Section 3.1 - Measures of Center

The **arithmetic mean** (or simply mean) of a data set is the measure of center found by adding the data values and then dividing the sum by the total number of values.

- The mean is a fairly reliable measure of center---when samples are selected from the same population, sample means tend to be more consistent than other measures of center.
- The mean takes every data value into account.
- The mean is particularly sensitive to extreme data values---it *is not a resistant* measure of center. (A statistic is **resistant** if the presence of extreme values does not have a great effect on its value.)

The **median** is the numerically middle data value if there are an odd number of values or the mean of the two middle values if there are an even number.

- The data must be arranged numerically to determine the median.
- The median *is a resistant* measure of center---it is not dramatically affected by extreme values.

The **mode** is the data value that occurs with the greatest frequency (assuming there is one).

- If two data values occur with same greatest frequency, there are two modes and the data set is **bimodal**.
- If there are more than two modes, the data set is **multimodal**.
- If there are no repeated data values, there is **no mode**.

The **midrange** is the measure of center found by computing the value midway between the minimum data value and the maximum: add the minimum and maximum and then divide by two.

- The midrange is rarely used because is a very sensitive to extreme values.

Warning: The word *average* is often used for the mean, but it can be used to refer to any of the measures of center.

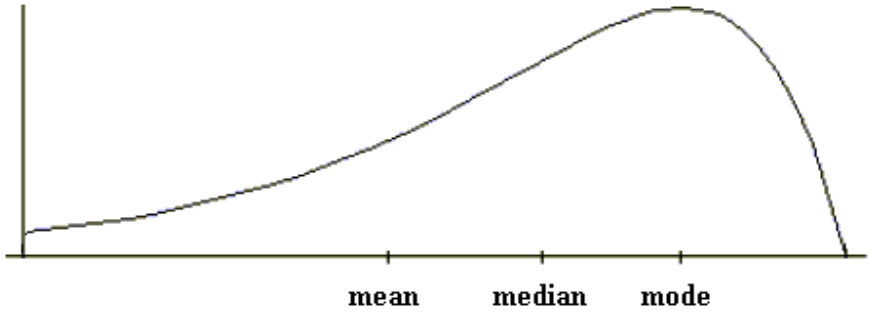
If certain data values "carry more weight" than others, it may be appropriate to compute a **weighted mean**. A weighted mean is computed by multiplying each data value by its associated "weight", adding, and then dividing by the total "weight".

Some examples of weighted means are grade point averages, expected values in probability, and means obtained from frequency distributions.

A comparison of the mean, median, and mode can provide information about how a distribution is **skewed**.

While the mean and median do not always characterize the shape of the distribution, the following are generally true...

Skewed left (negatively skewed)---mode greater than mean and median



Skewed right (positively skewed)---mode less than mean and median