MATH 108: Elementary Probability and Statistics

Ramapo College of New Jersey

Instructor: Dr. Atul Anurag

Semester: Fall 2025 Date: September 8, 2025

Measures of Central Tendency

Objectives

- Determine the arithmetic mean of a variable from raw data.
- Determine the median of a variable from raw data.
- Explain what it means for a statistic to be resistant.
- Determine the mode of a variable from raw data.

1 Introduction

A **measure of central tendency** numerically describes the average or typical value of a dataset. In everyday language, this is often referred to as the *average*.

Common Examples

- The average miles per gallon (MPG) of a 2016 Chevrolet Corvette Z06: 29 MPG.
- The average commute time to work (U.S., 2013): 25.4 minutes.
- Average household income (U.S., 2014): $53,657.
- Average height of American men: 5 feet 10 inches.

Note: The term average is often used informally to refer to the *mean*, but sometimes is used to refer to the *median* or *mode* depending on context.

2 Arithmetic Mean (The Mean)

Definition

The **arithmetic mean** of a variable is the sum of all the data values divided by the number of observations.

- **Population Mean** (μ): parameter.
- **Sample Mean** ($\bar{x}$): statistic.

Formulas

$$\text{Population Mean:} \quad \mu = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{\sum x_i}{N}$$

$$\text{Sample Mean:} \quad \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum x_i}{n}$$

Example 1: Population and Sample Mean

Given test scores of 10 students: `82, 77, 90, 71, 62, 68, 74, 84, 94, 88`

- **(a) Population Mean:**

$$\mu = \frac{82+77+90+71+62+68+74+84+94+88}{10} = \frac{790}{10} = 79$$

- **(b) Random Sample (n = 4):** `62, 88, 77, 68`
- **(c) Sample Mean:**

$$\bar{x} = \frac{62+88+77+68}{4} = \frac{295}{4} = 73.8$$

Example 2: Population Mean

The heights (in cm) of all 5 students in a class are: 160, 165, 170, 175, 180.
Find the **population mean**.

$$\mu = \frac{160+165+170+175+180}{5} = \frac{850}{5} = 170$$

Example 3: Sample Mean

A random sample of 4 exam scores from a class are: 72, 85, 90, 88.
Find the **sample mean**.

$$\bar{x} = \frac{72+85+90+88}{4} = \frac{335}{4} = 83.75$$

Population and Sample Mean Comparisons

Let the population be: $\{4, 6, 8, 10, 12\}$

$$\text{Population Mean } \mu = \frac{4+6+8+10+12}{5} = 8$$

Case 1: Sample Mean = Population Mean

Sample: $\{6, 8, 10\}$

$$\bar{x} = \frac{6+8+10}{3} = 8 \Rightarrow \bar{x} = \mu$$

Case 2: Sample Mean < Population Mean

Sample: $\{4, 6, 8\}$

$$\bar{x} = \frac{4+6+8}{3} = 6 \Rightarrow \bar{x} < \mu$$

Case 3: Sample Mean > Population Mean

Sample: $\{8, 10, 12\}$

$$\bar{x} = \frac{8+10+12}{3} = 10 \Rightarrow \bar{x} > \mu$$

3 Median

Definition

The **median** is the value that lies in the middle of the dataset when arranged in ascending order. It is denoted by M.

Steps to Compute the Median

1. Arrange the data in ascending order.
2. Count the number of observations, n.
3. Determine the middle position:
   - If n is **odd**: Median = value at position $\frac{n+1}{2}$
   - If n is **even**: Median = average of values at positions $\frac{n}{2}$ and $\frac{n}{2}+1$

Example 1: Median of an Unordered Dataset (Even n)

Find the median of: 88, 72, 91, 85, 77, 90

Steps to Compute the Median

1. Arrange the data in ascending order: 72, 77, 85, 88, 90, 91
2. Count the number of observations, $n = 6$
3. Since n is even:
$$\text{Median} = \frac{\text{3rd value} + \text{4th value}}{2} = \frac{85 + 88}{2} = 86.5$$

Example 2: Median with Repeated Values (Odd n)

