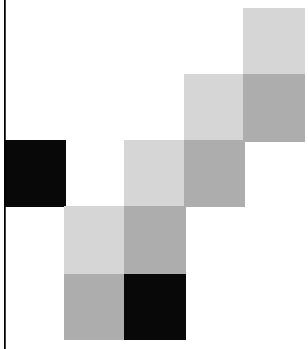


Measures of Location

MATH 3342

Section 1.3

Location and Variability

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

Measures of Center

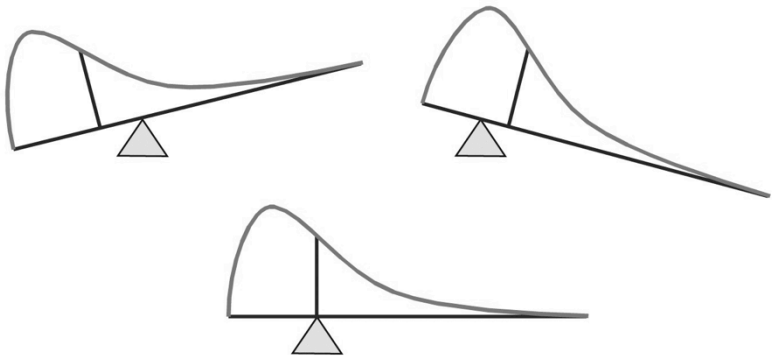
- Provide some notion of what value the distribution is grouped around.
- The most common measures are the **mean** and **median**.

The Sample Mean

- Also called the arithmetic **average**.
- Represents the _____ 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 $\tilde{x}$ 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, $\tilde{x}$ is the center observation.
3. If n is even, $\tilde{x}$ 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

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

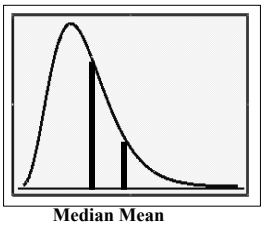
Population Analogues

- The mean of a population is denoted by μ
  - **Discussed in more depth in Ch. 3 and 4**
- The median of a population is denoted by $\tilde{\mu}$

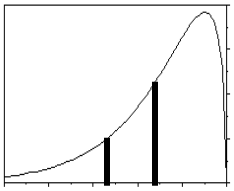
Mean vs. Median

- The _____ is GREATLY affected by outliers.
- The _____ 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.

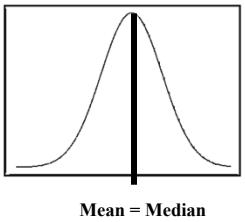
Skewness



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

Quartiles

- Quartiles divide the data into four equal-sized groups.
- 25% of the data are less than Q_1, the first quartile.
- The second quartile is the median $\tilde{x}$.
- 75% of the data are less than Q_3, the third quartile.

Calculating Quartiles

1. Arrange data in ascending order.
2. Locate the median.
3. Q_1 is the median of all data to the left of the overall median.
4. Q_3 is the median of all data to the right of the overall median.

The Five-Number Summary

Minimum, Q_1, $\tilde{x}$, Q_3, Maximum

- A quick summary of both center **and** variability.

Example: Quartiles

Original:

6 -5.2 -1 3.5 3.2 -1.7 -1.5 1 4.9 1.4 7 3.7 -12

Sorted:

-12 -5.2 -1.7 -1.5 -1 1 1.4 3.2 3.5 3.7 4.9 6 7

To the left of the median:

-12 -5.2 -1.7 -1.5 -1 1

$\tilde{x}$

To the right of the median:

3.2 3.5 3.7 4.9 6 7

Five-Number Summary:

Location for Qualitative Data

- The sample proportion is the ratio of successes to number of observations:

$$\hat{p} = \frac{\#\,successes}{n} = \frac{x}{n}$$

- The population proportion is p.

Example

- A department wants to know what proportion of their 400 returning students have laptops.
- They poll 40 of these students.
- 28 of them said yes.

- What is the sample proportion of returning students that have laptops?