

§2.3 Numerical Measures of Central Tendency

There are several ways of measuring the "middle" of a data set or the clustering of data.

Suppose we have n values in a data set:

$\{x_1, x_2, x_3, \dots, x_n\}$, then

Definition the mean value of a data set is given by the sum of the numeric values in the data divided by the number of data values; that is

$$\bar{x} = \frac{\sum_{k=1}^n x_k}{n}$$

Definition the median value of a data set is the value, M , for which half of the data values are less than M and half of the data values are greater than M .

Note that if the data set has an odd number of values, the M is the middle value. If there are an even number of values, then M is the mean of the two middle values.

Definition the mode of a data set is the value or values that has the greatest frequency.

Example 10 households are surveyed for number of children (under age 18)^{that} reside in the household.

The results are the data set $\{2, 3, 0, 2, 1, 0, 3, 0, 1, 6\}$. Create a frequency chart, calculate the mean, median and mode for this data.

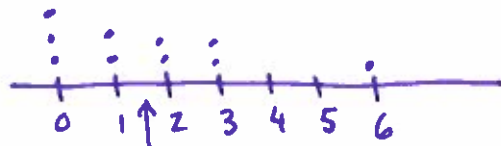
Number of Children	
Class	frequency
0	3
1	2
2	2
3	2
4	0
5	0
6	1
total:	10

} 5 households
} 5 households

The mean: $\bar{x} = (2+3+0+2+1+0+3+0+1+6)/10$
 $= 1.8 = \frac{2(1)+2(2)+2(3)+6}{10}$

The median: $M = (1+2)/2 = 1.5$

Note:



even number of data points, split between values 1 and 2 children.

The mode is 0. 0 has the highest frequency

Notation: \bar{X} = sample mean, μ = population mean.
 M = sample median, η = population median.

Definition the modal class is the class with the highest frequency.

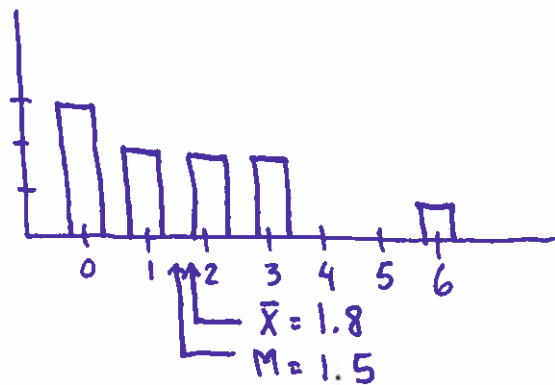
Detecting Skewness in data:

If the data is skewed to the right, then typically the median is less than the mean.

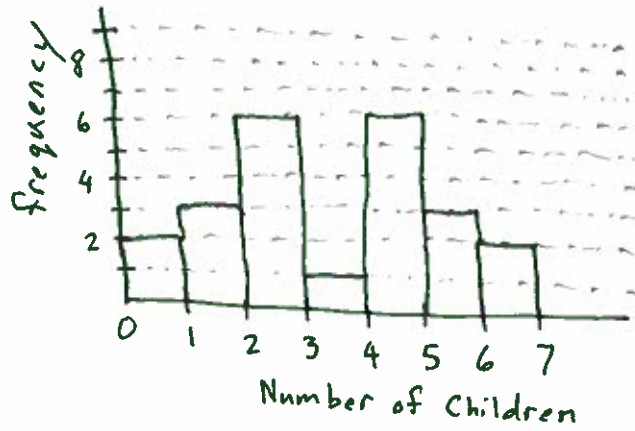
If the data is symmetric, then the mean is equal to the median.

Example The above data on children in the household is skewed to the right.

B/c $M < \bar{X}$



Example We go to a different neighborhood and count children. This survey is summarized as follows



Class	freq.
0	2
1	3
2	6
3	1
4	6
5	3
6	2
total:	23

$\bar{X} = 3$, $M = 3$, mode = $\{2, 4\}$
 The data is symmetric

Skewed Data graphically:

