Bringing British Education to You
www.nccedu.com

Foundation Mathematics

Topic 8 – Lecture 1: Understanding Dispersion

The Range and Quartiles

The Mean Deviation

Scope and Coverage

This topic will cover:

- Calculation of the range and quartiles
- Calculation of the mean deviation

Learning Outcomes

By the end of this topic students will be able to:

- Calculate the range and quartile values for data sets
- Calculate the mean deviation for data sets

Measures of Dispersion - 1

- Although averages are a good method of determining the central point of a distribution of data, they give no real information about the "spread" of the distribution.

Value	Distribution Frequency X	Distribution Frequency Y
x	f	f
1	3	0
2	6	0
3	10	10
4	12	30
5	10	10
6	6	0
7	3	0
Total	50	50

- This table gives us a great deal of information about the frequency and distribution of our values

Measures of Dispersion - 2

- What we can see from our data is that the averages for both X and Y are the same.

Value	Distribution Frequency X	Distribution Frequency Y
x	f	f
1	3	0
2	6	0
3	10	10
4	12	30
5	10	10
6	6	0
7	3	0
Total	50	50

- However, further examination shows us that although the averages are the same (4), the way in which the data is dispersed is very different.

Measures of Dispersion - 3

- The differences that we see in these distributions is caused by the spread or dispersion.
- Measures of dispersion can give us an idea about the spread of a variable.

The Range

- The range is simply the difference between the highest and lowest observations within a data set.
- It gives us a quick way of determining the spread of dispersion. We can calculate the mean and range.
  - E.g. X = 3, 5, 6, 7, 15, 24, 8, 2, 4
- The mean value $\bar{x}$ = 74/9 = 8.22 the range is 24 – 2 = 22
- However, the range is easily affected by extreme values and it gives no indication of the spread of the extremes.

Quartiles

- We order data on the basis of numerical value and classify it according to where it is positioned across our dispersion
- Our value may therefore be in:
- The lower quartile - the value which 25% of the observations are below
- The upper quartile - the value which 25% of the observations are above
- The middle point or second quartile is in fact the median

Quartiles – Example - 1

- These concepts are best illustrated by an example.
- Look at the raw data below.
  - 30, 4.5 ,5.3, 25, 23.6, 11.3, 18, 19.3, 19, 29.2, 28.9, 20, 18.2, 28, 21.3, 24.2, 9.2, 6,13,
- The first step when calculating quartiles for ungrouped data is to put the data into ascending order.
  - 4.5, 5.3, 6, 9.2, 11.3, 13, 18, 18.2, 19, 19.3, 20, 21.3, 23.6, 24.2, 25, 28, 28.9, 29.2, 30.

Quartiles – Example - 2

- We are dealing with ungrouped data and as we know that our second quartile is the same as the median (the mid point of the data set) we are able to calculate the mid point using the formula
- Therefore our calculation our quartiles becomes $\frac{(n+1)}{2}$

$$LQ = \frac{1}{4}(n+1)$$

$$MQ = \frac{1}{2}(n+1)$$

$$UQ = \frac{3}{4}(n+1)$$

In all cases n= total frequency of observations

Quartiles Example - 3

- For our data set
  - 4.5, 5.3, 6, 9.2, 11.3, 13, 18, 18.2, 19, 19.3, 20, 21.3, 23.6, 24.2, 25, 28, 28.9, 29.2, 30.
- We can identify that there are 19 observations therefore n = 19
  - Q1 = (19 + 1) /4 = 5th value
  - Q2 = 2(19 + 1)/4 = 10th value
  - Q3 = 3(19 + 1) /4 = 15th value
- These produce values of
  - Q1 = 11.3, Q2 = 19.3, Q3 = 25.

