Sample Mean and Population Mean

Week 4
Lecture 1
Spring 2025
Rida Maryam

Introduction

- In statistics, we often need to find the **average** of a group of numbers to summarize large amounts of data.
- If we only have data from a **sample**, we calculate the **sample mean**, which is used to estimate the population mean.

The Sample Mean and Population Mean

• The mean helps summarize large data sets.

• Population Mean (μ): The average of an entire population.
(If we have data for an **entire population**, we calculate the **population mean**.)

• Sample Mean ($\bar{x}$): The average of a subset of the population.
(If we only have data from a **sample**, we calculate the **sample mean)**

• Sample means are used to estimate population means.

Real-Life Importance

- **Business:** A company may want to know the average salary of employees but can only survey 500 workers instead of the entire workforce.
- **Health:** Researchers may estimate the average blood pressure of adults by testing a sample of 1,000 people rather than millions.

Formula of Population and Sample Mean

- Population Mean (μ):

  Formula: $\mu = \Sigma X / N$
- Sample Mean ($\bar{x}$):

  Formula: $\bar{x} = \Sigma X / n$
- Difference: Sample mean is based on limited data, while population mean includes all data.

Example 1: Population Mean Calculation

Scenario: A company has 5 employees with salaries: 3000, 3200, 2900, 3100, 3050.

Solution:

$$\mu = (3000 + 3200 + 2900 + 3100 + 3050) / 5$$

$$\mu = 3050$$

Interpretation: The average salary of all employees is $3050.

Example 2: Sample Mean Calculation

Scenario: A sample of 3 employees has salaries: 3000, 3200, 2900.

Solution:

$\bar{x} = (3000 + 3200 + 2900) / 3$

$\bar{x} = 3033.33$

Interpretation: Sample mean is close but not exactly equal to the population mean due to sampling variability.

Importance of the Sample Mean

- Estimating large population values when full data is unavailable.
- Used in healthcare (e.g., average heart rate), economics (e.g., average rent), and education (e.g., test scores).

Class Activity: Compute and Compare Means

Task:

1. Compute the population mean for given age data: 22, 24, 25, 28, 30, 32, 35, 38, 40, 42.
2. Select 5 numbers and compute the sample mean.
3. Compare both means and discuss differences.

Confidence Intervals and Their Interpretation

Week 4
Lecture 2
Spring 2025
Rida Maryam

Confidence Intervals and Their Interpretation

- Sample means are not always equal to population means, but they provide useful approximations.
- Every sample is slightly different, leading to some **uncertainty** in our estimates.
- **Confidence Intervals (CIs)** help us measure this uncertainty by providing a range of values within which the true population parameter is likely to fall.
- They account for sample variability and help in decision-making.
- Used in statistics, research, medicine, economics, and business.

Definition of Confidence Interval

• Confidence Interval (CI): A range of values that likely contains the true population mean.

Formula:

$$\bar{x} \pm Z \times (\sigma / \sqrt{n})$$

Where:

$\bar{x}$ (Sample Mean) = Mean of the sample

Z (Z-score) = A number based on confidence level (e.g., 1.96 for 95%)

σ (Population Standard Deviation) = Spread of the population data

n (Sample Size) = Number of observations in the sample

Z-values for Confidence Intervals

Confidence Level	Z Value
70%	1.036
75%	1.150
80%	1.282
85%	1.440
90%	1.645
95%	1.960
98%	2.326
99%	2.576
99.5%	2.807
99.9%	3.291
99.99%	3.891
99.999%	4.417

Confidence Interval Interpretation:

- A **95% confidence interval** means that if we take 100 different samples, about **95 of them** will contain the true population mean.
- A **larger confidence level (99%)** gives a **wider** interval, while a **smaller confidence level (90%)** gives a **narrower** interval.

Example 1: 95% Confidence Interval for Average Height

- A researcher surveys **100 students** and finds that their average height is **170 cm**, with a **standard deviation of 10 cm**.

Find the **95% confidence interval** for the true average height of all students.

Solution:

Given:

$\bar{x}$=170 cm (sample mean)

σ=10 cm (population standard deviation)

n=100 (sample size)

Z-score for 95% confidence = **1.96**

$$\bar{x} \pm Z \times (\sigma / \sqrt{n})$$

Margin of Error = 1.96 × (10/√100) = 1.96 × 1 = 1.96

CI = 170 ± 1.96 → (168.04, 171.96)

Interpretation:

We are 95% confident that true height is between 168.04 cm and 171.96 cm.

