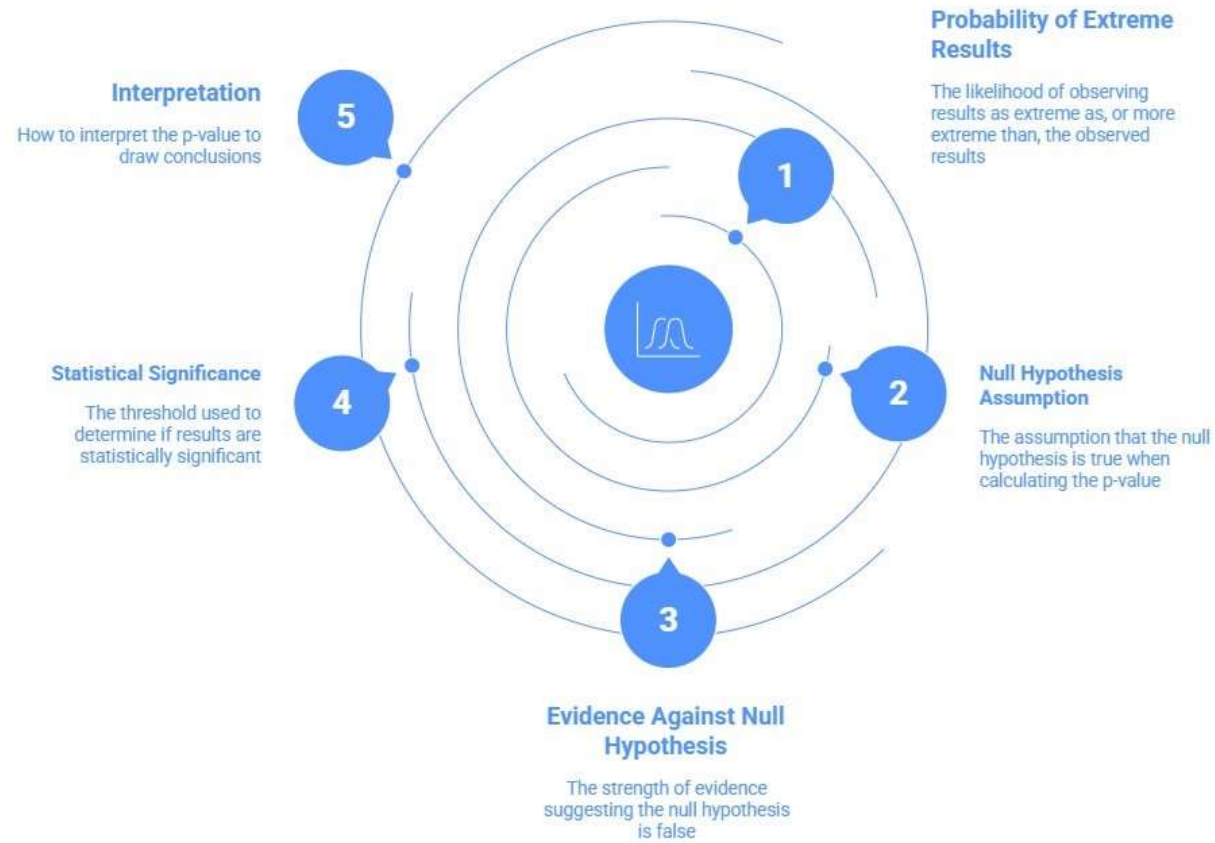
**COURSE NAME: 23ADO201- DATA SCIENCE FUNDAMENTALS  
(OPEN ELECTIVE )**

**III YEAR / VI SEMESTER**

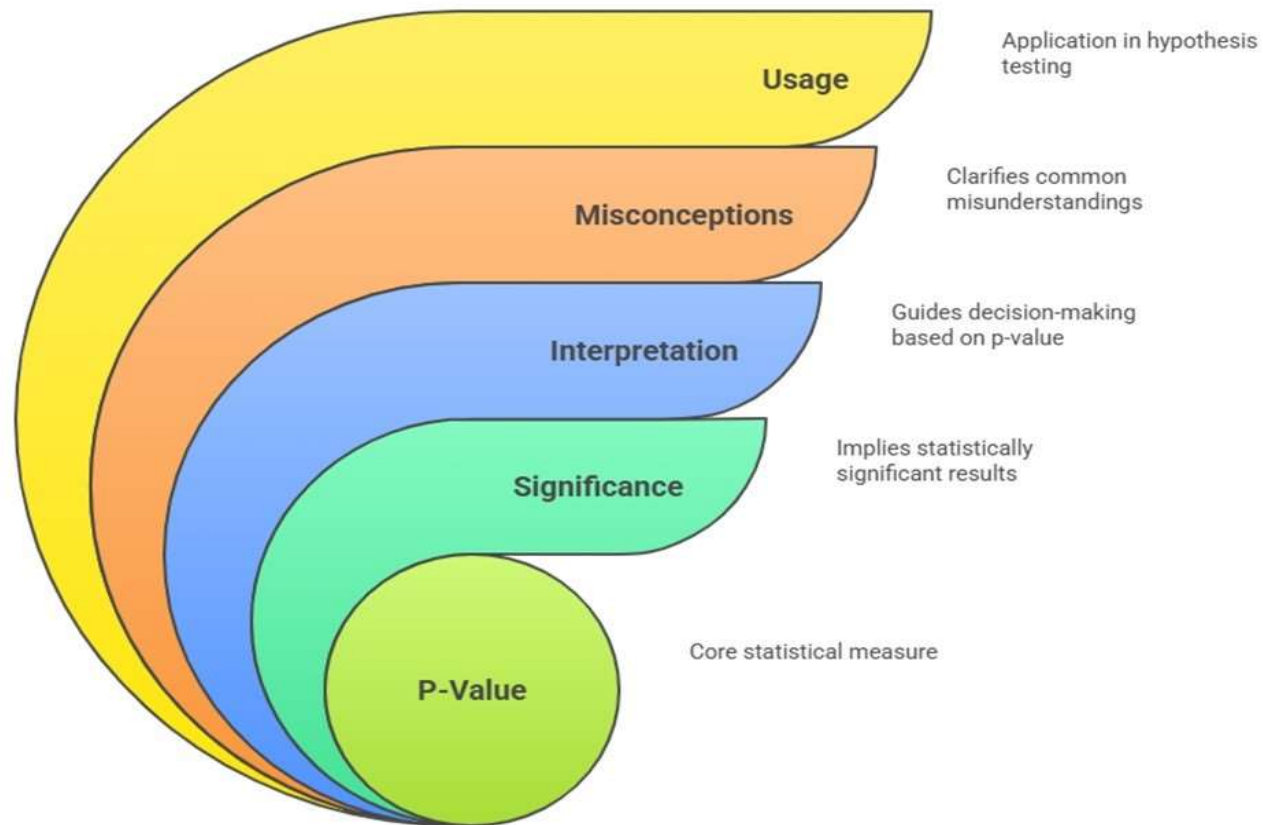
**Unit 4 - ANALYSIS OF VARIANCE**

**Topic : P-value**

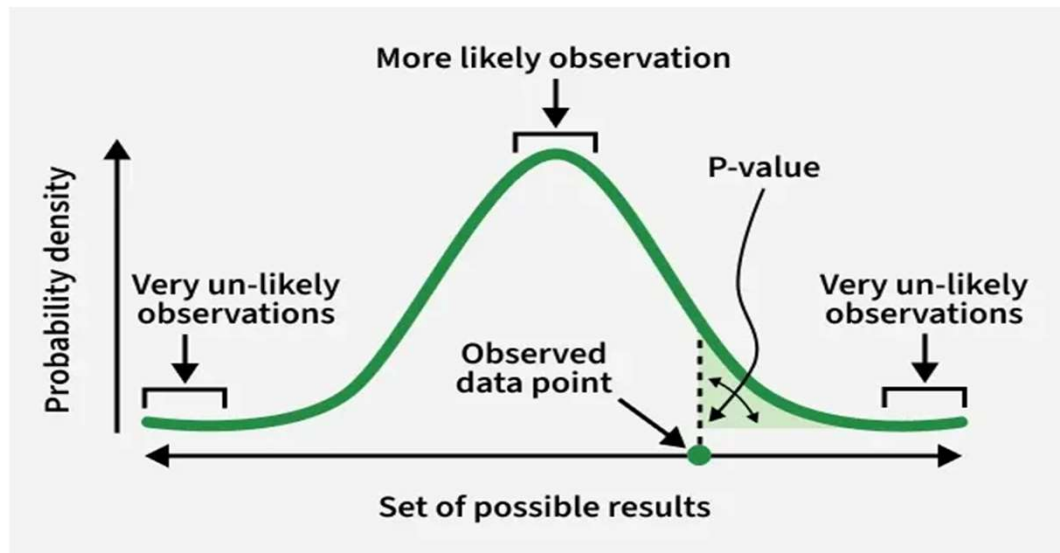
## Understanding the P-value



## P-Value Interpretation



A p-value (probability value) is a statistical measure used in [hypothesis testing](#) to help decide whether the results of an experiment are meaningful or likely due to random chance.



