

Biostatistics

Lecture 4

Prepared by Dr. Ibrahim AL-Jaafari

Descriptive Statistics

Measures of Central Tendency

Mean

*** For Ungrouped Data**

Mean = The sum of all values ÷ Number of values

Note : The mean is used to represent the average when the data are normally distributed or symmetrical shaped.

Descriptive Statistics

Measures of Central Tendency

Mean

*** For Ungrouped Data**

Example / Find the mean of the following data (158, 189, 265, 127, 191)

$$158 + 189 + 265 + 127 + 191 = 930$$

$$930 \div 5 = 186$$

Descriptive Statistics

Measures of Central Tendency

Mean

*** For Grouped Data**

$$\text{Mean} = \Sigma mf/n$$

Σ = Sum

F = Frequency

M = Midpoint of the class

n = Total of frequencies

Descriptive Statistics

Measures of Central Tendency

Mean

*** For Grouped Data**

Example / Find the mean of the following grouped data

Age	Frequency
10 - 12	4
13 - 15	12
16 - 18	20
19 - 21	14
	Total = 50

Descriptive Statistics

Measures of Central Tendency

Mean

*** For Grouped Data**

Solution / $\text{Mean} = \Sigma mf / n = 832 / 50 = 16.64$

Age	Frequency	Class Midpoint	Midpoint $\times$ Frequency
10 - 12	4	11	$11 \times 4 = 44$
13 - 15	12	14	$14 \times 12 = 168$
16 - 18	20	17	$17 \times 20 = 340$
19 - 21	14	20	$20 \times 14 = 280$
	Total = 50		Total = 832

Descriptive Statistics

Measures of Central Tendency

Mean

Advantages and disadvantages of using the mean

The most common measure of central tendency

Easy to use

It is effected by adding or deleting a new values

It is effected by outlier or exterem values

No need to order the data

It takes into account the whole values to be calculated

The sum of deviations from the mean must be equal to zero

Descriptive Statistics

Measures of Central Tendency

Median

Median = The midpoint value of a set of ordered data from lowest to highest

How to calculate median ?

1-) If the number of values is odd.

Order the values from lowest to highest

Find the location of the median $Median\ Location = \frac{N + 1}{2}$

Descriptive Statistics

Measures of Central Tendency

Median

Example:

Find the median for the following values (1, 6, 5, 2, 4)

Order the data (1, 2, 4, 5, 6)

The location of the median = $5 + 1 / 2 = 3$ (The third value of the data set)

Median = 4

Descriptive Statistics

Measures of Central Tendency

Median

2-) If the number of values is even.

Order the values from lowest to highest

Find the location of the median = $\text{Number of values} + 1 / 2$

Descriptive Statistics

Measures of Central Tendency

Median

Example:

Find the median for the following values (7, 1, 6, 5, 2, 4)

Order the data (1, 2, 4, 5, 6, 7)

The location of the median = $6 + 1 / 2 = 3.5$

Median = $4 + 5 / 2 = 4.5$

Descriptive Statistics

Measures of Central Tendency

Median

Advantages and disadvantages of using the median

Easy to use

It is not effected by adding or deleting a new values

It is not affected by outlier or extreme values

The data should be ordered from lowest to the highest

It does not take into account the whole values to be calculated

The sum of deviations from the mean should not be equal to zero

Note : The median is used to represent the average when the data are not normally distributed or symmetrical, for instance the skewed distribution to the right or to the left.

Descriptive Statistics

Measures of Central Tendency

Mode

Mode is the most frequent value in the distribution, there is no way to calculate the mode.

Uni-modal can be found when the data have only one peak, just one mode in the distribution.

Bi-modal can be found when the data have two peaks, this means two modes in the distribution.

Multi-modal can be found when the data have more than two peaks, it means more than two modes.

Descriptive Statistics

Measures of Central Tendency

Mode

Example:

Find the mode for the following data ?

