Sampling and Data: Frequency, Relative Frequency, and Cumulative Frequency[*]

Susan Dean

Barbara Illowsky, Ph.D.

This work is produced by The Connexions Project and licensed under the
Creative Commons Attribution License [†]

Abstract

This module introduces the concepts of frequency, relative frequency, and cumulative relative frequency, and the relationship between these measures. Students will have the opportunity to interpret data through the sample problems provided.

Twenty students were asked how many hours they worked per day. Their responses, in hours, are listed below:

5; 6; 3; 3; 2; 4; 7; 5; 2; 3; 5; 6; 5; 4; 4; 3; 5; 2; 5; 3

Below is a frequency table listing the different data values in ascending order and their frequencies.

Frequency Table of Student Work Hours

DATA VALUE	FREQUENCY
2	3
3	5
4	3
5	6
6	2
7	1

Table 1

A **frequency** is the number of times a given datum occurs in a data set. According to the table above, there are three students who work 2 hours, five students who work 3 hours, etc. The total of the frequency column, 20, represents the total number of students included in the sample.

A **relative frequency** is the fraction of times an answer occurs. To find the relative frequencies, divide each frequency by the total number of students in the sample - in this case, 20. Relative frequencies can be written as fractions, percents, or decimals.

---

[*]Version 1.17: Aug 24, 2010 7:00 pm GMT-5

[†]http://creativecommons.org/licenses/by/2.0/

http://cnx.org/content/m16012/1.17/

Frequency Table of Student Work Hours w/ Relative Frequency

DATA VALUE	FREQUENCY	RELATIVE FREQUENCY
2	3	$\frac{3}{20}$ or 0.15
3	5	$\frac{5}{20}$ or 0.25
4	3	$\frac{3}{20}$ or 0.15
5	6	$\frac{6}{20}$ or 0.30
6	2	$\frac{2}{20}$ or 0.10
7	1	$\frac{1}{20}$ or 0.05

Table 2

The sum of the relative frequency column is $\frac{20}{20}$, or 1.

Cumulative relative frequency is the accumulation of the previous relative frequencies. To find the cumulative relative frequencies, add all the previous relative frequencies to the relative frequency for the current row.

Frequency Table of Student Work Hours w/ Relative and Cumulative Relative Frequency

DATA VALUE	FREQUENCY	RELATIVE FREQUENCY	CUMULATIVE RELATIVE FREQUENCY
2	3	$\frac{3}{20}$ or 0.15	0.15
3	5	$\frac{5}{20}$ or 0.25	$0.15 + 0.25 = 0.40$
4	3	$\frac{3}{20}$ or 0.15	$0.40 + 0.15 = 0.55$
5	6	$\frac{6}{20}$ or 0.30	$0.55 + 0.30 = 0.85$
6	2	$\frac{2}{20}$ or 0.10	$0.85 + 0.10 = 0.95$
7	1	$\frac{1}{20}$ or 0.05	$0.95 + 0.05 = 1.00$

Table 3

The last entry of the cumulative relative frequency column is one, indicating that one hundred percent of the data has been accumulated.

NOTE: Because of rounding, the relative frequency column may not always sum to one and the last entry in the cumulative relative frequency column may not be one. However, they each should be close to one.

The following table represents the heights, in inches, of a sample of 100 male semiprofessional soccer players.

Frequency Table of Soccer Player Height

HEIGHTS (INCHES)	FREQUENCY OF STUDENTS	RELATIVE FREQUENCY	CUMULATIVE RELATIVE FREQUENCY
59.95 - 61.95	5	$\frac{5}{100} = 0.05$	0.05
61.95 - 63.95	3	$\frac{3}{100} = 0.03$	$0.05 + 0.03 = 0.08$
63.95 - 65.95	15	$\frac{15}{100} = 0.15$	$0.08 + 0.15 = 0.23$
65.95 - 67.95	40	$\frac{40}{100} = 0.40$	$0.23 + 0.40 = 0.63$
67.95 - 69.95	17	$\frac{17}{100} = 0.17$	$0.63 + 0.17 = 0.80$
69.95 - 71.95	12	$\frac{12}{100} = 0.12$	$0.80 + 0.12 = 0.92$
71.95 - 73.95	7	$\frac{7}{100} = 0.07$	$0.92 + 0.07 = 0.99$
73.95 - 75.95	1	$\frac{1}{100} = 0.01$	$0.99 + 0.01 = 1.00$
	Total = 100	**Total = 1.00**	

Table 4

The data in this table has been **grouped** into the following intervals:

- 59.95 - 61.95 inches
- 61.95 - 63.95 inches
- 63.95 - 65.95 inches
- 65.95 - 67.95 inches
- 67.95 - 69.95 inches
- 69.95 - 71.95 inches
- 71.95 - 73.95 inches
- 73.95 - 75.95 inches

NOTE: This example is used again in the Descriptive Statistics[1] chapter, where the method used to compute the intervals will be explained.

