

Math 146 3.1 - Measures of Center

Measures of Center are used to describe the typical or “average” value of sample data. Depending on the distribution of the data, different measures of center are used to represent the typical or middle value.

Mean

The mean (or arithmetic mean) is the most common measure of center, and is the one most people refer to as the *average*.

Sample Mean: $\bar{x} = \frac{\sum X}{n}$ Population Mean: $\mu = \frac{\sum X}{N}$ (μ is pronounced mu)

Mean of Grouped Data: $\bar{x} = \frac{\sum(xf)}{\sum(f)}$ where x is the class midpoint and f is the frequency.

Note: sigma, \sum , means the sum of ...

Pros:

- Sample means from the same population vary less than other measures of center.
- Uses every data value in the sample.

Cons:

- One extreme outlier can have a substantial effect on the value of the mean.
-

Median

The median is the “middle value”, or the value that divides the data set in half.

Odd n : 12 17 20 21 **25** 31 37 39 42

The median is $\tilde{x} = 25$. That value has four data points to the left, and four points to the right.

Even n : 14 17 20 \wedge 26 30 32

The value mid-way between 20 and 26 splits the data set with three values on each side: $\tilde{x} = \frac{20+26}{2} = 23$. The median can be represented by \tilde{x} , M, or Med.

Pros: Resistant to extreme values. Replacing 32 with 100 in the above set does not change the median.

Cons: Doesn't use every data value. Data set must be sorted first.

Mode

The mode is the value that occurs with the greatest frequency. If two values occur with the same greatest frequency we say the data is **bi-modal**; more than two can be called **multi-modal**.

Pros: Can be used with both quantitative and qualitative data.

Cons: A data set may not have a mode, or multiple modes.

Midrange

The midrange is the value midway between the minimum and maximum value: $MR = \frac{\min + \max}{2}$.

Pros: Easy to calculate.

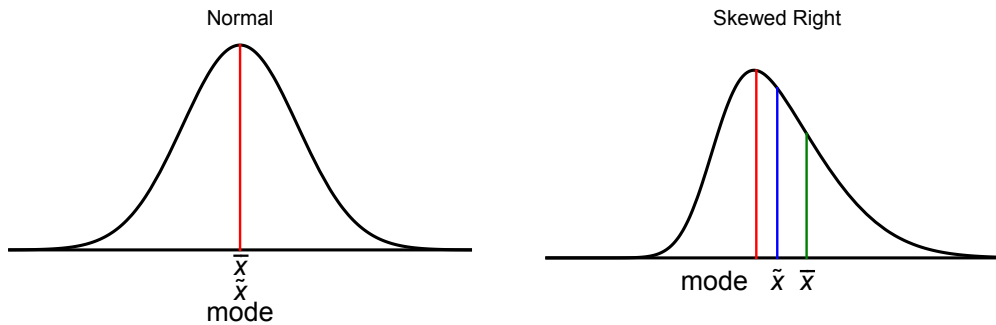
Cons: Only uses the maximum and minimum value and so is very depended on those values.

For the odd n data set, the MR is 27 which is close to the median of 25. Both are probably okay to use to describe a “typical” value.

If we replace the 32 with 100 in the even n data set, {14, 17, 20, 26, 30, 100} the MR is 57, which is certainly NOT representative of a “typical” value.

Location of Mean, Median, Mode

If the data are symmetric and bell shaped, the mean, median and mode have the same value. For data that is skewed right, the mean is most affected by the extreme right values, then the median.



Example 1 Find the Mean, median, mode and midrange of the Exam 1 scores data set. Knowing the shape of the distribution, which measure of center is most representative of a typical score?

42 42 38 42 46 34 40 23 41 38
 39 46 50 49 45 46 36 52 34 43
 44 26 41 45 37

Example 2 Instead of having the raw data above, suppose we only had the summary data:

Score	f
20 – 24	1
25 – 29	1
30 – 34	2
35 – 39	5
40 – 44	8
45 – 49	6
50 – 54	2

Find the mean, median, and mode using the summary data and compare the values to example 1.