Find the median of: 55, 60, 60, 65, 70, 75, 80

Steps to Compute the Median

1. Data is already ordered.
2. Count the number of observations, $n = 7$
3. Since n is odd:
$$\text{Median} = \text{value at position } \frac{7+1}{2} = 4 \Rightarrow \text{Median} = 65$$

Example 3: Median with an Outlier (Even n)

Find the median of: 10, 12, 13, 14, 15, 100

Steps to Compute the Median

1. Data is already in order.
2. Count the number of observations, $n = 6$
3. Since n is even:
$$\text{Median} = \frac{\text{3rd value} + \text{4th value}}{2} = \frac{13 + 14}{2} = 13.5$$

4 Resistant Statistics

A **resistant statistic** is not significantly affected by extreme values (outliers).

- **Median** is resistant.
- **Mean** is **not** resistant—it can be skewed by outliers.

Example

Consider the data: `10, 11, 12, 13, 100`

- Mean $= (10 + 11 + 12 + 13 + 100)/5 = 29.2$
- Median $= 12$

Conclusion: The outlier (100) greatly increases the mean but does not affect the median.

5 Mode

Definition

The **mode** is the value that appears most frequently in a dataset.

- A dataset can be:
  - **Unimodal**: One mode
  - **Bimodal**: Two modes
  - **Multimodal**: More than two modes
  - **No mode**: All values occur with the same frequency
- The mode is the only measure of central tendency that can be used with **qualitative** data.

Example 1

Dataset: `62, 68, 71, 74, 77, 82, 84, 88, 90, 94`
Each value occurs only once: **No mode**

Categorical Example

Responses to favorite color: `Blue, Blue, Red, Green, Blue, Red`

- Mode = **Blue** (appears 3 times)

Example 2: Finding the Mode

The following values represent the number of cups of coffee consumed by 8 employees in a day: 2, 3, 3, 4, 2, 5, 3, 2. Find the **mode**.
Frequency of each value:

- 2 appears 3 times
- 3 appears 3 times

- 4 and 5 appear once

There are two modes:

$$\text{Mode} = 2 \text{ and } 3 \quad \text{(Bimodal)}$$

6 Shapes of Distributions

Definition

The **shape of a distribution** describes the overall pattern of the data values when graphed. It gives insight into the nature of the data, including symmetry, skewness, and uniformity.

Common Shapes

- **Symmetric (Normal)**: Data is evenly spread around the center, often forming a bell-shaped curve. The mean and median are approximately equal.
- **Skewed Right (Positively Skewed)**: The data has a long tail on the right side, indicating some unusually high values. The mean is greater than the median.
- **Skewed Left (Negatively Skewed)**: The data has a long tail on the left side, indicating some unusually low values. The mean is less than the median.
- **Uniform**: Data values are spread evenly across the range, with roughly equal frequencies.

Example: Symmetric (Normal) Distribution

Consider the dataset: 65, 70, 75, 80, 85

- Mean $= \frac{65+70+75+80+85}{5} = 75$
- Median $= 75$
- Since the mean and median are equal and data is evenly distributed, the shape is **symmetrical**.

Example: Skewed Right (Positively Skewed)

Consider the dataset: 10, 12, 13, 15, 50

- Mean $= \frac{10+12+13+15+50}{5} = 20$
- Median $= 13$
- The mean is greater than the median due to the high value (50), creating a **right-skewed** distribution.

Example: Skewed Left (Negatively Skewed)

Consider the dataset: 5, 40, 42, 43, 44

- Mean $= \frac{5+40+42+43+44}{5} = 34.8$
- Median $= 42$
- The mean is less than the median due to the low value (5), creating a **left-skewed** distribution.

Summary Tables

Shape	Key Features
Symmetric	Mean $\approx$ Median, Bell-shaped
Skewed Right	Mean $>$ Median, Long right tail
Skewed Left	Mean $<$ Median, Long left tail
Uniform	All bars roughly equal height

Measure	Symbol / Representation	Resistant?
Mean	μ (population), $\bar{x}$ (sample)	No
Median	M	Yes
Mode	—	Yes

End of Lecture #3