# Median and Cumulative Frequency.

Please note every effort has been made to make this pack accessible but due to the nature of some of the content, the charts and graphs may not be accessible by screen readers. If you have any queries or need any help, please email learningdevelopment@northampton.ac.uk

## The Median Temperatures.

The **Median** is a set of numbers which is in the middle when the numbers have been arranged in size order. Look at the following examples.

**Example 1.**

 Find the median value of the numbers 3, 7, 2, 9, 10, 6, 4.

**Answer.**

 First arrange the numbers in order of size: 2, 3, 4, 6, 7, 9, 10.

 There are several numbers so in this case the median is the fourth one.

 Median = 6

If it should happen that there are an even number of values, then the median is the average of the two centre numbers.

**Example 2.**

 Find the median value of the following lengths

 12 cm, 19 cm, 4 cm,13 cm, 20 cm, 39 cm.

**Answer.**

 Arrange the lengths in size order: 4 cm,12 cm, 13 cm, 19 cm, 20 cm, 39 cm.

There are six items in the set (an even number) so there are two middle values, the third and the fourth. The median is the average of these.

 Median = $\frac{13+19}{2}=\frac{32}{2}=16cm$

In general, to **find the median value of distribution**, if there are n terms in the distribution the median is given by the $\frac{n+1}{2}$th term. (When the terms have been arranged in size order).

 If there were 35 values the median would be the $\frac{35 + 1}{2}=$18th value.

If there were 36 values the median would be the $\frac{36+1}{2}=$ 18½th value. In other words, the average of the 18th and 19th values.

**Exercise 1.**

1. Find the median of 38, 19, 2, 8, 15, 22, 13, 6, 5.

2. What is the median of: 9, 8, 17, 47, 3, 4, 15, 40, 18, 28.

### To Find the Median of Frequency Distribution.

**Example 3.**

Look at the following distribution. It refers to the number of apples found in a pre-packed 2 kg bag.

16 apples in 3 bags.

17 apples in 5 bags.

18 apples in 6 bags.

19 apples in 5 bags.

20 apples in 2 bags.

Total number of bags is 21.

**Answer.**

 The median of 21 values is given by the $\frac{21+1}{2}=11th$ term .

To find out which is the 11th term we could list the number of apples in each bag in size order as we have done before:

Number of apples: 16, 16, 16, 16, 17, 17, 17, 17,17, 18, 18, 18, 18, ….etc.

 The 11th term is 18.

As you can see this method will be long and tedious for large distribution. A better method is to use cumulative frequency distribution.

## The Cumulative Frequency Distribution.

Look at the cumulative frequency distribution for our example. Where the frequency is the number of bags containing *x* number of apples:

Less than 16 apples in 0 bags.

Less than 17 apples in 3 bags.

Less than 18 apples in 8 bags.

Less than 19 apples in 14 bags.

Less than 20 apples in 19 bags.

Less than 21 apples in 21 bags.

**Note:**

1. The total frequency of the distribution is the same as the cumulative frequency of the last class.
2. This cumulative frequency distribution gives you the following information:

a) No bags contain less than 16 apples; 3 bags have less than 17 apples; 8 bags have less than 18 apples; 14 bags have less than 19 apples and so on.

b) The median value of this distribution is 18 apples. We know this because the median is the 11th value and this must lie in the class ‘less than 19’.

### Finding the Median Of A Grouped Cumulative Frequency Distribution.

Suppose that we want to find the median height of the class of the first school children. This is the cumulative frequency distribution.

**Example 4.**

The frequency (f) is the number of pupils recorded for each height range:

 80 to 89cm for 2 pupils.

 90 to 99cm for 10 pupils.

100 to 109cm for 16 pupils.

110 to 119cm for 24 pupils.

120 to 129cm for 16 pupils.

130 to 139cm for 11 pupils.

140 to 149cm for 1 pupil.

Total frequency $=80$

**Answer.**

In order to find the median, we must first form the cumulative frequency distribution. Where cf represents Cumulative Frequency.

 Height of less than 89.5cm has cf = 2

 Height of less than 99.5cm with cf = 12

 Height of less than 109.5cm with cf = 28

 Height of less than 119.5cm with cf = 52

 Height of less than 129.5cm with cf = 68

 Height of less than 139.5cm with cf = 79

 Height of less than 149.5cm with cf = 80

**Notice** how the class limits of the cumulative frequency distribution are chosen. They are taken from the original distribution. The limit of each class is given by the lower boundary of the next higher class.