In simple terms it answers the question “If nothing unusual is happening, how surprising are these results?”

- A small p-value means the observed results are unlikely to occur by chance alone providing strong evidence against the null hypothesis.
- A large p-value suggests the results are consistent with random variation and do not provide enough evidence to reject the null hypothesis.

## How to Calculate P-Value



### State the Null Hypothesis

Define the baseline assumption of no effect

### Choose the Alternative Hypothesis

Specify the expected effect or difference

### Select the Test and Calculate the Test Statistic

Choose a suitable test and compute the statistic

### Determine the Sampling Distribution

Identify the probability distribution of the statistic

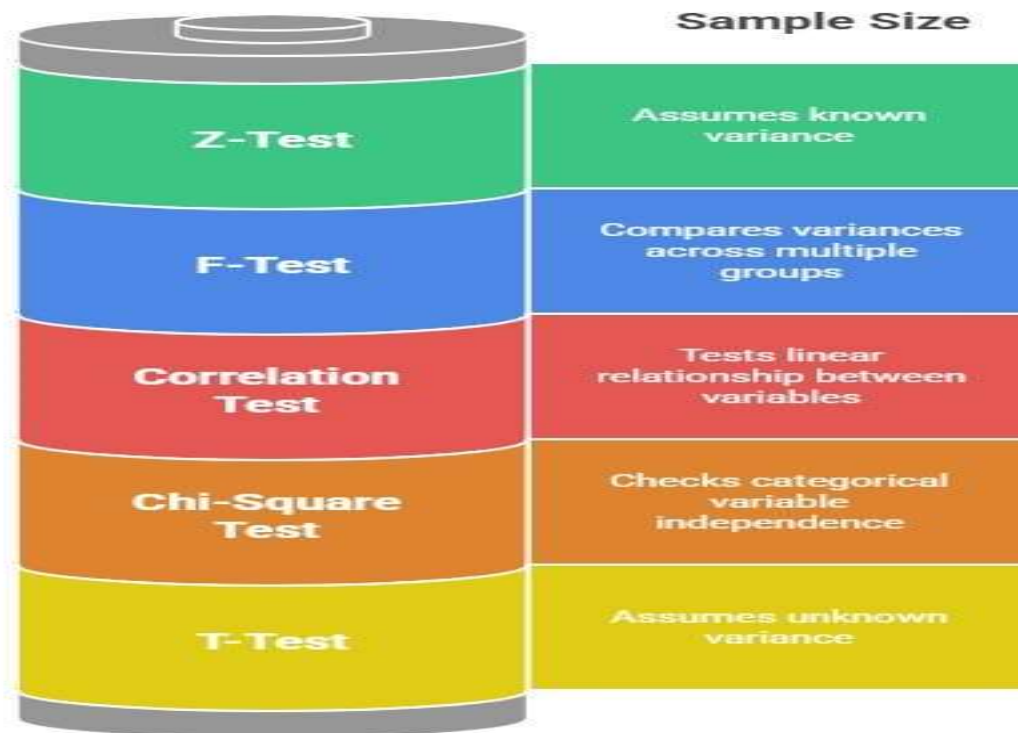
### Calculate the P-Value

Find the probability of observing an extreme result

### Make a Decision

Compare the p-value to the significance level

## Statistical tests ranked by sample size and variance knowledge



4/12/2026

4/12/2026

## Calculating P-Value Using Two-Sample T-Test

Suppose a researcher wants to investigate whether there is a significant difference in mean height between males and females in a university population.

Data:

- Group 1 (Males):  $n_1=30$ ,  $\bar{x}_1=175$ ,  $s_1=5$
- Group 2 (Females):  $n_2=35$ ,  $\bar{x}_2=168$ ,  $s_2=6$

1. **Null Hypothesis ( $H_0$ ):** There is no significant difference in mean height between males and females.

2. **Alternative Hypothesis ( $H_1$ ):** There is a significant difference in mean height between males and females.

