Measure of Central Tendency

Measures of Central Tendency

- A measure of central tendency is a summary measure that attempts to describe a whole set of data with **a single value** that represents the **middle or centre** of its distribution.
- There are three main measures of central tendency: the **mean, median, and mode**. Each of these measures describes a different indication of the typical or central value in the distribution.

Measures of Central Tendency

Mean

- The arithmetic mean is the most widely used average.
- For any set of data on the variable x, the mean is denoted by $\bar{x}$ and is obtained by the sum of observations by their number
- $\bar{x} = \frac{1}{n}\sum_{1}^{n} x_i$

Example

Calculate the mean for the following dataset:

$$1\ 2\ 2\ 4\ 5\ 10$$

Solution:

$$\mu = \frac{1+2+2+4+5+10}{6} = \frac{24}{6} = 4$$

Now change just one data point in the dataset

$$1\ 2\ 2\ 4\ 5\ 70$$

Solution:

$$\mu = \frac{1+2+2+4+5+70}{6} = \frac{84}{6} = 14$$

Properties of Mean

- It is greatly affected by extreme values
- Sum of differences about the mean is zero

$$\sum_{i=1}^{n}(x_i - \mu) = 0$$

- Mean is unique for a given dataset

Measures of Central Tendency

Median

- The middlemost value of the set when the elements are arranged in either ascending or descending order is called the median of a set of observations on a variable x.
- $\tilde{x} = \frac{x_{(m)} + x_{(m+1)}}{2}$

Measures of Central Tendency

Median

- It is the “middle” of the data
- The median is the number such that exactly half of the dataset is less than or equal to it and exactly half is greater than or equal to it
- To get the median, we must first rearrange the data into an ordered array
- If n is odd, the median is the middle observation of the ordered array.
- If n is even, it is midway between the two central observations

Measures of Central Tendency

Median

- Robert hit **11** balls at Grimsby driving range. The recorded distances of his drives, measured in yards, are given below. Find the median distance for his drives.

  85, 125, 130, 65, 100, 70, 75, 50, 140, 95, 70

- Ordered data

  50, 65, 70, 70, 75, 85, 95, 100, 125, 130, 140

- Single middle value

- Median drive = 85 yards

- *In golf stroke mechanics, a drive, also known as a tee shot, is a long-distance shot played from the tee box, intended to move the ball a great distance down the fairway towards the green.*
- *A tee is a stand used to support a stationary ball so that the player can strike it*

Measures of Central Tendency

Median

- Robert hit **12** balls at Grimsby driving range. The recorded distances of his drives, measured in yards, are given below. Find the median distance for his drives.

  85, 125, 130, 65, 100, 70, 75, 50, 140, 95, 70, 135

- Ordered data

  50, 65, 70, 70, 75, 85, 95, 100, 125, 130, 135, 140

- Two middle values so take the mean

- Median drive = 90 yards

- *In golf stroke mechanics, a drive, also known as a tee shot, is a long-distance shot played from the tee box, intended to move the ball a great distance down the fairway towards the green.*
- *A tee is a stand used to support a stationary ball so that the player can strike it*

Properties of Median

- Median is unique for a dataset
- Median is not affected by extreme values
- Any observation selected at random is just as likely to be greater than the median as less than the median

Measures of Central Tendency

Mode

- The mode is the most commonly occurring value in a distribution.
- Mode is the value that occurs most frequently
- Might not be unique. More than one mode per dataset
- Example:

$$1\ 1\ 1\ 2\ 3\ 4\ 5$$

Mode = 1

$$5\ 5\ 5\ 6\ 8\ 10\ 10\ 10$$

Mode = 5, 10

- This is a "$bimodal$" dataset

Comparisons of Measures of Central Tendency

Comparison

- **The “Hotshot” Sales Executive**
  - Kurt works as a sales manager at vsellhomes.com. In the monthly sales review, Kurt reports that he will achieve his quarterly target of $1M.
  - Kurt claims his average deal size is $100,000 and he has 10 deals in his pipeline.
  - At the end of quarter, even after closing 8 deals Kurt fails to meet his target number and falls short by more than $500,000.

Comparison

- **The Reality of the “Hotshot” Sales Executive**
  - Average deal size in pipeline = $100,000
  - Deal #10 is of significantly higher value than all the
  - other deals and impacts the average calculation
  - Median = $55,000 more realistic measure

Deal #	Deal Value	Deal Status
1	70,000	Open
2	50,000	Closed
3	55,000	Closed
4	60,000	Closed
5	55,000	Closed
6	50,000	Closed
7	50,000	Closed
8	60,000	Closed
9	50,000	Closed
10	5,00,000	Open

Comparison

- A report says that “**the median credit card debt of American households is zero**.’’ We know that many households have large amounts of credit card debt. In fact, **the mean household credit card debt is close to $8000**.

- Is the report correct?

Comparison

- The mean and the median are rigidly defined. The position of the mode is somewhat similar to that of the median.
- All the measures are easy to interpret and not too difficult to compute.
- Only the mean directly depends on all the observations. A change in any one of the observations influences the value of the mean. The median and mode are not so sensitive.
- The mean is, generally, the best measure of central tendency. In case of extreme values, median is better measure of central tendency.

Scale of Measurement and Measure of Central Tendency

	Indicates Difference	Indicates Direction of Difference	Indicates Amount of Difference	Absolute Zero
Nominal	✓			
Ordinal	✓	✓		
Interval	✓	✓	✓	
Ratio	✓	✓	✓	✓

Nominal	Mode
Ordinal	Mode and Median
Interval	Mode, Median and Mean
Ratio	Mode, Median and Mean