
بسم الله الرحمن الرحيم
In the name of ALLAH
the most Beneficent and the most merciful

MEASURES OF CENTRAL TENDENCY

BY

DR. ABDUL RAUF

MEASURES OF CENTRAL TENDENCY

When a series of observations of continuous series are made, it is found that a large no. of them concentrate at the center of series and a small numbers of them lie at the periphery. This tendency of the values to aggregate in the center of distribution series, is called “central tendency” also called “statistical average”.

In other words it measures central value of the distribution.

MEASURES OF CENTRAL TENDENCY

There are 3 types of central tendency

MEAN

MEDIAN

MODE

Mean, Median, Mode are all types of average. An average summarises groups of data.

MEAN

Mean is arithmetic mean unless otherwise specified.

Other means are geometric mean & harmonic mean.

These are uncommon one.

MEAN

- The arithmetic mean is widely used in statistical calculation. It is sometimes simply called Mean.
- To obtain the mean, the individual observations are first added together, and then divided by the number of observations.

MEAN

- The operation of adding together is called 'summation' and is denoted by the sign ∑ or S.
- The individual observation is denoted by the sign η
- and the mean is denoted by the sign $\bar{x}$ (called “x bar").

MEAN

- The mean ($\bar{x}$) is calculated thus :
- The diastolic blood pressure of 10 individuals was
  - 83, 75, 81, 79, 71, 95, 75, 77, 84, 90.
- The total was 810.
- The mean is 810 divided by 10 which is 81.0.

Mean

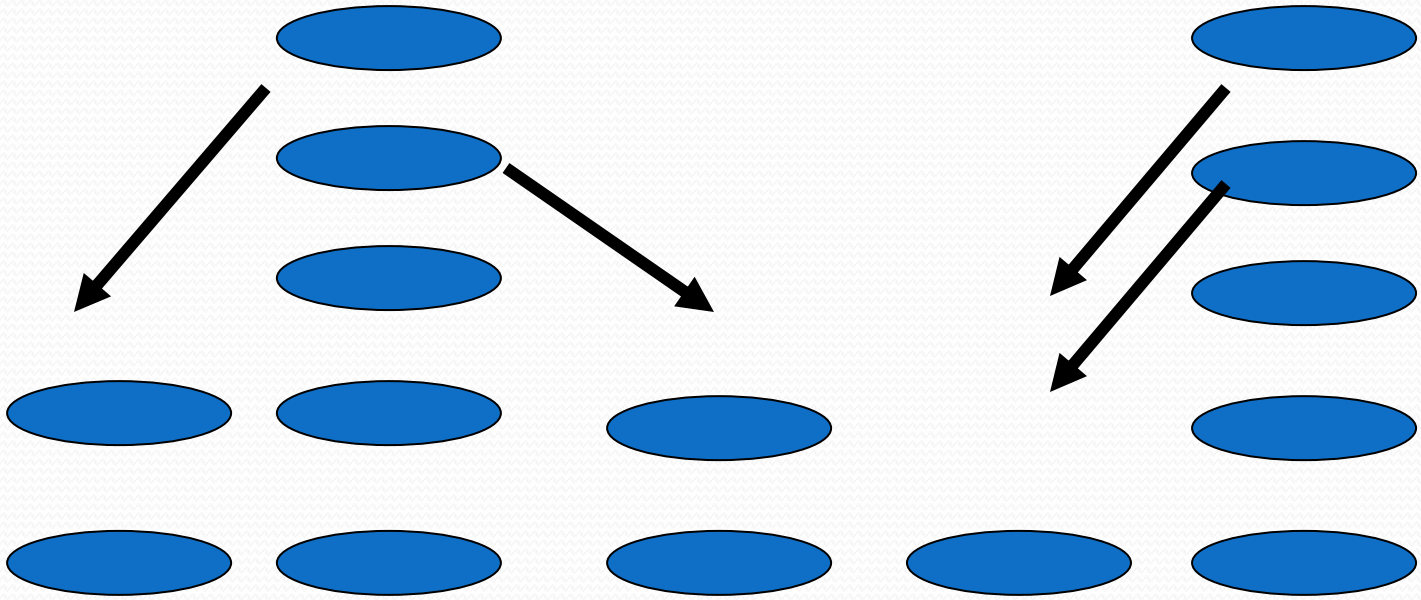
The average of a group of numbers.