Example 2: 99% Confidence Interval for Test Scores

- A professor tests **50 students**, and the sample mean score is **78**, with a standard deviation of **12**. Find the **99% confidence interval** for the true mean test score.
- Z-score for 99% CI = 2.576

  Margin of Error = 2.576 × (12/√50) = 2.576 × 1.7 = 4.38

  CI = 78 ± 4.38 → (73.62, 82.38)

- **Interpretation:** We are 99% confident that true mean score is between 73.62 and 82.38.

Why Use Confidence Intervals?

- A single **sample mean** (e.g., 78) may not fully represent the population.
- A **confidence interval** gives us a **range of possible values**, reducing uncertainty.

Effect of Sample Size on Confidence Intervals

Larger samples → Narrower intervals (more precise estimates).
Smaller samples → Wider intervals (more uncertainty).

- Example: A survey of 1,000 people gives more accurate results than 50 people.

Applications of Confidence Intervals:

- **Elections:** Predicting a candidate’s support (e.g., "45% ± 3%")
- **Healthcare:** Estimating the effect of a new drug
- **Economics:** Estimating the average household income in a city

Class Activity: Compute Confidence Intervals

• Given: Screen time data from 40 students, mean = 220 min, standard deviation = 50 min.

Tasks:

- Compute 95% and 99% confidence intervals.
- Compare the widths of both intervals.
- Discussion: Why does increasing confidence level widen the interval?

Estimating Variance of the Sample Mean

Week 4
Lecture 3
Spring 2025
Rida Maryam

Introduction

- The **sample mean** ($\bar{x}$) is used to estimate the **population mean** (μ).
- However, different samples taken from the same population can have **different means**, causing **variability**.
- This variability is measured by the **variance of the sample mean**, which helps us understand the accuracy of our estimate.

Definition of Variance of the Sample Mean

- Population Variance: $\sigma^2 = \Sigma(Xi - \mu)^2 / N$
  - Measures how spread out the values are in the entire population.
- Sample Variance: $s^2 = \Sigma(Xi - \bar{x})^2 / (n-1)$
- Variance of Sample Mean: $Var(\bar{x}) = \sigma^2 / n$
- Standard Error (SE): $SE = \sigma / \sqrt{n}$
  - The standard deviation of the sample mean (also called the **Standard Error**)

➢ **Larger samples** have **lower variance**, meaning the sample mean is more accurate.

Example 1: Computing Sample Variance

A teacher records the test scores of **5 students**: **78, 85, 92, 88, 80**.
Find the **sample variance** and **standard deviation**.
Solution:
1. Compute the **sample mean:** $\bar{x}$

$\bar{x}$=78+85+92+88+80/5=423/5=84.6

2. Compute each deviation from the mean, square them, and find variance.

(78−84.6)2=(−6.6)2=43.56
(85−84.6)2=(0.4)2=0.16
(92−84.6)2=(7.4)2=54.76
(88−84.6)2=(3.4)2=11.56
(80−84.6)2=(−4.6)2=21.16

3. Compute the **sample variance**:

s2=43.56+0.16+54.76+11.56+21.16/5-1=131.2/4=32.8

4. Compute the **sample standard deviation**:

s= √ 32.8=5.73

Interpretation:

The test scores vary by about **5.73 points** from the mean.

Example 2: Computing Variance of Sample Mean

• A researcher collects **10 samples** from a population where **σ=15**
Find the **variance of the sample mean** and **standard error**.
Solution:

Population standard deviation (σ) = **15**

Sample size (n) = **10**

1. Compute the **variance of the sample mean**:
   $Var(\bar{x}) = \sigma^2 / n = 225 / 10 = 22.5$

2. Compute the **standard error**:
   Standard Error: $SE = \sigma / \sqrt{n} = 15 / 3.16 = 4.75$

Interpretation:
The standard error of **4.75** means that the sample mean will typically vary by **4.75 units** around the true population mean.

Effect of Sample Size on Variance

• Larger sample size → Smaller variance → More precise estimates.
• Smaller sample size → Higher variance → Less precise estimates.

• Example: A survey of 5 people vs. 500 people.

Importance of Estimating Variance of the Sample Mean

- Helps us **understand uncertainty** in sample estimates.
- Used in **hypothesis testing and confidence intervals.**
- Informs us how **precise our sample mean is** when estimating the population mean.

Applications of Sample Mean Variance

- Medical Trials: Measuring the effect of a new drug.
- Business Analytics: Predicting customer income.
- Surveys: Estimating public opinion polls.
- Research: Estimating population parameters.

Class Activity: Compute Variance

- A dataset contains the **exam scores of 6 students**:
  75, 80, 85, 70, 90, 95

1. Compute the **sample variance** and **standard deviation**.
2. Compute the **variance of the sample mean** if the population standard deviation is **10** and sample size is **6**.
3. Discuss: What happens if we increase the sample size to 20?