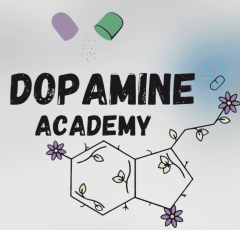


Statistics

Subject: Lecture 3



لجان الرفعات



Important characteristics of data Description:

1. Centre : a representative or average value that indicates where the middle value of the data set is located.
2. Variation : a measure of the amount that values vary among themselves.
3. Distribution : The nature or shape of the distribution of data as bell-shaped, uniform, skewed
4. Outliers : Sample values that lie very far away from the vast majority of other sample values.

measure of central tendency :

1. Mean = Arithmetic mean = Average

2. Median : half way of data.

3. Mode : most common value observed. للنوال

* most frequently

• In a normal distribution, mean and median are the same.

* Sample lower case n

• population mean is usually Unknown.

population upper case N

$$\text{Mean of intervals} = \frac{\sum (f_i \cdot x_i)}{\sum f_i (n)}$$

x_i : mid point of each class interval.

f_i : Frequency

Subject :

Ex :

Driving Times (min)

Number of teachers

x mid point

0 - 10

3

$0 + 10 / 2$

5

10 - 20

10

$10 + 20 / 2$

15

20 - 30

6

$20 + 30 / 2$

25

30 - 40

4

$30 + 40 / 2$

35

40 - 50

2

$40 + 50 / 2$

45

Calculate the mean of the driving times ?

25

15

+ 150

150

140

90

545

$$\text{Mean of intervals} = \frac{\sum (f_i \cdot x_i)}{n} = \frac{545}{25} = 21.8$$

properties of the mean ?

1. Uniqueness

2. Is used in computing other statistics, such as variance.

3. Simplicity : it is easy to compute.

4. Is affected (not robust) by outliers.

properties of the median ?

1. Uniqueness

2. Simplicity

3. Robust measure

The median of grouped data - simple frequency Table :-

1) Find the CF

2) Find the $n/2$

n : Sum of the frequencies

3) median is the value that having $CF \geq n/2$

Subject :

Ex:	Number of Accidents	2	7	12	17	22	
$\Sigma f(n)$	Number of weeks	5	22	13	8	2	50

Calculate the median for the number of accidents per week ?

F CF

$$\frac{n}{2} = \frac{50}{2} = 25$$

الأصغر يقابل الجدول هو دي

22 (27)

13 46

$$MD = CF \geq n/2$$

8 48

2 50

$$MD = 7$$

50 ✓

The median of intervals :

1. find the class median $CF \geq n/2$

2. median = $L_m + \left(\frac{\frac{n}{2} - F_m}{f_m} \right) \cdot i$ width : the difference between the

lower boundary of the class median F of the class median CF before lower endpoint of an interval and the lower endpoint of the next interval.

OR (Upper limit - lower limit) + 1

$L - 0.5$ (10 - 20)

$$L_m = 9.5$$

$$L_m = 20.5$$

Ex: Time to travel to work frequency CF

$$i = 11 - 1 = 10$$

$$\frac{n}{2} = \frac{50}{2} = 25$$

1 - 10

8

8

11 - 20

14

22

Class median

21 - 30

12

34

$$MD = 20.5 + \left(\frac{25 - 22}{12} \right) \cdot 10$$

31 - 40

9

43

41 - 50

7

50

$$MD = 23$$

$$n = \Sigma f = 50$$

► Subject :

- Sometimes there are more than one mode.

Properties of the mode :-

1. Sometimes it is not unique.
2. Simple
3. Robust measure.
4. It may be used for describing qualitative data.
5. It can be calculated for both quantitative and qualitative data.

interval

The mode of grouped data (Frequency Table) :-

1. Find the modal class : is the class with highest frequency.

$$\text{Mode} = L_{mo} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) \cdot i$$

↓ ↓ ↓

lower boundary of modal class $f_{mc} - f_{bmc}$ $f_{mc} - f_{ame}$ class width

Ex: Time to travel
to work

Frequency

Calculate the mode ?

	1 - 10	8
modal class	11 - 20	14
	21 - 30	12
	31 - 40	9
	41 - 50	7

$$\begin{aligned} \text{Mode} &= L_{mo} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) i \\ &= 10.5 + \frac{6}{8} \times 10.5 \\ &= 18 \end{aligned}$$

$$\Delta_1 = f_{mc} - f_{bmc} = 14 - 8 = 6$$

$$\Delta_2 = f_{mc} - f_{ame} = 14 - 12 = 2$$

$$i = 10$$

$$L_{mo} = 10.5$$