Cumulative frequency (9)

Contents

Introduction

In the notes on "Averages & Spread" we saw that it is too difficult to find a median from a grouped frequency table. In this case, a cumulative frequency graph was required. Let us now see what we mean by cumulative frequency and use this to find the median and the answer to a variety of other calculations.

1 What is cumulative frequency?

Consider the grouped frequency table below:

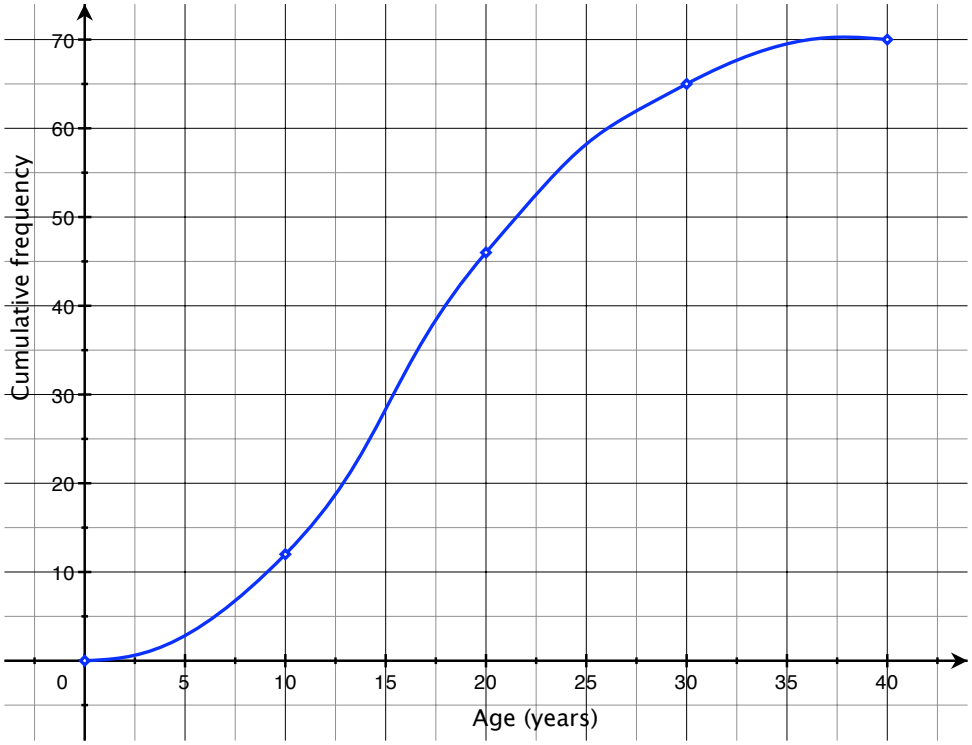
Age	Frequency
$0-10$	12
$10-20$	34
$20-30$	19
$30-40$	5

If you were asked the question "How many people are under 20?" you could see that this is all of the people in the $0-10$ group and all of those in the $10-20$ group, giving a total of $12+34=46$.
This question describes exactly what **cumulative frequency** is: it is how many people are **less than (or equal to)** the upper class boundary (that is, the top end) of any group:

Age	Cumulative frequency
≤ 10	$=12$
≤ 20	$12+34=46$
≤ 30	$46+19=65$
≤ 40	$65+5=70$

