

# Statistics and Data Analysis

## Unit 03 – Lecture 02: Hypothesis Testing (t-test): Concepts and Setup

Tofik Ali

School of Computer Science, UPES Dehradun

February 14, 2026

<https://github.com/tali7c/Statistics-and-Data-Analysis>

# Quick Links

Overview

t-test Basics

p-values

Exercises

Demo

Summary

# Agenda

1 Overview

2 t-test Basics

3 p-values

4 Exercises

5 Demo

6 Summary

# Learning Outcomes

- Define null and alternative hypotheses clearly

# Learning Outcomes

- Define null and alternative hypotheses clearly
- Compute a one-sample t statistic (given summary)

# Learning Outcomes

- Define null and alternative hypotheses clearly
- Compute a one-sample t statistic (given summary)
- Explain p-value and significance level  $\alpha$

# Learning Outcomes

- Define null and alternative hypotheses clearly
- Compute a one-sample t statistic (given summary)
- Explain p-value and significance level  $\alpha$
- Distinguish one-tailed vs two-tailed tests

# Learning Outcomes

- Define null and alternative hypotheses clearly
- Compute a one-sample t statistic (given summary)
- Explain p-value and significance level  $\alpha$
- Distinguish one-tailed vs two-tailed tests
- State key assumptions behind the t-test



# t-test Basics: Key Points

- $H_0/H_1$  setup

# t-test Basics: Key Points

- $H_0/H_1$  setup
- Test statistic measures how far the sample is from  $H_0$

# t-test Basics: Key Points

- $H_0/H_1$  setup
- Test statistic measures how far the sample is from  $H_0$
- Assumptions: independence, outliers, normality/CLT

# t-test Basics: Key Formula

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \quad \text{df} = n - 1$$

# p-values: Key Points

- p-value = probability of data (or more extreme) assuming  $H_0$

# p-values: Key Points

- p-value = probability of data (or more extreme) assuming  $H_0$
- Small p-value  $\rightarrow$  evidence against  $H_0$

# p-values: Key Points

- p-value = probability of data (or more extreme) assuming  $H_0$
- Small p-value  $\rightarrow$  evidence against  $H_0$
- p-value is not effect size

# Exercise 1: Write hypotheses

Claim: mean score is 60. Write  $H_0$  and  $H_1$  for a two-sided test.



# Solution 1

- $H_0: \mu = 60$
- $H_1: \mu \neq 60$

## Exercise 2: Compute t

Given  $n=25$ ,  $\bar{x}=53$ ,  $s=10$ , test  $H_0: \mu=50$ . Compute t.

## Solution 2

- $SE = 10/\sqrt{25} = 2$
- $t = (53-50)/2 = 1.5$
- $df = 24$

## Exercise 3: Tail choice

You want to show a new method increases mean score. One-tailed or two-tailed?

## Solution 3

- One-tailed (right):  $H_1: \mu \geq \mu_0$

# Mini Demo (Python)

Run from the lecture folder:

```
python demo/demo.py
```

Outputs:

- images/demo.png
- data/results.txt

# Demo Output (Example)



# Summary

- Key definitions and the main formula.



# Summary

- Key definitions and the main formula.
- How to interpret results in context.

# Summary

- Key definitions and the main formula.
- How to interpret results in context.
- How the demo connects to the theory.

# Exit Question

Why can a very small p-value still be unimportant in practice?