 This is easier to understand if you look at what it means in this example.

The heights are given to the nearest cm so that 80 to 89 cm really means 79.5 to 89.49 cm (in other words less than 89.5cm); 90 to 99cm means 89.5 to 99.49 and so on.

There are 80 children in the distribution so we can see that the median will come somewhere in the class ‘less than 119.5 cm’. That is, somewhere between 109.5 cm and 119.5 cm. We can get a better estimate of the median than this by using the cumulative frequency curve. This is explained in the next section.

### The Cumulative Frequency Curve (sometimes called an ogive).

Please note the graphs in this section may not be accessible by some screen readers. If you have any queries or need any help, please email learningdevelopment@northampton.ac.uk

The ogive is formed by plotting points from the cumulative frequency distribution and joining them with a smooth curve. The median is then the value of the variable which corresponds to ½(n + 1)th value. If n is reasonably large, we can say this is approximately equal to half to the total frequency. We can find out many things about a distribution by looking at the cumulative frequency curve/ the median is one of them, we shall look at some others in the next example.

**Ogive to show all the heights of 80 children**



**Notice** that this is a ‘less than’ curve. Cumulative frequency curves usually have this ‘flattened S’ shape and that was the reason why they were called ogives (an architectural term). The horizontal scale measures the variable of the distribution. The vertical scale tells us how many are ‘less than’ a particular value. Some examples of the problem we could solve by using the ogive are on the following pages.

Always use graph paper when drawing cumulative frequency curves.

**Example 5.**

 Use the cumulative frequency curve to find the median of the distribution of 80 first school children.

**Answer.**

 First draw the ogive.

 

The median is the value which divides the distribution into two halves, that is the 40th value, and this is read from the graph.

Median = 114 cm.

In this example we have used the ogive to find the median. Other statistics can also be obtained from it and some of these are explained briefly below. The most important of these for this course are the **quartiles.**

## Quartiles.

The median divides a distribution into two equal parts. The quartiles divide it into four equal parts. The second quartile is the same as the median.

Here is a distribution already arranged in size order.

1, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 5, 6, 6, 7, 7, 7, 8

There are 19 values so the median (M) is the $\frac{9+1}{2}=$ 10th value.

The lower quartile (LQ) is the ¼ (19 + 1) = 5th value

The upper quartile (UQ) is the ¾ (19 +1) = 15th value

Therefore, M=5, LQ=3,UQ=6.

For a continuous distribution or a grouped frequency distribution it is better to find the quartiles in the same way that we found the median in example 5. That is: first form a cumulative frequency distribution and from it draw a cumulative frequency curve: then find quartiles from this graph.

### Percentiles and Deciles.

Quartiles divide distributions into four equal parts. Percentiles are the statistics which divide them into 100 equal parts and deciles into 10 equal parts. (The upper quartile could also be called the 75th percentile)

The collective name for quartiles, deciles and percales is **quantiles**.

**Example 6.**

Using the ogive drawn for example 4 find the upper and lower quartile for the distribution of heights of first school children.



**Answer.**

The lower quartile (LQ) is given by the ¼(80) = 20th value

The upper quartile (UQ) is given by the ¾(80) = 60th value

From the graph:

LQ = 104 cm

UQ = 124 cm

**Example 7.**

Looking again at the frequency distribution of the 80 first school children, example 4. We could use the cumulative frequency curve to answer the following questions.

a) How many of the pupils are less than 125 cm tall?

b) How many are greater than 125 cm?

c) What height limit should we set if we wish to give some special vitamin supplements to the smallest 30 children?

d) What should the minimum height be if we wanted to find the tallest 40% of the children?

**Answer.**

.

The arrows show you which way to read from the graph.

All of these answers have been read from the ogive.

a) 61

b) The number greater than 125 cm is equal to the total minus the number less than 125 cm. 80 – 61 = 19

c) 111 cm

d) 40% of 80 = 40 x 80 = 3200

 $=\frac{3200}{100}=32$

Therefore, 32 children will be taller than the limit set which means that $80-32=48 $will be less than this limit. We can see from the ogive that the limit should be set at 118 cm.