In this sample, there are **5** players whose heights are between 59.95 - 61.95 inches, **3** players whose heights fall within the interval 61.95 - 63.95 inches, **15** players whose heights fall within the interval 63.95 - 65.95 inches, **40** players whose heights fall within the interval 65.95 - 67.95 inches, **17** players whose heights fall within the interval 67.95 - 69.95 inches, **12** players whose heights fall within the interval 69.95 - 71.95, 7 players whose height falls within the interval 71.95 - 73.95, and **1** player whose height falls within the interval 73.95 - 75.95. All heights fall between the endpoints of an interval and not at the endpoints.

Example 1
From the table, find the percentage of heights that are less than 65.95 inches.

Solution
If you look at the first, second, and third rows, the heights are all less than 65.95 inches. There are $5 + 3 + 15 = 23$ males whose heights are less than 65.95 inches. The percentage of heights less than 65.95 inches is then $\frac{23}{100}$ or 23%. This percentage is the cumulative relative frequency entry in the third row.

---

[1]"Descriptive Statistics: Introduction" <http://cnx.org/content/m16300/latest/>

Example 2

From the table, find the percentage of heights that fall between 61.95 and 65.95 inches.

Solution

Add the relative frequencies in the second and third rows: 0.03 + 0.15 = 0.18 or 18%.

Example 3

Use the table of heights of the 100 male semiprofessional soccer players. Fill in the blanks and check your answers.

1. The percentage of heights that are from 67.95 to 71.95 inches is:
2. The percentage of heights that are from 67.95 to 73.95 inches is:
3. The percentage of heights that are more than 65.95 inches is:
4. The number of players in the sample who are between 61.95 and 71.95 inches tall is:
5. What kind of data are the heights?
6. Describe how you could gather this data (the heights) so that the data are characteristic of all male semiprofessional soccer players.

Remember, you **count frequencies**. To find the relative frequency, divide the frequency by the total number of data values. To find the cumulative relative frequency, add all of the previous relative frequencies to the relative frequency for the current row.

1 Optional Collaborative Classroom Exercise

Exercise 1

In your class, have someone conduct a survey of the number of siblings (brothers and sisters) each student has. Create a frequency table. Add to it a relative frequency column and a cumulative relative frequency column. Answer the following questions:

1. What percentage of the students in your class has 0 siblings?
2. What percentage of the students has from 1 to 3 siblings?
3. What percentage of the students has fewer than 3 siblings?

Example 4

Nineteen people were asked how many miles, to the nearest mile they commute to work each day. The data are as follows:

2; 5; 7; 3; 2; 10; 18; 15; 20; 7; 10; 18; 5; 12; 13; 12; 4; 5; 10

The following table was produced:

Frequency of Commuting Distances

DATA	FREQUENCY	RELATIVE FRE-QUENCY	CUMULATIVE RELATIVE FRE-QUENCY
			continued on next page

3	3	$\frac{3}{19}$	0.1579
4	1	$\frac{1}{19}$	0.2105
5	3	$\frac{3}{19}$	0.1579
7	2	$\frac{2}{19}$	0.2632
10	3	$\frac{4}{19}$	0.4737
12	2	$\frac{2}{19}$	0.7895
13	1	$\frac{1}{19}$	0.8421
15	1	$\frac{1}{19}$	0.8948
18	1	$\frac{1}{19}$	0.9474
20	1	$\frac{1}{19}$	1.0000

Table 5

Problem   *(Solution on p. 6.)*

1. Is the table correct? If it is not correct, what is wrong?
2. True or False: Three percent of the people surveyed commute 3 miles. If the statement is not correct, what should it be? If the table is incorrect, make the corrections.
3. What fraction of the people surveyed commute 5 or 7 miles?
4. What fraction of the people surveyed commute 12 miles or more? Less than 12 miles? Between 5 and 13 miles (does not include 5 and 13 miles)?

Solutions to Exercises in this Module

Solution to Example 3, Problem (p. 4)

1. 29%
2. 36%
3. 77%
4. 87
5. quantitative continuous
6. get rosters from each team and choose a simple random sample from each

Solution to Example 4, Problem (p. 5)

1. No. Frequency column sums to 18, not 19. Not all cumulative relative frequencies are correct.
2. False. Frequency for 3 miles should be 1; for 2 miles (left out), 2. Cumulative relative frequency column should read: 0.1052, 0.1579, 0.2105, 0.3684, 0.4737, 0.6316, 0.7368, 0.7895, 0.8421, 0.9474, 1.
3. $\frac{5}{19}$
4. $\frac{7}{19}$, $\frac{12}{19}$, $\frac{7}{19}$

Glossary

Definition 1: Frequency
 The number of times a value of the data occurs.

Definition 2: Relative Frequency
 The ratio of the number of times a value of the data occurs in the set of all outcomes to the number of all outcomes.

Definition 3: Cumulative Relative Frequency
 The term applies to an ordered set of observations from smallest to largest. The Cumulative Relative Frequency is the sum of the relative frequencies for all values that are less than or equal to the given value.