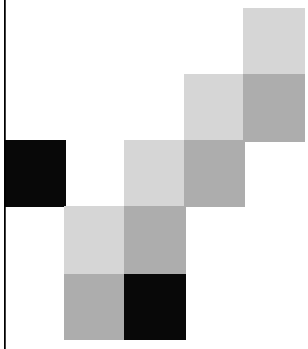


Measures of Center

MATH 2300

Section 3.1

Location and Variability

- To get a ***useful*** numerical description of a distribution, we need to have both some measure of **center (**or **location)** and some measure of **variability**.

Measures of Center

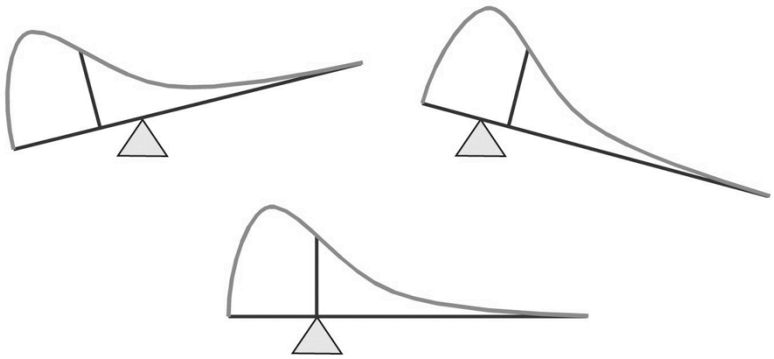
- Provide some notion of what value the distribution is grouped around.
- The most common measures are the **mean**, **median**, and **mode**.

The Sample Mean

- Also called the arithmetic **average**.
- Represents the ***balance point*** for a sample of *quantitative* data.

$$\bar{x} = \sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

The Mean as Balance Point



Housing Price Example

All Values Given in Thousands of Dollars

{House Prices} = {144; 98; 204; 177; 155; 316; 100}

$$\bar{x} = \frac{144 + 98 + 204 + 177 + 155 + 316 + 100}{7}$$

$$= \frac{1{,}194}{7} = 170.6$$

The Sample Median

- The median Q_2 is the **midpoint** of a distribution.
- It is the number such that half the observations are smaller and the other half are larger.

Calculating the Median

1. Arrange the observations in order from smallest to largest.
2. If *n* is odd, Q_2 is the center observation.
3. If *n* is even, Q_2 is the mean of the two centered observations.

Housing Price Example

{house prices} = {$144; 98; 204; 177; 155; 316; 100}

NOTE: *n* is odd

Ordered:

$98; 100; 144; 155; 177; 204; 316

Middle Value

$$Q_2 = 155$$

Another Housing Price Example

{house prices} = {144; 98; 204; 177; 155; 316; 100; 177; 177; 170}

NOTE: *n* is even

Ordered:

$98; 100; 144; 155; 170; 177; 177; 177; 204; 316

Middle Values

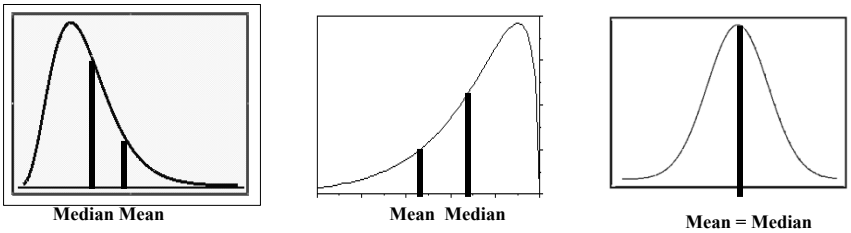
$$Q_2 = (170 + 177)/2$$

$$= 173.5$$

Mean vs. Median

- The mean is GREATLY affected by outliers.
- The median is NOT affected by outliers.
- If the mean and median are (almost) equal, then the distribution is (approximately) symmetric.
- If mean < median, distribution is left skewed.
- If mean > median, distribution is right skewed.

Skewness



Right-Skewed **Left-Skewed** **Symmetric**

Trimmed Means

- Remove a percentage of the highest and lowest values and calculate the mean from what remains.
- More **resistant** to outliers than the mean.
- Less **resistant** to outliers than the median.

Example

{house prices} = {$144; 98; 204; 177; 155; 316; 100}

Mean = 170.6 Median = 155

Ordered:

~~$98~~; 100; 144; 155; 177; 204; ~~316~~

14.3% Trimmed Mean = 156

The Mode

- The observation with greatest frequency in the data set
  - **If no value occurs more than once, then the data set has *no* mode.**
- Not useful for really small data sets
- Sometimes, it only makes sense to describe as a class, rather than an individual value