2, 5, 2, 1, 5

Mean = (2+5+2+1+5)/5 =3

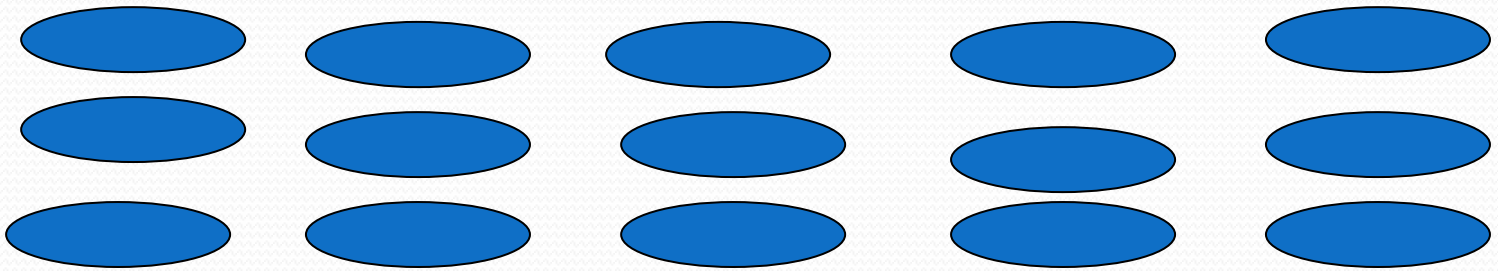
Mean is found by evening out the numbers

2, 5, 2, 1, 5



Mean is found by evening out the numbers

mean = 3



MEAN

For grouped data

$$\bar{x} = \frac{\sum fx}{\sum f}$$

Example

Missing teeth	No. of patients
0	5
1	3
2	4
3	6
4	2
5	1

MEAN

Missing teeth(x)	patients(f)	fx
0	5	0
1	3	3
2	4	8
3	6	18
4	2	8
5	1	5
∑=	21	42

MEAN

$\bar{x} = \sum fx / \sum f$

By putting values

$= 42/21 = 2$

so mean is = 2 missing teeth

Advantages

- It is defined by a mathematical formula
- It is easy to calculate & understand
- It is the most widely used "average"
- It depends on all values of the data and a change in any value, changes the value of the mean.

Disadvantages

1. It is influenced by extreme values
2. It is "not an appropriate average for highly skewed distribution"
3. Sometimes it may even look ridiculous;
   for instance, the average number of children born to a woman in a certain place was found to be 4. 76,
   which never occurs in reality.

Definition

Median

is in the

Middle

Median

- The median is an average of a different kind, which does not depend upon the total and number of items.
- To obtain the median, the data is first arranged in an ascending or descending order of magnitude, and then the value of the middle observation is located, which is called the median

- **Median** – The middle number in a set of ordered numbers.

1, 3, 7, 10, 13

Median = 7

Median

Formula for ungrouped Data

Median = If total numbers of items is odd no.

$$= \frac{(n+1)\text{th value}}{2}$$

Median= If total numbers of items is even no.

= average of middle 2 observations

What is the median of these numbers?

16, 10, 7

First arrange the data in ascending order.

Solution: 7, 10 , 16

Median=10

What is the median of these numbers?

29, 8, 4, 11, 19

Sol:4, 8, 11, 19, 29

Median=11

- Note : If there are two middle numbers, find the median of these two numbers.

18, 19, 21, 25, 27, 28

$$\frac{21+25}{2} = 23$$ ⟵ median

What is the median of these numbers?

31, 7, 2, 12, 14, 19

2, 7, 12, 14, 19, 31

Median=12 + 14 = 26

$$\frac{26}{2} = 13$$

Median

1st arrange in ascending order as

2,3,6,7,8,8,9,9,10,11,11,13,42,52

As there are 14 patients, the average of the periods of stay corresponding to 7th& 8th patients is calculated as median.

$$\text{Median} = \frac{9+9}{2} = 9 \text{ days}$$

Median

1. Obtain class boundaries.
2. Find cumulative frequencies of all the classes in data.
3. Compute n/2 and compare this quantity with cumulative frequencies. Choose cumulative frequency (CF) which is equal to or just exceeds the n/2.
4. The class corresponding to this CF is known as median class in which median lies.

Median

Formula for grouped data

$$\text{Median} = l + \frac{h\,(n-c)}{f}$$

l =Lower limit of the median group

h =Length of class interval (C-I) e.g.(2-4)

n =Total no. of frequencies

f= frequency of the median group

C =Cumulative frequency proceeding to the median group.

Example

The distribution of marks obtained by 50 students in Biostatistics, are shown as below.

Calculate "median"

marks	10-19	20-29	30-39	40-49	50-59	60-69	70-79
students	7	9	4	1	16	9	4

	C – B	f	c . f
10 – 19	9.5 – 19.5	7	7
20 – 29	19.5 – 29.5	9	16
30 – 39	29.5 - 39.5	4	20
40 – 49	39.5 – 49.5	1	21
50 – 59	49.5 – 59.5	16	37
60 – 69	59.5 – 69.5	9	46
70 – 79	69.5 – 79.5	4	50

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - c\right)$$

$$= 49.5 + \frac{10}{16}(25 - 21)$$

$$= 49.5 + \frac{10}{16}(4)$$

$$= 49.5 + 10/4$$

$$= 49.5 + 2.5$$

median=52

Advantages

- It eliminates the effect of extreme values
- It is easy to calculate & understand
- Only the values of the middle item need to be known

Disadvantages

- If you change the extreme value, median does not have any effect.
- It can not be calculated unless the data is arranged in some order (ascending order or descending order)

Median

- The relative merits of median and mean may be examined from the following example :
- The income of 7 people per day in Rupees was as follows : · 5, 5, 5, 7, 10, 20, 102 = (Total 154)
- The mean is 154 divided by 7 which is 22; the median is 7 which is the value of the middle observation.

