Alexandria University
Faculty of Engineering

CS 112: Statistics
Spring 2025
Lab 3
Assigned: 18/4/2025
Due: 30/4/2025

Estimating the Average Height of Students Using Confidence Intervals

Objective

By the end of this assignment, students will be able to simulate population data, extract random samples, calculate confidence intervals for the mean, visualize distributions and intervals, and understand the impact of sample size for confidence interval estimation.

Requirements

- Write and run an interactive Python notebook (ipynb)
- You can use Colab or your local machine
- Write a clean and commented code and do the explanations in markdown cells.
- Use matplotlib, numpy, scipy, seaborn or stats libraries

Submission

- You have to submit the ipynb
- Work on this lab individually

Problem Statement

In this assignment, you take on the role of a researcher studying the average student height. Since measuring the entire population is impractical, you rely on sampling and statistical inference methods to estimate the population average with a degree of confidence.

Steps to Implement

Data

Generate a **random** population of **size** 10000, with a **mean** of 170 and **a standard deviation** of 10. This will be the population that you will sample from.

Task 1: Visualize the population

- Plot a histogram of the population distribution.
- Add a vertical line at the population mean.

Task 2: Take a Sample and Compute 95% CI

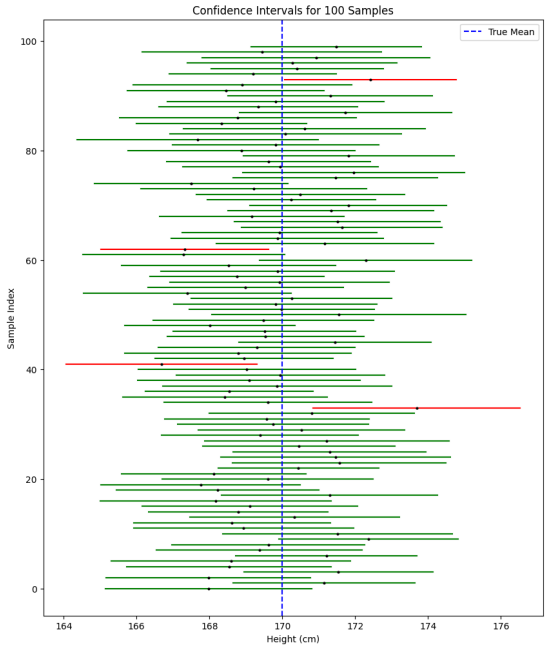
- Randomly select 50 students from the population we generated. **(DO NOT GENERATE NEW SAMPLES)**
- Compute sample mean, standard deviation, and 95% confidence interval.
- **Write the equation** used for calculating the 95% confidence interval

Task 3: Repeat Sampling 100 Times

- Repeat the sampling and CI calculation 100 times.
- Count how many intervals contain the **true population mean**.
- Count how many intervals **DO NOT** contain the **true population mean**. Write them in a **CSV** file that has the lower bound, upper bound, sample mean, and sample standard deviation

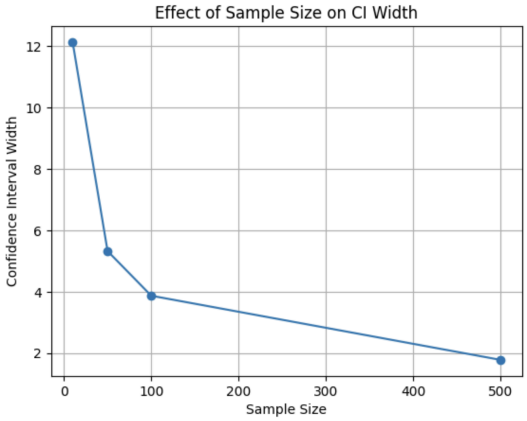
Task 4: Visualize All Confidence Intervals

- Plot 100 confidence intervals.
- Color them differently depending on whether they contain the true mean.



Task 5: Effect of Sample Size

- Calculate the 95% confidence interval for different sample sizes: 10, 50, 100, 500.
- Plot confidence interval width as a function of sample size.



Reporting

- You are required to deliver every task with its requirements and code.
- Answer the following questions:
  - What does the confidence interval represent in this context?
  - Why do some CIs not contain the true mean in Task 3?
  - How does sample size affect the confidence interval width? What does that imply?