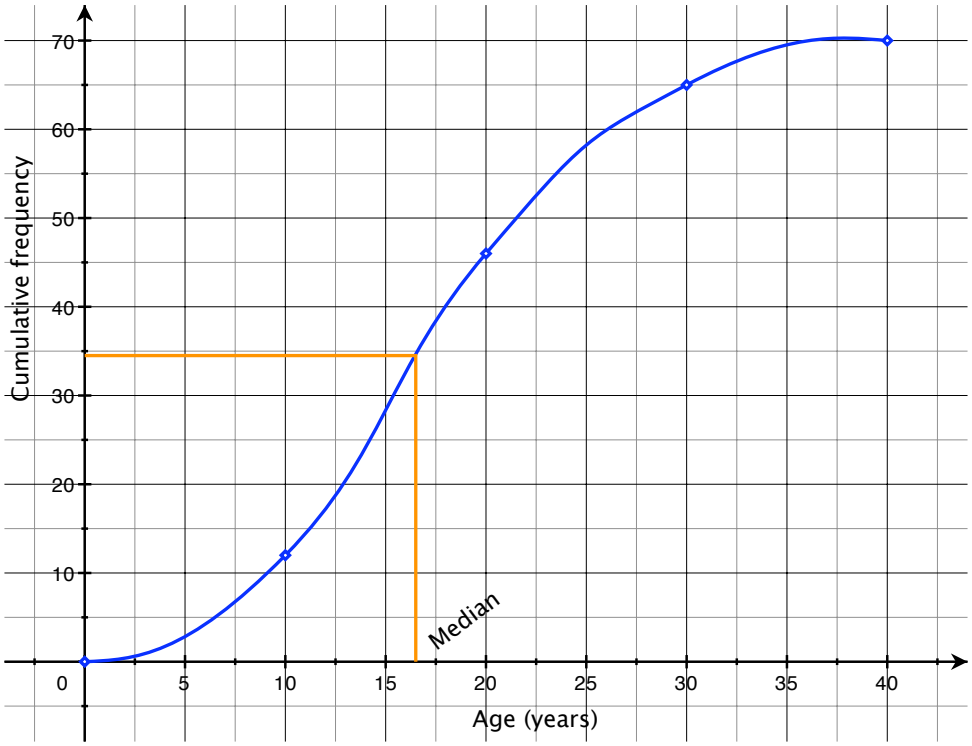
2 Cumulative frequency curve

We can plot the cumulative frequencies against the *upper class boundary* of each group to produce a cumulative frequency graph (join with straight lines to give a polygon or curves to give a curve)

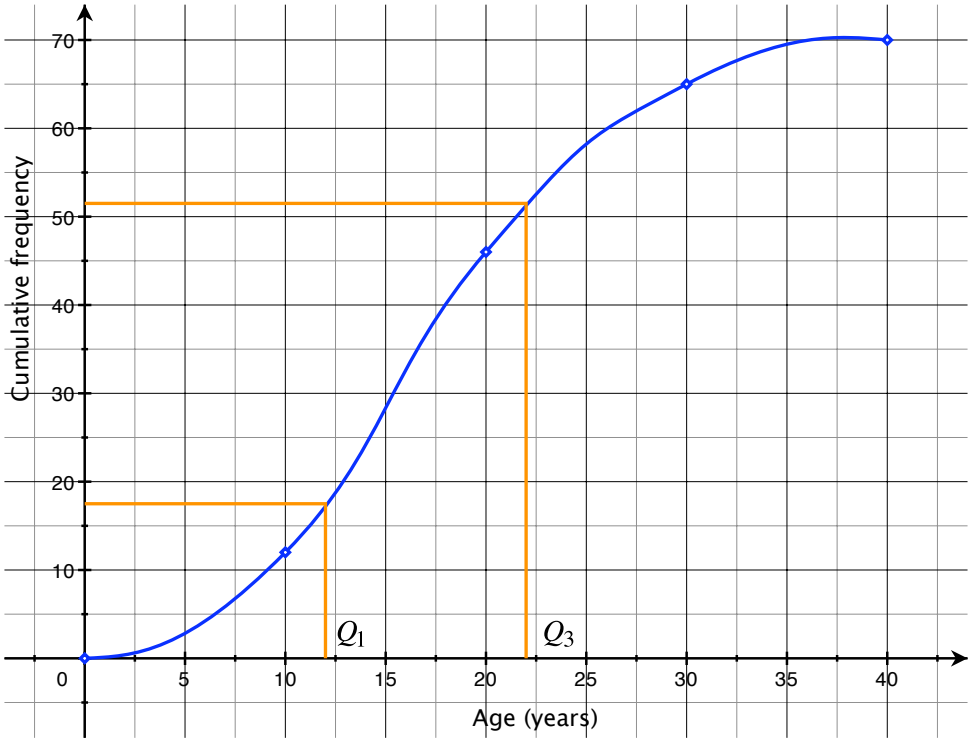


3 Using the cumulative frequency graph to answer questions

Median. Since there are 68 data values, the median is found at the 34.5th position (see lesson on boxplots for formula on positon of median, lower quartile and upper quartile). If we look across and down from the 34.5th position, we find the median is approximately 17 years old.



Quartiles. Similarly, the lower quartile can be found at the 17.5th position and the upper quartile at the 51.5th position:



We can see that the lower quartile is $Q_1 = 12$ and the upper quartile is $Q_3 = 22$. Hence, the interquartile range is $23 - 12 = 11$ years.

