

CHAPTER 3

THE FREQUENCY DISTRIBUTION

OBJECTIVES

After completing this chapter, you should

- be able to understand, construct, and use a frequency distribution.

CHAPTER REVIEW

A first approach to organizing data is to arrange the numbers in order of value. By convention, we list the highest number first, followed by the next highest, and so on. The data can be condensed slightly by constructing a *frequency distribution* in which the scores are ranked from highest to lowest and the number of times each score occurs (its *frequency*) is listed beside it. Scores are symbolized by the letter X , whereas frequencies appear under the letter f .

A *continuous variable* is one whose measurement can take an infinite number of values; a *discrete variable* is one capable of assuming only specific values. Gaps in the measurement of a continuous variable are apparent rather than real; that is, the gaps are an artifact of the crudeness of measurement. Data from a continuous variable presented in the form of whole numbers have gaps between the numbers; these numbers with gaps are called *apparent limits*. *Real limits* close the gaps by subtracting a half unit from the lower apparent limit and adding a half unit to the upper apparent limit.

To compare frequency distributions from samples varying in size, we need to convert the frequencies to percentages; that is, we need to put them on an equivalent basis. The resulting distribution is called a *percentage* or *relative frequency distribution*. A *cumulative frequency distribution* is constructed by starting with the distribution's lowest interval and accumulating frequencies as you ascend. Thus, for any interval, the cumulative frequency tells the number of scores contained in that interval plus the sum of the

frequencies in intervals below the one considered. Converting each of the accumulated frequencies to a percentage results in a *cumulative percentage distribution*.

SYMBOLS

<i>Symbol</i>	<i>Stands For</i>
X	a score
f	frequency of a score
N	sum of the frequencies, or total sample size
$\text{Cum } f$	cumulative frequency

FORMULAS

Formula 3-1. Equation to convert frequencies into percentages

$$\%age \text{ (percentage)} = \frac{f}{N}(100)$$

f is the frequency or the number of times each score occurs; N is the number of observations or the sum of the frequency column in a frequency distribution.

Formula 3-2. Equation to convert cumulative frequencies to cumulative percentages

$$\text{Cum } \%age = \frac{\text{Cum } f}{N}(100)$$

$\text{Cum } f$ stands for the cumulative frequency, which is the total number of observations up to and including the observations in the interval you're considering.

TERMS TO DEFINE AND/OR IDENTIFY

score

frequency distribution

frequency

continuous variable

discrete variable

apparent limits

real limits

percentage frequencies

cumulative frequency
distribution

cumulative percentage

cumulative percentage
distribution

FILL-IN-THE-BLANK ITEMS

Introduction

The first step in organizing some data might be to arrange the scores by order of value, listing them from the (1) _____ to the (2) _____. The numbers constituting the data are called (3) _____ and are symbolized by the letter (4) _____.

Defining the Frequency Distribution

An arrangement in which the scores are listed in descending order and the number of times each score occurs is listed beside it is called a (5) _____. The number of times each score occurs is called its (6) _____ and is symbolized by the letter (7) _____. In order to further condense the data, scores occurring with a zero frequency are often (8) _____ in constructing the frequency distribution.

Continuous variables and discrete variables: Real limits and apparent limits

A (9) _____ variable is one whose measurement can take an infinite number of values; a variable that can take only specific values is called a (10) _____ variable. Data from a continuous variable presented as whole numbers have gaps between the numbers, resulting in (11) _____ limits. Closing the gaps by subtracting (12) _____ a unit from the lower limit and adding (13) _____ a unit to the upper limit results in (14) _____ limits.

Percentage or Relative Frequency and Cumulative Frequency Distributions

One way to compare frequency distributions from samples of unequal size is to convert the frequencies to (15) _____ frequencies. To do this, you divide each frequency by (16) _____ and multiply the result by 100. N stands for the (17) _____ of the frequencies or the total sample (18) _____.

To construct a cumulative frequency distribution, start with the distribution's lowest interval and (19) _____ frequencies as you ascend. For any interval, the cumulative frequency tells you the number of scores in the interval plus the sum of the frequencies in all (20) _____ intervals. Cumulative frequency is symbolized by (21) _____.

PROBLEMS

1. A group of students is given four brief presentations of a list of 15 nouns. A recognition test is given after a 20-minute delay. The number of nouns correctly identified by each student is listed next. Arrange them into a frequency distribution. Include a Cum f column and a Cum %age column.

6	15	12
8	15	14
12	13	13
15	15	13
13	10	15

2. Another group of students is given similar exposure to the same list of 15 nouns but is given a recall test rather than a recognition test of retention. The resulting scores are listed here. Construct a frequency distribution. Why would it be unnecessary to convert the frequencies in Problems 1 and 2 to percentages if you wanted to compare the two distributions?