Median

- In this example, the income of the seventh individual has seriously affected the mean, whereas it has not affected the median.
- In an example of this kind median is more nearer the truth, and therefore more representative than the mean.

MODE

- MODE is the most frequently occurring value in a set of observations.
- If the data is presented in a curve form, then the peak of the curve will represent the mode.

MODE

Following is the weight(in Kg)of new born babies

3, 3.1, 3.2, 3, 2.9, 2.8, 2.6, 3, 2.5, 2.7

What is the "**mode**"

Mode = 3

MODE

Example

Following is the ages of 10 medical students:

18, 18, 19, 19, 20, 20, 20, 21, 22, 23

What is the “mode”?

mode = 20 years of age

MODE

Example

To check the accuracy of the clinical diagnosis of malaria, blood slides of 33 patients were examined for malaria parasites. There were three possible results:

Negative

P. falciparum

P vivax

MODE

Example

The results are presented in the following frequency distribution.

Negative	19
P. falciparum	13
P. vivax	1
Total	33

What is the mode?

The mode is "Negative."

MODE

Example

Health personnel from 148 different rural health institutions were asked the following question.

"How often have you run out of drugs for the treatment of malaria in the past two years?"

This was a closed question with the following possible answers.

Never

1 to 2 times (rarely),

3 to 5 times (occasionally),

more than 5 times (frequently)

MODE

. The numbers of responses in each category were totaled to give the following frequency distribution.

• Never	47
• Rarely	71
• Occasionally	24
• Frequently	6
Total	**148**

What is the mode?

The mode is "rarely."

MODE

Example

82 clinics in one district were asked to submit the number of patients treated for malaria in one month. The researchers presented both the frequency distribution and percentages (or relative frequencies) as follows

MODE

NUMBER OF PATIENTS	NUMBER OF CLINICS	RELATIVE FREQUENCY
0 to 19	25	31%
20 to 39	3	4%
40 to 59	5	6%
60 to 79	11	14%
80 to 99	19	24%
100 to 119	10	12%
120 to 139	4	5%
140 to 159	3	4%
Total	80*	100%

MODE

Data form two clinics are missing

Note: Usually you do not include missing data in the calculation of percentages

However, the number of missing data (e.g., people who did not respond to a question) is a useful identification of the adequacy of your data collection. Therefore, this number should be mentioned, as a note to your table.

MODE

What is the mode?

The mode is "0 to 19", as this outcome is recorded most frequently (25 times out of 80).

MODE

There can be more than one mode for a series of data. In a distribution with two most frequent values, there will be 2 modes: Bimodal distribution

Mode= average of 2 modes

MODE

Grouped data

$$\text{Mode} = l + \frac{fm - f_1}{(fm - f_1) + (fm - f_2)} \times h$$

l = lower class boundary of modal group

fm = frequency of modal group

f_1 = proceeding frequency of modal group

f_2 = following frequency of modal group

h = class interval of modal group

MODE

Example

Following are the number of men in various age groups with some form of paid employment in a village. The age recorded for each man is the number of completed years lived.

Calculate "mode"

Age =	14-20	21-30	31-40	41-50	51-60	61-70
men =	12	14	26	35	23	5

MODE

age	f	Class boundaries
14 – 20	12	13.5 – 20.5
21 – 30	14	20.5 – 30.5
31 – 40	26	30.5 - 40.5
41 – 50	35	40.5 -50.5
51 – 60	23	50.5- 60.5
61 – 70	5	60.5 – 70.5
71 - 90	1	70.5 – 90.5

MODE

$$\text{Mode} = l + \frac{fm - f_1}{(fm - f_1) + (fm - f_2)} \times h$$

$$= 40.5 + \frac{35 - 26}{(35 - 26) + (35 - 23)} \times 10$$

$$= 40.5 + \frac{9}{9 + 12} \times 10 \quad ,= 40.5 + \frac{9}{21} \times 10$$

$$\text{Mode} = \frac{40.5+90}{21} = 44.8$$

MODE

In distribution with extreme values

- Most affected measure of central tendency;

MEAN

- Least affected measure of central tendency;

MODE

- Most preferable measure of central tendency;

MEDIAN

MODE

Example

The incidence of malaria in an area is

20,20,50,56,60, 5000, 345, 678456,,898,

Incidence in ascending order is

20,20,50,56,60,345,456,678,898,5000

Mean= $\sum$ x/n= 7583/10= 758.3

Median= average of 5^{th} & 6^{th} value=(60+345)/2

Median= 202.5

Mode= 20

ADVANTAGES

It is easy to calculate

It is least influence by extremes of values

DISADVANTAGES

It may not exist in a small group of values.

The exact location is often uncertain and is often not clearly defined. Therefore, mode is not often used in biological or medical statistics


Thank You