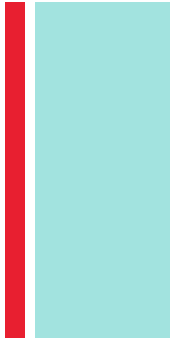


+ FUN Friday! Sept 2, 2022

- **Warm-Up: Letter to Future ME**
- **Review measures of Center & Spread**
- **SU-DO-KU? Game of Skunk?**
- **HW Time, Video time**
- **TEST review DUE Next week!**
- **Questions?**



Measures of Center and spread



- What are common measures of center for a numerical distribution of data?

mean & median

- What common measures of spread for a numerical distribution of data?

range, interquartile range (IQR),
& *standard deviation*

■ Measuring Center: The Mean

- The most common measure of center is the ordinary arithmetic average, or **mean**.

Definition:

To find the **mean** \bar{x} (pronounced “x-bar”) of a set of observations, add their values and divide by the number of observations. If the n observations are $x_1, x_2, x_3, \dots, x_n$, their mean is:

$$\bar{x} = \frac{\text{sum of observations}}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

In mathematics, the capital Greek letter Σ is short for “add them all up.” Therefore, the formula for the mean can be written in more compact notation:

$$\bar{x} = \frac{\sum x_i}{n}$$

■ Measuring Center: The Median

- Another common measure of center is the **median**. In section 1.2, we learned that the median describes the midpoint of a distribution.

Definition:

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

To find the median of a distribution:

- 1) Arrange all observations from smallest to largest.
- 2) If the number of observations n is odd, the median M is the center observation in the ordered list.
- 3) If the number of observations **n is even**, the median M is the average of the two center observations in the ordered list.

■ Measuring Center

- Use the data below to calculate the mean and median of the commuting times (in minutes) of 20 randomly selected New York workers.

Example, page ??

10	30	5	25	40	20	10	15	30	20	15	20	85	15	65	15	60	60	40	45
----	----	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

$$\bar{x} = \frac{10 + 30 + 5 + 25 + \dots + 40 + 45}{20} = 31.25 \text{ minutes}$$

0	5
1	005555
2	0005
3	00
4	005
5	
6	005
7	
8	5

Key: 4|5
represents a
New York
worker who
reported a 45-
minute travel
time to work.

$$M = \frac{20 + 25}{2} = 22.5 \text{ minutes}$$

Comparing the Mean and the Median

- The mean and median measure center in different ways, and both are useful.
 - *Don't confuse the "average" value of a variable (the mean) with its "typical" value, which we might describe by the median.*

Comparing the Mean and the Median

The mean and median of a roughly symmetric distribution are close together.

If the distribution is exactly symmetric, the mean and median are exactly the same.

In a skewed distribution, the mean is usually farther out in the long tail than is the median.

+ Measures of spread

Range: the spread of all the data, calculated as the difference between the largest and smallest observations in the data.

Standard deviation: average or “typical” deviation from the mean for a set of data. Calculated by finding the average of the squared deviations from the mean.

Interquartile range (*IQR*): the spread of the middle 50% of the data, calculated by difference in $Q_3 - Q_1 = IQR$

■ Measuring Spread: The Interquartile Range (*IQR*)

- A measure of center alone can be misleading.
- A useful numerical description of a distribution requires both a measure of center and a measure of spread.

How to Calculate the Quartiles and the Interquartile Range

To calculate the **quartiles**:

- 1) Arrange the observations in increasing order and locate the median M .
- 2) The **first quartile** Q_1 is the median of the observations located to the left of the median in the ordered list.
- 3) The **third quartile** Q_3 is the median of the observations located to the right of the median in the ordered list.

The **interquartile range** (*IQR*) is defined as:

$$IQR = Q_3 - Q_1$$