77 , 69 , 74 , 81 , 71 , 68 , 74 , 73

Mode = 74

Descriptive Statistics

Measures of Central Tendency

Mode

Advantages and disadvantages of using the mode

Easy to use

It is not effected by adding or deleting a new values

It is not effected by outlier or exterem values

The data should not be ordered from lowest to the highest

It does not take into account the whole values to be calculated

The sum of deviations from the mean should not be equal to zero

Descriptive Statistics

Measures of Central Tendency

The relationship between mean, median and mode

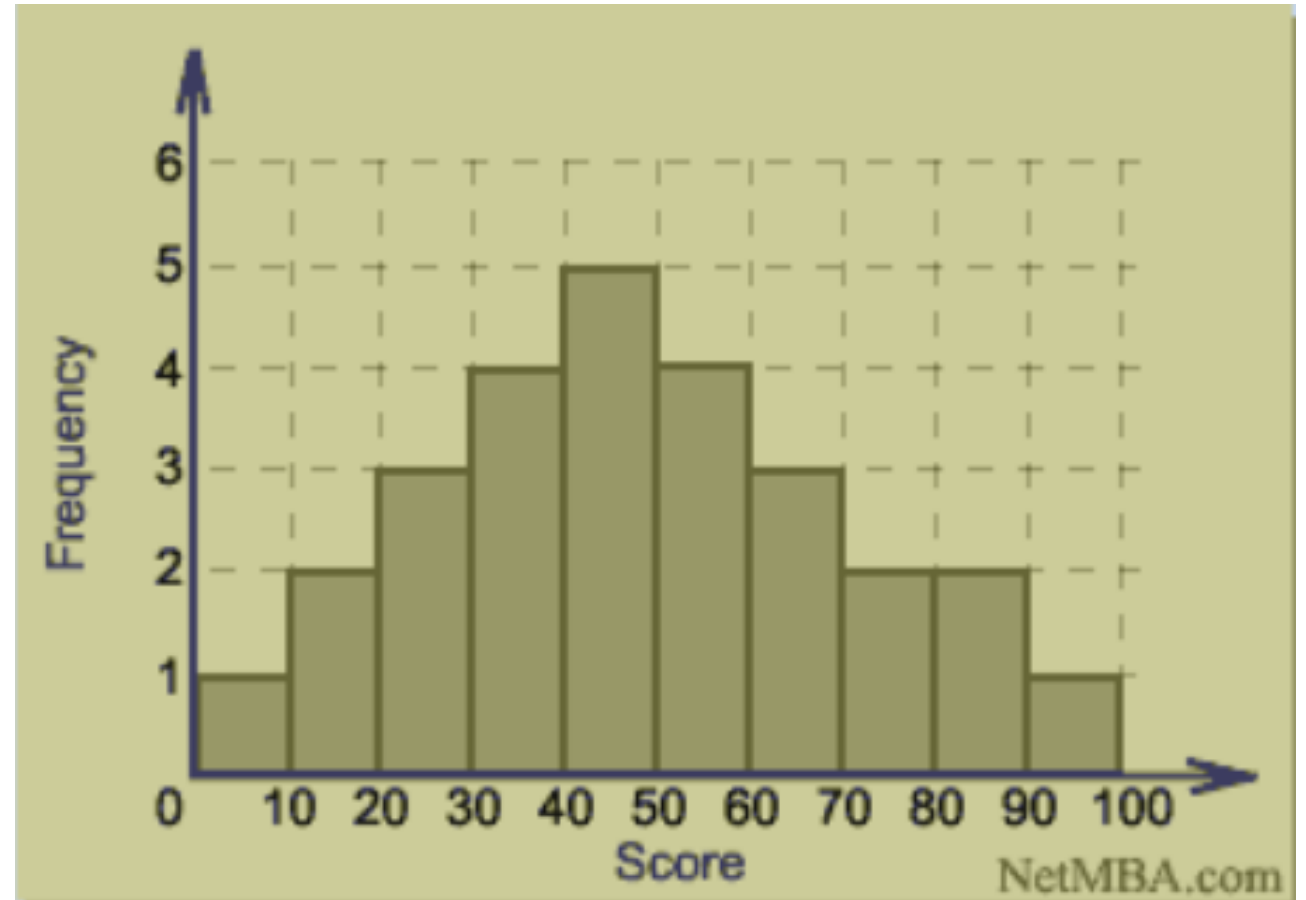
1-) The data is normally distributed (Symmetry)

The values of mean, median and mode are very close to each other in Uni-modal and the shape of the distribution is symmetrical (bell shaped).