0	1	2
1	1	4
0	2	2
0	1	1
3	2	1

3. To test their reaction times, 50 students from an introductory psychology class are asked to press a button when a green light flashes on. Their reaction times in hundredths of a second are as follows:

25	33	28	22	18
31	23	29	25	15
17	21	27	37	16
19	19	25	30	25
15	18	18	26	24
16	21	17	21	23
16	20	15	15	23
17	17	14	16	25
25	15	20	17	26
24	12	21	19	20

- a. Construct a frequency distribution.

- b. Add the real limits and a %age frequency column to your frequency distribution.

4. Find the real limits for each of the following scores:

	Apparent Limits	Real Limits
a.	25	
b.	11.7	
c.	12.55	
d.	7.853	

5. A random sample of 20 depressive patients has been given the MMPI. Their scores on Scale 2 (Depression Scale) are as follows:

30	25	26	31
45	37	35	36
32	32	30	26
28	34	34	42
33	32	35	39

Construct a frequency distribution. Include a Cum f column and a %age f column.

6. For practice, some graduate students in a personality appraisal course administer the MMPI to 85 of their friends. Their scores on Scale 4 (Psychopathic Deviate) are shown in frequency distribution A as follows. Distribution B shows the Scale 4 results from 32 participants in a DWI (Driving While Intoxicated) class. What would you have to do to the frequency distributions in order to compare them? Add the appropriate column to each distribution.

<u>Distribution A</u>		<u>Distribution B</u>	
<u>X</u>	<u>f</u>	<u>X</u>	<u>f</u>
79	1	79	1
77	1	78	2
76	1	77	2
74	2	76	2
65	2	75	3
64	3	74	2
62	2	73	2
60	2	71	1
57	5	70	2
56	4	69	3
54	6	68	4
53	7	65	1
52	6	60	1
51	7	58	2
50	10	55	1
49	7	50	1
48	6	44	1
47	4	39	1
45	2		
44	1		
42	3		
40	2		
39	1		

7. A large number of students took a 50-item test in a section of introductory psychology. A frequency distribution of the results is shown here. Show the real limits, cumulative frequencies, and cumulative percentages. What is N ?

X	f	X	f
49	1	35	8
48	6	34	3
47	6	33	1
46	11	32	7
45	13	31	3
44	16	30	4
43	8	29	3
42	15	28	4
41	16	27	1
40	13	26	3
39	9	25	2
38	5	23	1
37	3		
36	7		