3. **Test Statistic:** For two independent samples, the two-sample t-test is used:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where

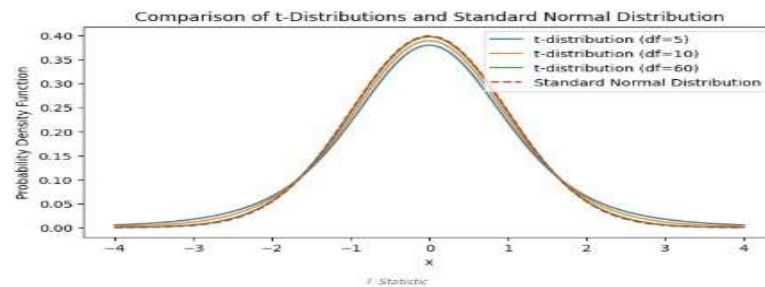
- $\bar{x}_1$ : mean of the first sample
- $\bar{x}_2$ : mean of the second sample
- $s_1$ : First sample standard deviation
- $s_2$ : Second sample standard deviation
- $n_1$ : First sample sample size
- $n_2$ : Second sample's sample size

$$t = \frac{175 - 168}{\sqrt{\frac{0.8333}{30} + \frac{1.0286}{35}}} = \frac{7}{\sqrt{0.8333 + 1.0286}} = \frac{7}{1.364} \approx 5.13$$

The calculated t-statistic is 5.13.

**4. Distribution and Degrees of Freedom:** The t-distribution is used for small samples or unknown population standard deviation.

$$df = (n_1 + n_2) - 2 = (30 + 35) - 2 = 63$$



The t-distribution is symmetric and bell-shaped similar to the normal distribution. As the degrees of freedom increase the t-distribution approaches the shape of the standard normal distribution. Practically it affects the critical values used to determine statistical significance and confidence intervals.

#### 5. P-Value Calculation

The p-value for a two-tailed test is calculated using the t-distribution.

```
import scipy.stats as stats
t_statistic = 5.13
degrees_of_freedom = 63
p_value = 2 * (1 - stats.t.cdf(abs(t_statistic), degrees_of_freedom))
print("P value:", p_value)
```

**Output:**

P value: 2.9918663893013786e-06

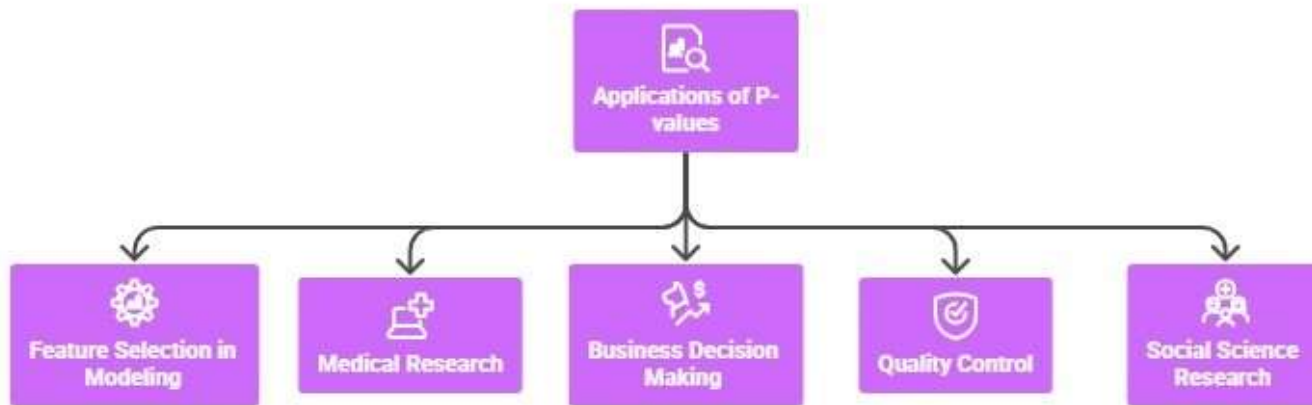
#### 5. Decision Rule

At a significance level of  $\alpha = 0.05$

- If p-value  $\leq \alpha$  reject  $H_0$
- If p-value  $> \alpha$  fail to reject  $H_0$

Since the p-value is much smaller than 0.05, the null hypothesis is rejected.

## Applications of P-values



# Activity



## Activity Title: Understanding p-values through Coin Toss

### Objective

Help students understand what a **p-value** means in hypothesis testing using a real-life experiment.