Descriptive Statistics Measures of Central Tendency

The relationship between
mean, median and mode

1-) The data is normally
distributed (Symmetry)



Descriptive Statistics

Measures of Central Tendency

The relationship between mean, median and mode

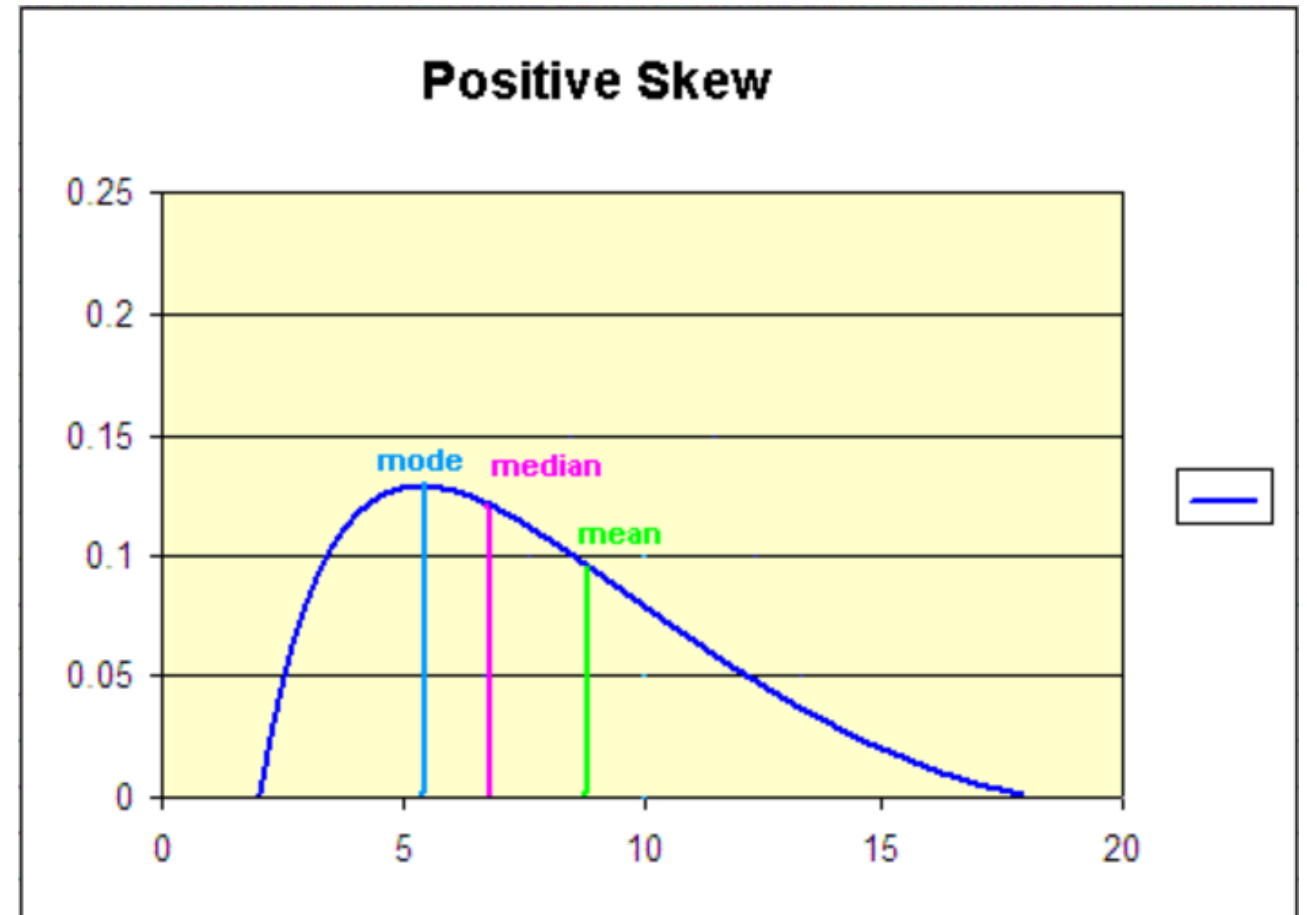
2-) The data is skewed to the right (positive skewed)

The mean is greater than median and the tail is longer on the right side.

