

# Statistics and Data Analysis

## Unit 03 – Lecture 01: Population, Sample, Sampling, Estimation and Confidence Intervals

Tofik Ali

School of Computer Science, UPES Dehradun

February 17, 2026

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

# Quick Links

Overview

Sampling

Confidence Intervals

Exercises

Demo

Summary

# Agenda

- 1 Overview
- 2 Sampling
- 3 Confidence Intervals
- 4 Exercises
- 5 Demo
- 6 Summary

# Learning Outcomes

- Differentiate population parameters and sample statistics

# Learning Outcomes

- Differentiate population parameters and sample statistics
- Explain sampling bias vs random error

# Learning Outcomes

- Differentiate population parameters and sample statistics
- Explain sampling bias vs random error
- Describe common sampling methods (SRS, stratified, cluster)

# Learning Outcomes

- Differentiate population parameters and sample statistics
- Explain sampling bias vs random error
- Describe common sampling methods (SRS, stratified, cluster)
- Compute and interpret a basic confidence interval for a mean

# Sampling: Key Points

- Population vs sample



# Sampling: Key Points

- Population vs sample
- Bias vs random error

# Sampling: Key Points

- Population vs sample
- Bias vs random error
- Representative sampling matters

# Confidence Intervals: Key Points

- Interpretation: long-run coverage

# Confidence Intervals: Key Points

- Interpretation: long-run coverage
- Width depends on  $n$  and variability

# Confidence Intervals: Key Points

- Interpretation: long-run coverage
- Width depends on  $n$  and variability
- CIs support decision-making with uncertainty

# Confidence Intervals: Key Formula

$$\bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

# Exercise 1: Parameter vs Statistic

Give one example of a parameter and one of a statistic.

# Solution 1

- Parameter: population mean
- Statistic: sample mean



## Exercise 2: CI Interpretation

In one sentence, what does a 95% CI mean (correctly)?

## Solution 2

- About 95% of such intervals contain the true mean in repeated sampling.

## Exercise 3: Bias Scenario

Why does convenience sampling create bias?

## Solution 3

- Because some groups are over/under-represented systematically.

# Mini Demo (Python)

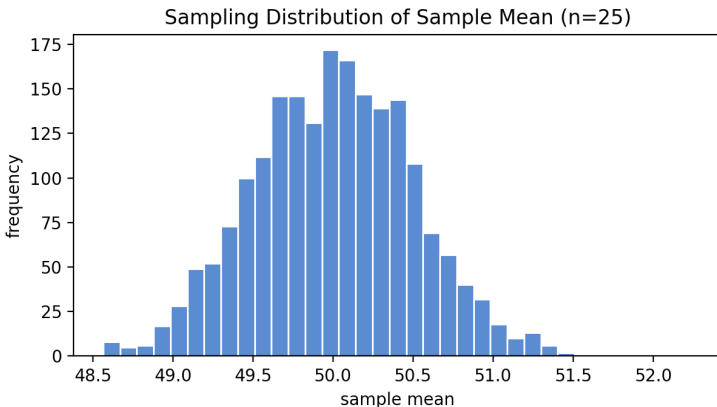
Run from the lecture folder:

```
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

If your CI is too wide, what two actions reduce its width (without cheating)?