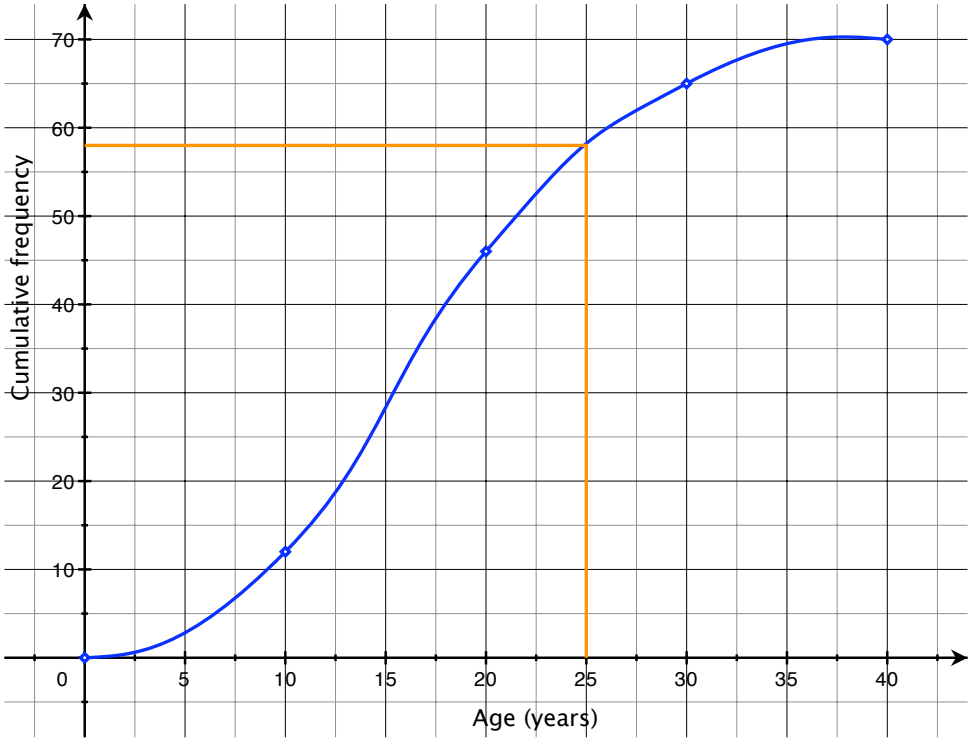
Comment (see boxplot lesson)

On average, the people shown in this graph are 17 years old. This is shown by a median of 17 years. This means that these people are quite young, on average.
The middle 50% of people have an 11 year age gap. This is shown by an interquartile range of 11. This means that there is not a great variety in ages of the majority of people who will fall below 30 years of age.

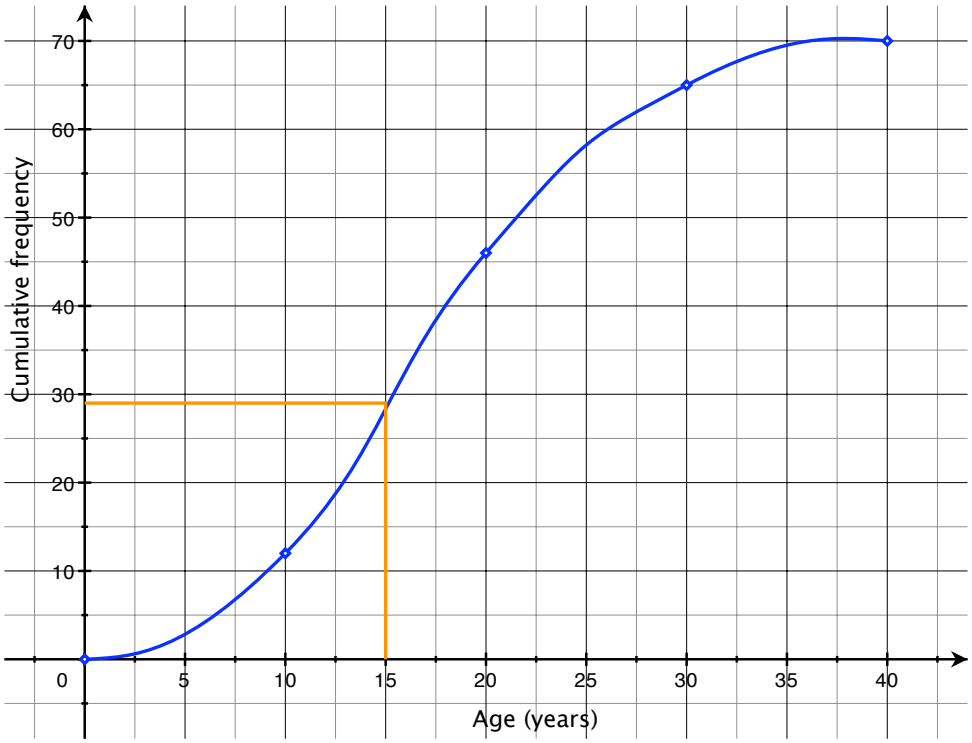
4 Other calculations

Remember the fact that cumulative frequency means **less than**. If you have to answer a **less than** question from the graph this should be quite straight forward.

Example. How many people were less than 25 years old?



Roughly 58 people were less than 25 years of age.
If the question involves **more than**, it may involve much more thought. E.g. How many people were **more than** 15 years old?



The cumulative frequency we read off is 29, but this is the number of people who are less than 15 years. Since there are 70 people altogether,

$$\begin{aligned}\text{More than } 15 \text{ years} &= 70 - 29 \\ &= 41 \text{ people (ie the ones left over).}\end{aligned}$$