Quartiles - Grouped Data

- When dealing with grouped data we need to apply a simple formula. Consider the following data table:

Weight (kg)	**Frequency of observation**	**Cumulative Frequency**
$0 \leq w < 10$	2	2
$10 \leq w < 20$	3	5
$20 \leq w < 30$	7	12
$30 \leq w < 40$	15	27
$40 \leq w < 50$	8	35
$50 \leq w < 60$	4	39
$60 \leq w < 70$	1	40
Total Frequency	40	

Quartiles – Applying Formula - 1

- As our exact value for our quartile we need to apply a straightforward formula

$$Quartile = L + \frac{(Q - P)I}{f}$$

- Where:
  - L = Lower limit of quartile class
  - Q = Quartile position
  - P = Previous class frequency
  - I = Quartile class interval
  - F = Class frequency

Quartiles – Applying Formula - 2

- This gives us:
- For Quartile 1 $$20+\frac{(10-5)10}{7}=20+\frac{50}{7}=27.1kg$$
- For Quartile 2 $$30+\frac{(20-12)10}{15}=35.33kg$$
- For Quartile 3 $$40+\frac{(30-27)10}{8}=43.75kg$$

The Quartile Range

- The information obtained from the calculation can be further augmented by combining quartiles to produce new measures such as the quartile range and the quartile deviation range.
- The quartile range is the difference between Q3 and Q1.
- In the previous example the quartile range is simply calculated by subtracting the value of Q1 from Q3
  - 43.7 Kg (Q3) - 27.1Kg (Q1) =16.6 Kg.
- This indicates that 50% of observations occur within a spread of 16.6 Kg.

The Mean Deviation

- Calculating the **mean deviation** is very straightforward which will be demonstrated by the following example.
- In this example you will encounter the following notation I$x-\bar{x}$I . The vertical lines I I simply mean that we ignore the sign of the value inside, this means that the value of this calculation will always be positive. Therefore:

  I 4 – 34 I is in fact 30 for our purposes

The Mean Deviation – Example - 1

- Given the following Data
  - 10,12,13,14,15,16,17,18,21,11,13
- We can calculate that our mean $\bar{x}$ = 160/11 = 14.55

Presenting our data
In a table gives us

x	$\bar{x}$	$\lvert x - \bar{x} \rvert$
10	14.5	4.5
12	14.5	2.5
13	14.5	1.5
14	14.5	0.5
15	14.5	0.5
16	14.5	1.5
17	14.5	2.5
18	14.5	3.5
21	14.5	6.5
11	14.5	3.5
13	14.5	1.5
		28.5

The Mean Deviation – Example - 2

- We are then able to use the following formula

$$\frac{\sum |x - \bar{x}|}{n}$$

- To calculate our mean deviation which from our table gives us

28.5/11 = 2.59

- Each individual observation lies an average (arithmetic mean) of 2.59 units away from the mean value.

Mean Deviation for Grouped Data - 1

- For grouped data we are faced with the same issues as for other analysis of group data – our value is hidden in the class interval
- Our first task is to calculate the mean using $\bar{x} = \frac{\sum fx}{\sum f}$

Weight (Kgs)	Midpoint x	$\lvert x-\bar{x} \rvert$	f	$f\lvert x-\bar{x} \rvert$
0 ≤ w < 10	5	30	2	60
10 ≤ w < 20	15	20	3	60
20 ≤ w < 30	25	10	7	70
30 ≤ w < 40	35	0	15	0
40 ≤ w < 50	45	10	8	80
50 ≤ w < 60	55	20	4	80
60 ≤ w < 70	65	30	1	30
Totals			40	380

Mean Deviation for Grouped Data - 2

- The formula by which we can calculate our grouped data is

$$\frac{\sum f\left|x-\bar{x}\right|}{\sum f}$$

- By substituting in our values we get 380/40 = 9.5
- Therefore individual observations in the sample lie on average 9.5 Kg away from the mean of 35 Kg.

Topic 8 – Understanding Dispersion 1

Any Questions?

Bringing British Education to You
www.nccedu.com