### Materials Needed

- 1 coin
- Paper & pen (for recording results)

### Step-by-Step Procedure

#### 1. State the Hypotheses

- Null Hypothesis (H<sub>0</sub>):** The coin is fair (equal chance of heads & tails)
- Alternative Hypothesis (H<sub>1</sub>):** The coin is biased

#### 2. Conduct the Experiment

- Toss the coin **20 times**
- Record the number of heads

#### 3. Example Observation

Suppose you get:

- Heads = 16
- Tails = 4

#### 4. Discuss with Students

Ask:

•“If the coin is fair, is getting 16 heads out of 20 common or rare?”

#### 5. Explain p-value (Conceptually)

- The **p-value** is the probability of getting results **this extreme (or more extreme)** assuming the coin is fair.
- If this probability is **very small**, we doubt the coin is fair.

4/12/2026

10

### Visualizing the Idea

$$P(X \geq 16) \text{ where } X \sim \text{Binomial}(n = 20, p = 0.5)$$

This represents the probability of getting **16 or more heads** out of 20 tosses.

### Interpretation

- If **p-value < 0.05** → Reject H<sub>0</sub> (coin may be biased)
- If **p-value ≥ 0.05** → Do not reject H<sub>0</sub> (coin seems fair)

### Student Discussion Questions

- 1.What does a small p-value indicate?
- 2.Does a large p-value prove the coin is fair?
- 3.What happens if we increase the number of tosses?

4/12/2026

## MCQ



### MCQs on p-values

#### 1. What does a p-value represent?

- A) Probability that null hypothesis is true
- B) Probability of observed result under null hypothesis
- C) Probability of alternative hypothesis
- D) Level of significance

**Answer: B**

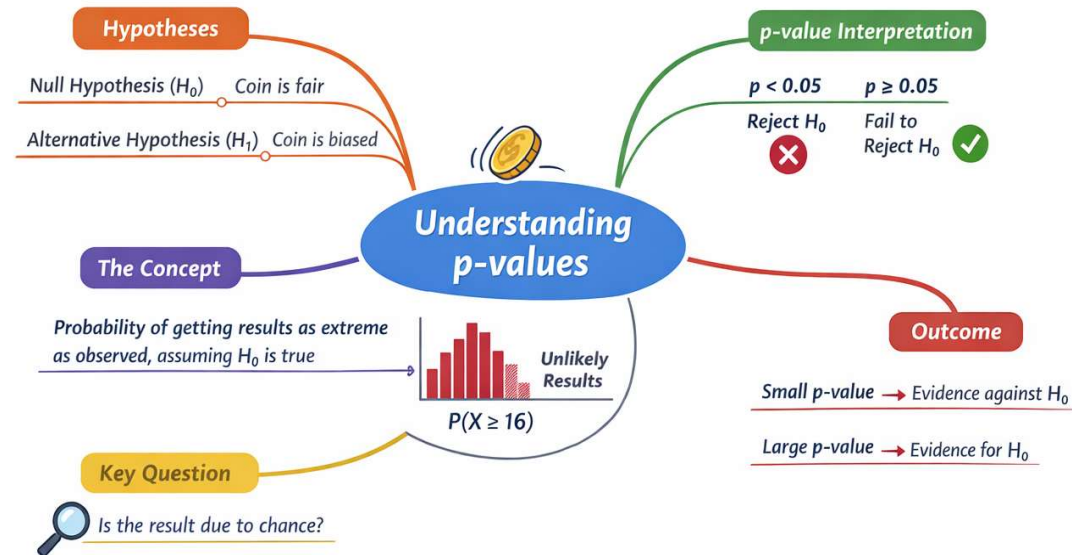
#### 2. A small p-value ( $\leq 0.05$ ) indicates:

- A) Strong evidence against null hypothesis
- B) Strong evidence for null hypothesis
- C) No evidence
- D) Data is incorrect

**Answer: A**

11

# MINDMAP



TEXT BOOKS	
1.	Srinivasan Desikan and Gopalaswamy Ramesh, —Software Testing – Principles and Practices, Pearson Education, 2006.
REFERENCES	
1.	Ilene Burnstein, —Practical Software Testing, Springer International Edition, 2003
2.	Edward Kit, Software Testing in the Real World – Improving the Process, Pearson Education, 1995
3.	Boris Beizer, Software Testing Techniques – 2nd Edition, Van Nostrand Reinhold, New York, 1990.
4.	Aditya P. Mathur, —Foundations of Software Testing _ Fundamental Algorithms and Techniques, Dorling Kindersley (India) Pvt. Ltd., Pearson Education, 2008.

**THANK  
YOU!**

14