Cumulative Frequency

by Sophia

WHAT'S COVERED

This tutorial will cover the topic of cumulative frequency. Our discussion breaks down as follows:

1. Cumulative Frequency

You likely already know about frequency, which refers to how often a data value occurs. Cumulative means the accumulation of everything that has occurred up to a certain point. Therefore, **cumulative frequency** is the collected frequency of data points.

EXAMPLE If a teacher says that a test is cumulative, that means that it's going to cover everything that you've learned that year, up to the point of the test.

In this context, cumulative frequencies involve separating the data into bins and determining how many observations fall within or below that bin.

EXAMPLE This is the distribution of temperatures by 10's for Chanhassen, Minnesota in the year 2009. Three days were between -10°F and -1°F, eight days that were between 0°F and 9°F for the high temperature, and so forth.

Temperature	Frequency
-10 - -1	3
0 - 9	8
10 - 19	25
20 - 29	39
30 - 39	30

40 - 49	51
50 - 59	46
60 - 69	39
70 - 79	80
80 - 89	40
90 - 99	4

With this information about the distribution of temperatures, you can determine cumulative frequencies by asking, “How many days were at or below 9°F for the high temperature?” Well, eight days fell within the zero to nine bin, and three that fell below it. This equals a total of 11.

For the third category, how many days were at or below this 19°F? Well, 25 were in that third bin and 11 were below it, which means it's a total of 36. You can continue this throughout the entire chart. Unsurprisingly, you will get 365 total days. All 365 days of the year were at or below 99 degrees in Chanhassen that year.

Temperature	Frequency	Cumulative Freq
-10 - -1	3	3
0 - 9	8	11
10 - 19	25	36
20 - 29	39	75
30 - 39	30	105
40 - 49	51	156
50 - 59	46	202
60 - 69	39	241
70 - 79	80	321
80 - 89	40	361
90 - 99	4	365

TERM TO KNOW

Cumulative Frequency

The number of data points that fall within or below a given bin of data.

2. Relative Cumulative Frequencies

Sometimes it is useful to consider **relative cumulative frequencies**, which is the percent of observations that fall in or below a certain bin.

You may have encountered relative frequency before, but not relative cumulative frequency. Fortunately, it's calculated the same way as relative frequency. To determine the relative cumulative frequency, divide each value by the total number of values.

In the above example, we are considering a full year or 365 days. So we will divide each cumulative frequency by 365 to get the relative cumulative frequency.

In the first bin, there were 3 out of 365 values that fell in this category. This means that 0.008 of the data fell in or below this bucket. Dividing 11 by 365 gives you about 0.03. Continuing on the rest of the chart, we get these values.

Temperature	Frequency	Cumulative Freq	Rel. Cumulative Freq
-10 - -1	3	3	/365 = 0.008
0 - 9	8	11	/365 = 0.030
10 - 19	25	36	/365 = 0.099
20 - 29	39	75	/365 = 0.205
30 - 39	30	105	/365 = 0.288
40 - 49	51	156	/365 = 0.427
50 - 59	46	202	/365 = 0.533
60 - 69	39	241	/365 = 0.660
70 - 79	80	321	/365 = 0.879
80 - 89	40	361	/365 = 0.989
90 - 99	4	365	/365 = 1.000

BIG IDEA

Note, the main overarching point here is to divide each cumulative by the total.

TERM TO KNOW

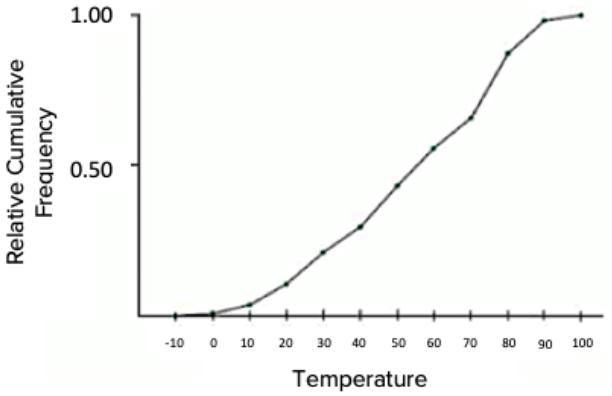
Relative Cumulative Frequency

The percent of data points that fall within or below a given bin of data.

3. Ogives

In the previous chart, you may notice that the final value of 1.000 means that 100% of the values, or all 365 days, fell at or below this bin. Graphically, this information can be presented in something called an ogive. It's also called a relative cumulative frequency graph, or sometimes a percentile graph. It's a line chart that uses these bins and the relative cumulative frequencies to show how many values were at or below these bins.

Temperature	Rel. Cum. Freq
-10 - -1	0.008
0 - 9	0.030
10 - 19	0.099
20 - 29	0.205
30 - 39	0.288
40 - 49	0.427
50 - 59	0.533
60 - 69	0.660
70 - 79	0.879
80 - 89	0.989
90 - 99	1.000



Use the left-hand edge of the bin, because by the time you've gotten to negative 10 degrees going left to right on this number line, you haven't encountered any of the days of the year yet. However, once you get to zero degrees, you've encountered three of the days, which is a certain amount of relative cumulative frequency. By the time you get to 100 degrees, you will have encountered every single day. Every day will have been at some point in or below that bin.

HINT

Ogives are increasing from left to right. If there's no data in a particular bucket, you get a flat line, or no increase.

SUMMARY

Cumulative frequency and relative cumulative frequency show the number or the percent, respectively, of the data that falls in or below a certain bin. We might also refer to this as an ogive. It is a useful way to show how certain values relate to other values or how they relate to the whole. For instance, is a day that is 70 degrees in Chanhassen considered a very hot day? How does it compare to the rest of the days of the year? The relative cumulative frequency can answer those questions.

Good luck!

Source: THIS TUTORIAL WAS AUTHORED BY JONATHAN OSTERS FOR SOPHIA LEARNING. PLEASE SEE OUR TERMS OF USE.

TERMS TO KNOW

Cumulative Frequency

The number of data points that fall within or below a given bin of data.

Relative Cumulative Frequency

The percent of data points that fall within or below a given bin of data.