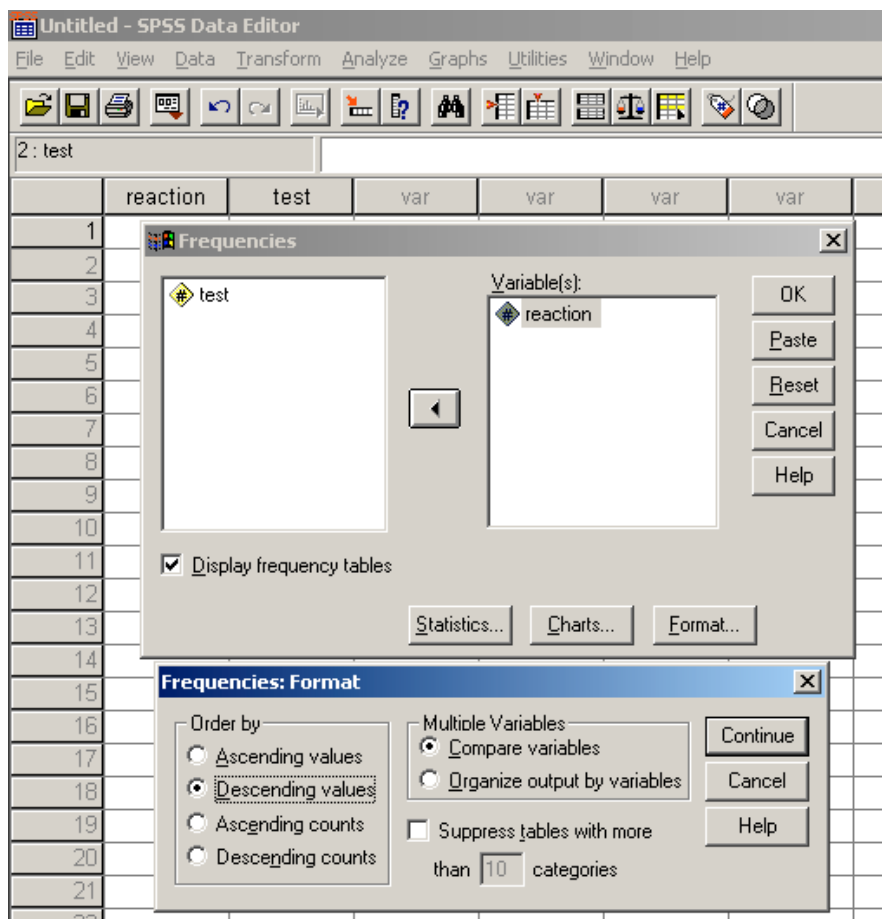
USING SPSS—EXAMPLE AND EXERCISE

If you do not already know how to use SPSS, read and work through Appendix 4 before proceeding with the exercise.

The SPSS FREQUENCIES procedure can be used to produce frequency distributions with percentage and cumulative percentage columns in either an ascending or descending arrangement.

Example: We will use SPSS to produce a frequency distribution for Problem 3. The steps are as follows:

1. Start SPSS, enter the data, and name the variable **reaction**.
2. Click *Analyze>Descriptive Statistics>Frequencies*.
3. Highlight **reaction** and move it into the Variable(s) box.
4. To produce a descending arrangement, click *Format>Order by Descending values>Continue*. The Frequencies dialog box and the Frequencies: Format dialog box should appear as follows:



5. Click OK.

The output showing the complete frequency distribution for Problem 3 should appear as follows. Note: If you are using the student version of SPSS, the syntax (immediately below) may not be available.

```
FREQUENCIES
  VARIABLES=reaction
  /FORMAT=DVALUE
  /ORDER ANALYSIS .
```

Frequencies

Statistics

REACTION

N	Valid	50
	Missing	0

REACTION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	37.00	1	2.0	2.0	2.0
	33.00	1	2.0	2.0	4.0
	31.00	1	2.0	2.0	6.0
	30.00	1	2.0	2.0	8.0
	29.00	1	2.0	2.0	10.0
	28.00	1	2.0	2.0	12.0
	27.00	1	2.0	2.0	14.0
	26.00	2	4.0	4.0	18.0
	25.00	6	12.0	12.0	30.0
	24.00	2	4.0	4.0	34.0
	23.00	3	6.0	6.0	40.0
	22.00	1	2.0	2.0	42.0
	21.00	4	8.0	8.0	50.0
	20.00	3	6.0	6.0	56.0
	19.00	3	6.0	6.0	62.0
	18.00	3	6.0	6.0	68.0
	17.00	5	10.0	10.0	78.0
	16.00	4	8.0	8.0	86.0
	15.00	5	10.0	10.0	96.0
	14.00	1	2.0	2.0	98.0
	12.00	1	2.0	2.0	100.0
Total		50	100.0	100.0	

Note that the cumulative percentage column accumulates the percentages in ascending rather than descending order, which is the opposite of the way we taught the procedure in the textbook. In fact, it may be better for you to leave the frequency distribution format in ascending order (see step 4), because the cumulative percent column in the output is always ascending. If you do it this way, the SPSS frequency distributions will be perfectly inverted for all columns and will be easier to compare with the text. In a few cases, particularly in this chapter and the next one on graphing, SPSS does not have enough flexibility to produce the output that we described in the textbook. Because SPSS doesn't have the flexibility, we will have to be flexible for it.

Exercise Using SPSS

1. Use the data in the second column (scores from 35 to 23) from Problem 7 to construct a descending frequency distribution using SPSS. Note that you will need to enter each score as many times as indicated by the frequency. For example, 35 would be entered 8 times.

CHECKING YOUR PROGRESS: A SELF-TEST

1. Match the following:

- | | |
|------------------------------|---|
| _____ X | a. ranking of scores from highest to lowest with the number of times each score occurs listed beside it |
| _____ f | b. number of observations; sample size |
| _____ continuous variable | c. stands for scores |
| _____ frequency distribution | d. number of score units in a class interval |
| _____ real limits | e. number of scores through a given score or interval |
| _____ N | f. difference between highest and lowest scores |
| _____ cumulative frequency | g. variable capable of assuming only specific values |
| | h. variable whose measurement can take an infinite number of values |
| | i. formed by adding or subtracting half a unit |
| | j. number of occurrences of a score |

2. Rounded to the nearest whole dollar, the amount of cash each student in a class had on a given day is shown here. Make a frequency distribution of these data. Give real limits and include a cumulative frequency column and a cumulative percentage column.

37	9	0	61	3	6
15	22	12	65	5	6
10	8	81	65	7	8
10	10	75	12	12	10
0	13	93	8	10	5
8	17	52	21	9	3
3	15	71	32	5	2