Descriptive Statistics Measures of Central Tendency

The relationship between
mean, median and mode

2-) The data is skewed to the
right (positive skewed)



Descriptive Statistics

Measures of Central Tendency

The relationship between mean, median and mode

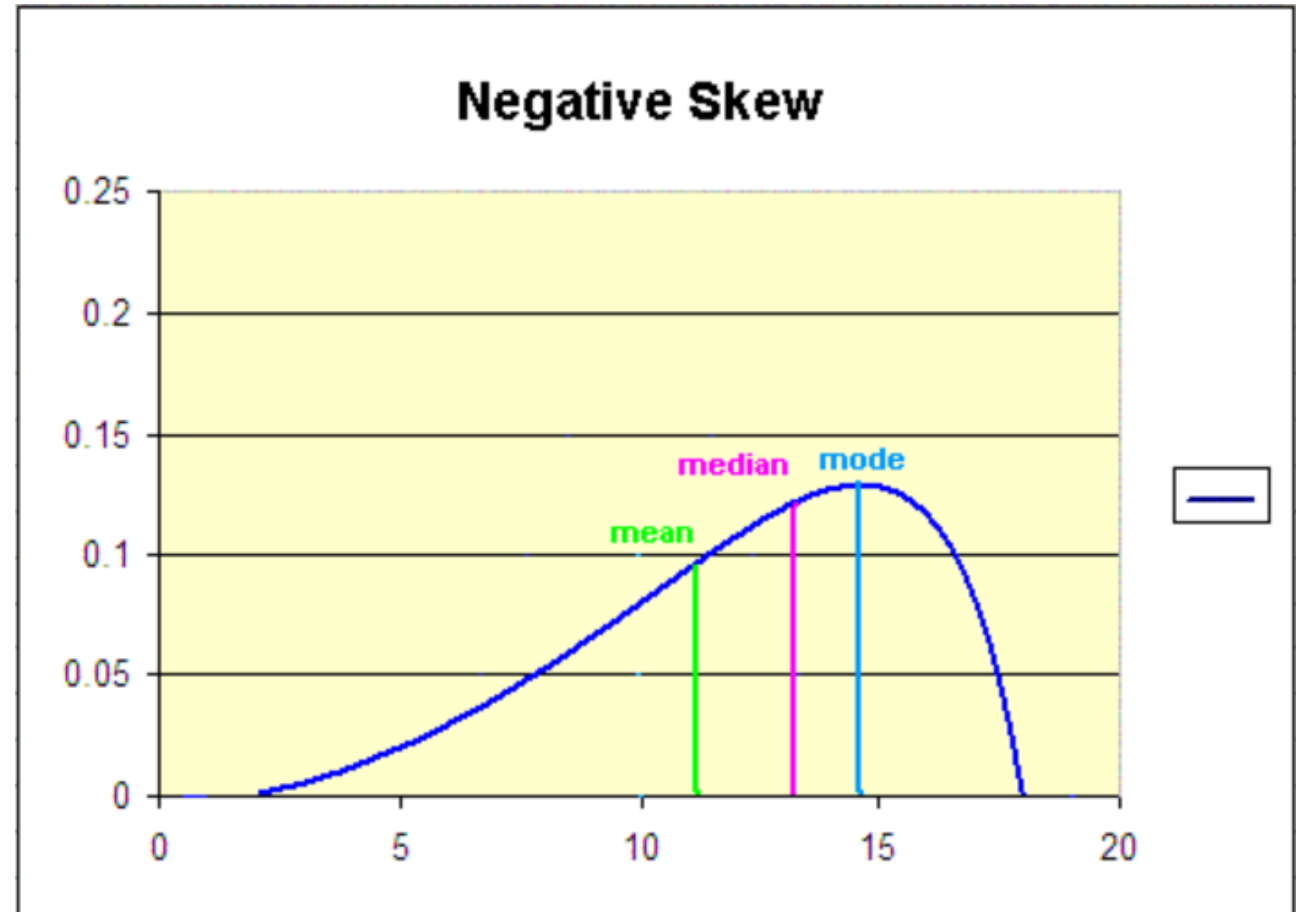
3-) The data is skewed to the left (negative skewed)

The mean is lower than median and the tail is longer on the left side.

Descriptive Statistics Measures of Central Tendency

The relationship between
mean, median and mode

3-) The data is skewed to
the left (negative skewed)



Reference

- Prem S. Mann 1998, Introductory Statistics, 7th edn, New York, USA.

Good Luck for All Students

- Please do not hesitate to contact me if you have any questions.
- Dr. Ibrahim AL-Jaafari
- www.Alghamdi-Biostatistics.com
- **Email.** Bio-stat@Hotmail.com
- Mobile Number : 0553777925