The **Semi-Interquartile Range, SIR**, = Upper Quartile subtract Lower Quartile $÷2$

**Exercise 2.**

1. Form the cumulative frequency distributions from the following:

1. Score 0 frequency 1

 Score 1 frequency 4

 Score 2 frequency 4

 Score 3 frequency 6

 Score 4 frequency 3

 Score 5 frequency 2

b) **Weight (kg)**  **Frequency (f)**

 130 to 134kg f = 7

 135 to 139kg f = 17

 140 to 144kg f = 40

 145 to 149kg f = 31

 150 to 155kg f = 5

2. Find a) the median b) the quartiles of the following

 M = Marks and f = frequency

 1 M f = 2

 2 M f = 1

 3 M f = 3

 4 M f = 5

 5 M f = 6

 6 M f = 8

 7 M f = 4

 8 M f = 3

 9 M f = 2

 10 M f = 1

3. The following data shows the frequency distribution of examination marks for 120 candidates.

 Exam Marks 10 to 19 2 candidates

20 to 29 6 candidates

30 to 39 7 candidates

40 to 49 14 candidates

50 to 59 20 candidates

60 to 69 35 candidates

70 to 79 29 candidates

80 to 89 6 candidates

90 to 99 1 candidate

 Construct a cumulative frequency table and use it to draw a cumulative frequency curve of the distribution. From the curve, estimate:

a) the median mark

b) the upper and lower quartiles

c) If the pass mark is 45, what percentage pass the paper?

d) If 75% of the candidates are to pass, what should the pass mark be?

4. A sample of 150 people were asked to give the number of hours they watched BBC 1 during the week. The results were as follows (f = frequency):

 Number of hours 0 to 5 f = 2

 6 to 10hrs f = 8

 11 to 15hrs f = 17

 16 to 20hrs f = 52

 21 to 25hrs f = 38

 26 to 30hrs f = 26

 31 to 35hrs f = 5

 36 to 40hrs f = 2

 Construct a cumulative frequency diagram and from it estimate:

 a) the median

 b) the percentage of people who watched 17 hours or more each week.

**ANSWERS.**

**Exercise 1.**

a) Median = 13

b) Median $=\frac{15+17}{2}=16$

**Exercise 2.**

1. a) Cumulative Frequency = cf

 Score less than 1 cf = 1

 less than 2 cf = 5

 less than 3 cf = 9

 less than 4 cf = 15

 less than 5 cf = 18

 less than 6 cf = 20

1. Cumulative Frequency = cf

 Score less than 134.5 cf = 7

 less than 139.5 cf = 24

 less than 144.5 cf = 64

 less than 149.5 cf = 95

 less than 154.5 cf = 100

2. a) Total frequency = 35

 Median $=\frac{35+1}{2}=$ 18th value

 Lower Quartile $=\frac{1}{4}(35+1)$ 9th value

 Upper Quartile $=3(35+\frac{1}{4})$27th value

 Forming cumulative frequency

 M = Marks; f = frequency and cf = cumulative frequency

 1 M f = 2 cf = 2

 2 M f = 1 cf = 3

 3 M f = 3 cf = 6

 4 M f = 5 cf = 11

 5 M f = 6 cf = 17

 6 M f = 8 cf = 25

 7 M f = 4 cf = 29

 etc.

b) Median = 6

 Lower Quartile = 4

 Upper Quartile = 7

 3. **cumulative frequency = cf**

 Exam Mark less than 20 cf = 2

 less than 30 cf = 8

 less than 40 cf = 15

 less than 50 cf = 29

 less than 60 cf = 49

 less than 70 cf = 84

 less than 80 cf = 113

 less than 90 cf = 119

 less than 100 cf = 120

From the graph:-

a) The median mark = 63.5

b) Lower Quartile = 50 Upper Quartile = 72.5

c) The pass mark = 45. 21 students score less than this.

 Therefore $120-21=99$

 $=\frac{99}{120} $ x 100 $=82\frac{1}{2}\% $of students pass the paper

d) 75% are to pass $\frac{75}{100}$ x 120 = 90

 Therefore 30 are to fail. Pass mark should be 50.

4. **cf = cumulative frequency**

 Less than 6hrs cf = 2

 Less than 11hrs cf = 10

 Less than 16hrs cf = 27

 Less than 21hrs cf = 79

 Less than 26hrs cf = 117

 Less than 31hrs cf = 143

 Less than 36hrs cf = 148

 Less than 41hrs cf = 150

a) Median = 75th value = 20.5 hours

b) 35 people watched 17 hours or less per week.

 Therefore $150-35=115$ people watched 17 hours or more each week.

 Expressed as a percentage,

 $\frac{115}{120}$ x 100 = 95.8%

This concludes the Statistics – Median Cumulative Frequency study pack.