3.1: Measures of Center

Now, we will begin studying some numerical measures that describe data sets. There are two basic types:

- Measures of central tendency (this section)
- Measures of dispersion (next section)

Σ = Greek letter capital Sigma. stands for sum

Summation Notation:

Summation notation is a compact way to write "add up n numbers" of "do something to n numbers first, and then add them up." The numbers are represented as $x_1, x_2, \ldots, x_n$"

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

Example 1: Consider the numbers 8, 2, 6, 10, 4, 9. Find $\sum_{i=1}^{6} x_i$ and $\sum_{i=1}^{6} x_i^2$.

$$\sum_{i=1}^{6} x_i = 8+2+6+10+4+9 = \boxed{39}$$

$$\sum_{i=1}^{6} x_i^2 = 8^2+2^2+6^2+10^2+4^2+9^2 = 64+4+36+100+16+81$$

$$= \boxed{301}$$

<u>The Mean: Ungrouped Data</u>:

The *mean* of a set of quantitative data is equal to the sum of all the measurements in the data set divided by the total number of measurements in the set.

If $x_1, x_2, \ldots, x_n$ is a set of n measurements, then the *mean*, or *average*, is given by

$$[\text{mean}] = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n} \quad \text{where}$$

$\bar{x} = [\text{mean}]$ if data set is a sample $\qquad \bar{x}$ (x-bar) = sample mean

$\mu = [\text{mean}]$ if data set is the population $\qquad \mu$ = population mean

Greek letter mu

The median:

Sometimes the mean can be misleading for a data set. Suppose that a math class had 7 students with test scores (out of a possible 100) of 88, 99, 7, 78, 89, 94, and 75.

Population mean: $\mu = \frac{88+99+7+78+89+94+75}{7} = 75.71$

Note: All but one student scored at least 75.
The low score of 7 pulled the mean down.

The *median* is unaffected by extreme values (outliers). Essentially it is the "middle" of the data set.

To find the median, you'll need to sort the data in numerical order.

> The Median (Ungrouped Data):
>
> - If the number of measurements is odd, the median is the middle measurement when the measurements are arranged in descending or ascending order.
> - If the number of measurements is even, the median is the mean of the two middle measurements when the measurements are arranged in descending or ascending order.

Example 2: Find the median of the test scores 88, 99, 7, 78, 89, 94, and 75.

~~7~~ ~~75~~ ~~78~~ 88 ~~89~~ ~~94~~ ~~99~~

median = 88

Example 3: Find the median of the test scores 88, 85, 99, 7, 78, 89, 94, and 75.

~~7~~ ~~75~~ ~~78~~ $\underbrace{85 \quad 88}$ ~~89~~ ~~94~~ ~~99~~

Median $= \frac{85+88}{2} = \boxed{86.5}$

Example 4: Provide some everyday examples in which the median is more useful than the mean.

Home prices

Income

The mode:

<u>The Mode</u>:

The *mode* is the most frequently occurring value in a data set, provided it occurs at least twice. There may be a unique mode, several modes, or no mode.

A data set with two modes is called *bimodal*.

<u>Example 5:</u> Find the median and mode for the following data sets.

a. {4, 5, 5, 5, 5, 6, 7, 8, 12}

mode = 5

b. {1, 2, 3, 3, 3, 5, 7, 7, 7, 23}

modes: 3, 7

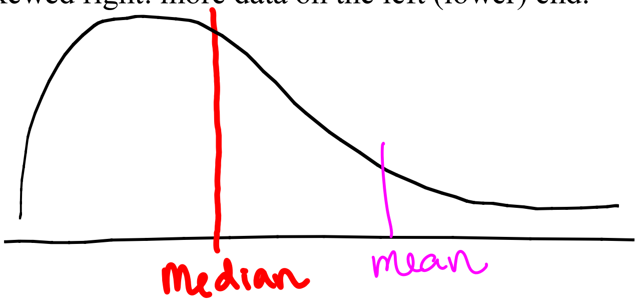
c. {1, 3, 5, 6, 7, 9, 11, 15}

no mode

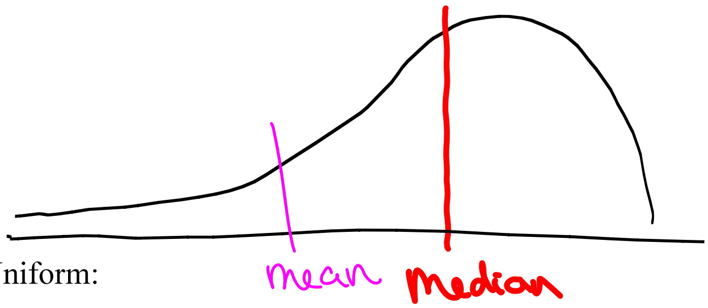
<u>Example 6:</u>

Mean, median, and mode for distributions of different shapes:

Skewed right: more data on the left (lower) end.

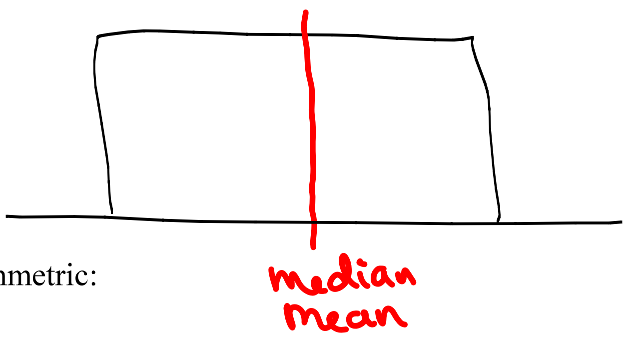


Skewed left: more data on the right (upper) end.



not symmetric

Uniform:



(this uniform distribution is an example of a symmetric distribution)

Symmetric: