Chapter 2: Frequency Distributions

Chapter Outline

2.1 Introduction to Frequency Distributions
2.2 Frequency Distribution Tables
    Obtaining ΣX from a Frequency Distribution Table
    Proportions and Percentages
    Grouped Frequency Distribution Tables
    Real Limits and Frequency Distributions
2.3 Frequency Distribution Graphs
    Graphs for Interval or Ratio Data (Histograms and Polygons)
    Graphs for Nominal or Ordinal Data (Bar Graphs)
    Graphs for Population Distributions (Relative Frequencies and Smooth Curves)
2.4 The Shape of a Frequency Distribution
2.5 Percentiles, Percentile Ranks, and Interpolation
    Cumulative Frequency and Cumulative Percentage
    Interpolation
2.6 Stem and Leaf Displays
    Comparing Stem and Leaf Displays with Frequency Distributions

Learning Objectives and Chapter Summary

1. Students should understand the concept of a frequency distribution as an organized display showing where all of the individual scores are located on the scale of measurement.

    Note that one goal of descriptive statistics is to organize research results so that researchers can see what happened. Also note that a frequency distribution does not simply summarize the scores, but rather shows the entire set of scores.

2. Students should be able to organize data into a regular or a grouped frequency distribution table, and understand data that are presented in a table.

    If scores are presented in a regular table, students should be able to retrieve the complete list of original scores.

    The purpose for a grouped table is to keep the presentation relatively simple and easy to understand. All of the guidelines for constructing a grouped table are intended to help make the result easy and simple. Note, however, that after the scores have been put into a grouped table, the individual score values are lost.

3. Students should be able to organize data into frequency distribution graphs, including bar graphs, histograms, and polygons. Also, students should be able to understand data that are presented in a graph.

Bar graphs (space between bars) are used to display data from nominal and ordinal scales. Polygons and histograms are used for data from interval or ratio scales. If scores are presented in a frequency distribution graph, students should be able to retrieve the complete list of original scores.

4. Students should understand that most population distributions are drawn as smooth curves showing relative proportions rather than absolute frequencies.

5. Students should be able to identify the shape of a distribution shown in a frequency distribution graph. Students should recognize symmetrical distributions (including but not limited to normal distributions), as well as positively and negatively skewed distributions.

6. Students should be able to describe locations within a distribution using percentiles and percentile ranks, and they should be able to compute percentiles and ranks using interpolation when necessary.

The first key to determining percentiles and percentile ranks is the idea that all cumulative values (both frequencies and percentages) correspond to the upper real limit of each interval.

The process of interpolation is based on two concepts:
1) Each interval is defined in terms of two different scales: scores and percentages. In Example 2.7, for example, one interval extends from $X = 4.5$ to $X = 9.5$ in terms of scores, and the same interval extends from 10% to 60% in terms of percentages.
2) A fraction of the interval on one scale corresponds to exactly the same fraction of the interval on the other scale. For example, a score of $X = 7$ is exactly half-way between 4.5 and 9.5, and the corresponding value of 35% is exactly half-way between 10% and 60%.

Other Lecture Suggestions

1. Begin with an unorganized list of scores as in Example 2.1, and then organize the scores into a table. If you use a set of 20 or 25 scores, it will be easy to compute proportions and percentages for the same example.

2. Present a relatively simple, regular frequency distribution table (for example, use scores of 5, 4, 3, 2, and 1 with corresponding frequencies of 1, 3, 5, 3, 2. Ask the students to determine the values of N and ΣX for the scores. Note that ΣX can be obtained two different ways: 1) by computing and summing the fX values within the table, 2) by retrieving the complete list of individual scores and working outside the table.

Next, ask the students to determine the value of ΣX^2. You probably will find a lot of wrong answers from students who are trying to use the fX values within the table. The common mistake is to compute $(fX)^2$ and then sum these values. Note that whenever it is necessary to do complex calculations with a set of scores, the safe method is to retrieve the list of individual scores from the table before you try any computations.

3. It sometimes helps to make a distinction between graphs that are being used in a formal presentation and sketches that are used to get a quick overview of a set of data. In one case, the graphs should be drawn precisely and the axes should be labeled clearly so that the graph can be easily understood without any outside explanation. On the other hand, a sketch that is intended for your own personal use can be much less precise. As an instructor, if you are expecting precise, detailed graphs from your students, you should be sure that they know your expectations.

4. Introduce interpolation with a simple, real-world example. For example, in Buffalo, the average snowfall during the month of February is 30 inches. Ask students, how much snow they would expect during the first half of the month. Then point out that the same interval (February) is being measured in terms of days and in terms of inches of snow. A point that is half-way through the interval in terms of days should also be half-way through the interval in terms of snow.

Exam Items for Chapter 2

Multiple-Choice Questions

1. What is the total number of scores for the distribution shown in the following table?

a. 4
b. 10
c. 14
d. 37

X	f
4	3
3	5
2	4
1	2

2. A sample of n = 15 scores ranges from a high of X = 11 to a low of X = 3. If these scores are placed in a frequency distribution table, how many X values will be listed in the first column?

a. 8
b. 9
c. 11
d. 15

3. For the following frequency distribution of quiz scores, how many individuals took the quiz?

a. n = 5
b. n = 15
c. n = 21
d. cannot be determined

X	f
5	6
4	5
3	5
2	3
1	2

4. (www) For the following distribution of quiz scores, if a score of X = 3 or higher is needed for a passing grade, how many individuals passed?

a. 3
b. 11
c. 16
d. cannot be determined

X	f
5	6
4	5
3	5
2	3
1	2

5. For the following distribution of quiz scores, How many individuals had a score of X = 2?

a. 1
b. 3
c. 5
d. cannot be determined

X	f
5	6
4	5
3	5
2	3
1	2

6. For the following frequency distribution of exam scores, what is the lowest score on the exam?

a. X = 70
b. X = 74
c. X = 90
d. cannot be determined

X	f
90-94	3
85-89	4
80-84	5
75-79	2
70-74	1

7. For the following frequency distribution of exam scores, how many students had scores lower than X = 80?

a. 2
b. 3
c. 7
d. cannot be determined

X	f
90-94	3
85-89	4
80-84	5
75-79	2
70-74	1

8. In a grouped frequency distribution one interval is listed as 50-59. Assuming that the scores are measuring a continuous variable, what are the real limits of this interval?

a. 50 and 59
b. 50.5 and 59.5
c. 49.5 and 59.5
d. 49.5 and 60.5

9. For the following distribution, how many people had scores less than X = 19?

a. 5
b. 10
c. 11
d. cannot be determined

X	f
20-25	2
15-19	5
10-14	4
5-9	1

10. (www) For the following distribution, what is the highest score?

a. 5
b. 20
c. 25
d. cannot be determined

X	f
20-25	2
15-19	5
10-14	4
5-9	1

11. For the following distribution, how many people had scores greater than X = 14?

a. 5
b. 7
c. 11
d. cannot be determined

X	f
20-25	2
15-19	5
10-14	4
5-9	1

12. (www) For the following distribution, what is the width of each class interval?

a. 4
b. 4.5
c. 5
d. 10

X	f
20-24	2
15-19	5
10-14	4
5-9	1

13. If the following distribution was shown in a histogram, the bar above the 15-19 interval would reach from _____ to _____.

a. X = 14.5 to X = 19.5
b. X = 15.5 to X = 18.5
c. X = 15.5 to X = 19.5
d. X = 15.0 to X = 19.0

X	f
20-25	2
15-19	5
10-14	4
5-9	1

14. (www) In a frequency distribution graph, frequencies are presented on the ___ and the scores (categories) are listed on the ___.

a. X axis/Y axis
b. horizontal line/vertical line
c. Y axis/X axis
d. class interval/horizontal line

15. What frequency distribution graph is appropriate for scores measured on a nominal scale?

a. only a histogram
b. only a polygon
c. either a histogram or a polygon
d. only a bar graph

16. The classrooms in the Psychology department are numbered from 100 to 108. A professor records the number of classes held in each room during the fall semester. If these values are presented in a frequency distribution graph, what kind of graph would be appropriate?

a. a histogram
b. a polygon
c. a histogram or a polygon
d. a bar graph

17. A researcher records the number of traffic tickets issued in each county along the New York State thruway. If the results are presented in a frequency distribution graph, what kind of graph should be used?

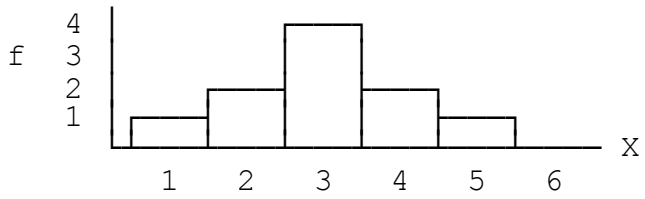
a. a bar graph
b. a histogram
c. a polygon
d. either a histogram or a polygon

18. What kind of frequency distribution graph shows the frequencies as bars, with no space between adjacent bars?

a. a bar graph
b. a histogram
c. a polygon
d. all of the above

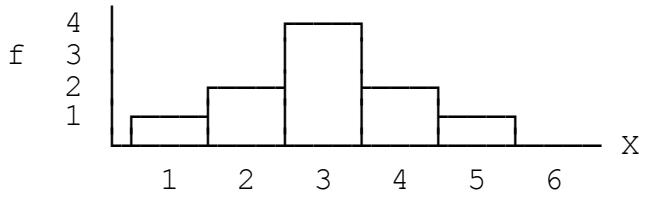
19. For the distribution in the following graph, what is the value of ΣX?

a. 15
b. 21
c. 30
d. cannot determine



20. What scale of measurement was used to measure the scores in the distribution shown in the following graph?

a. nominal
b. ordinal
c. interval or ratio
d. cannot determine



21. What kind of frequency distribution graph shows the frequencies as bars that are separated by spaces?

a. a bar graph
b. a histogram
c. a polygon
d. all of the above

22. (www) If a frequency distribution is shown in a bar graph, what scale was used to measure the scores?

a. nominal
b. nominal or ordinal
c. ratio
d. interval or ratio

23. The normal distribution is an example of

a. a histogram showing data from a sample
b. a polygon showing data from a sample
c. a bar graph showing data from a population
d. a smooth curve showing data from a population

24. If a set of exam scores forms a symmetrical distribution, what can you conclude about the students scores?

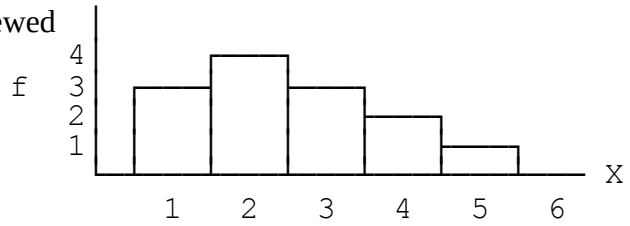
a. Most of the students had relatively high scores.
b. Most of the students had relatively low scores.
c. About 50% of the students had high scores and the rest had low scores.
d. It is not possible the draw any conclusions about the students' scores.

25. What term is used to describe the shape of a distribution in which the scores pile up on the left-hand side of the graph and taper off to the right?

a. symmetrical
b. positively skewed
c. negatively skewed
d. normal

26. What is the shape for the distribution shown in the following graph?

a. positively skewed
b. negatively skewed
c. symmetrical
d. normal



27. A skewed distribution typically has _____ tail(s) and a normal distribution has ____ tail(s).

a. 1, 1
b. 1, 2
c. 2, 1
d. 2, 2

28. (www) The students in a psychology class seemed to think that the midterm exam was very easy. If they are correct, what is the most likely shape for the distribution of exam scores?

a. symmetrical
b. positively skewed
c. negatively skewed
d. normal

29. In a distribution with positive skew, scores with the highest frequencies are _____.

a. on the right side of the distribution
b. on the left side of the distribution
c. in the middle of the distribution
d. represented at two distinct peaks

30. What is the shape of the distribution for the following set of data?
Scores: 1, 2, 3, 3, 4, 4, 4 5, 5, 5, 5, 6

a. symmetrical
b. positively skewed
c. negatively skewed
d. cumulative

31. (www) For the distribution in the following table, what is the 50th percentile?

a. X = 8
b. X = 7.5
c. X = 7
d. X = 6.5

X	c%
9	100%
8	80%
7	50%
6	25%

32. (www) For the distribution in the following table, what is the percentile rank for X = 8.5?

a. X = 90%
b. X = 80%
c. X = 65%
d. X = 50%

X	c%
9	100%
8	80%
7	50%
6	25%

33. (www) For the distribution in the following table, what is the 90th percentile

a. X = 9.5
b. X = 9
c. X = 8.5
d. X = 8

X	c%
9	100%
8	80%
7	50%
6	25%

34. (www) For the distribution in the following table, what is the percentile rank for X = 7?

a. X = 80%
b. X = 65%
c. X = 50%
d. X = 37.5%

X	c%
9	100%
8	80%
7	50%
6	25%

35. For the distribution in the following table, what is the 90th percentile?

a. X = 24.5
b. X = 25
c. X = 29
d. X = 29.5

X	c%
30-34	100%
25-29	90%
20-24	60%
15-19	20%

36. For the distribution in the following table, what is the percentile rank for X = 24.5?

a. 40%
b. 60%
c. 75%
d. 90%

X	c%
30-34	100%
25-29	90%
20-24	60%
15-19	20%

37. For the distribution in the following table, what is the 50th percentile?

a. X = 32
b. X = 35
c. X = 37
d. X = 39

X	c%
50-59	100%
40-49	90%
30-39	60%
20-29	20%

38. For the distribution in the following table, what is the percentile rank for X = 32?

a. 92%
b. 92.5%
c. 95%
d. 97.5%

X	c%
30-34	100%
25-29	90%
20-24	60%
15-19	20%

39. For the scores shown in the following stem and leaf display, what is the highest score in the distribution?

a. 8
b. 83
c. 84
d. 7042

8	314
7	945
6	7042
5	68
4	14

40. If the following scores were placed in a stem and leaf display, how many leaves would be associated with a stem of 6?

a. 1
b. 2
c. 3
d. 4

Scores: 26, 45, 62, 11, 21, 55, 66
64, 55, 46, 38, 41, 27, 29
36, 51, 32, 25, 34, 44, 59

True/False Questions

41. A researcher surveys a sample of n = 200 college students and asks each person to identify his or her favorite movie from the past year. If the data were organized in a frequency distribution table, the first column would be a list of movies.

42. A group of quiz scores ranges from 3 to 10, but no student had a score of X = 5. If the scores are put in a frequency distribution table, X = 5 would not be listed in the X column.

43. It is customary to list the score categories in a frequency distribution from the highest down to the lowest.

44. There is a total of n = 5 scores in the distribution shown in the following table.

X	f
5	2
4	8
3	5
2	3
1	2

45. For the following distribution of scores, 20% of the individuals have scores of X = 1.

X	f
5	2
4	8
3	5
2	3
1	2

46. For the following distribution of scores, $\Sigma X = 18$.

X	f
4	1
3	2
2	3
1	2

47. For the following distribution of scores, $\Sigma X^2 = 92$.

X	f
4	1
3	2
2	3
1	2

48. A grouped frequency distribution table lists one interval as, 20-29. The width of this interval is 9 points.

49. In a grouped frequency distribution table, one interval is identified as 30-34. This interval has a width of 5 points.

50. If a set of scores covers a range of 80 points, the grouped frequency table should use an interval width of 8 points.

51. A set of scores ranges from X = 18 to X= 91. If the scores are put in a grouped frequency distribution table with an interval width of 10 points, the top interval would be 91-100.

52. In a grouped frequency distribution table, the top value in each class interval should be a multiple of the interval width.

53. A set of scores ranges from a low of X = 18 to a high of X = 98. If the scores are put in a grouped frequency distribution table with an interval width of 10 points, the bottom interval should be 10-19.

54. A grouped frequency distribution table does not provide enough information to obtain a complete listing of the original set of scores.

55. For the following distribution, seven people have scores greater than X = 14.

X	f
20-24	2
15-19	5
10-14	4
5-9	1

56. In the following distribution, the scores are grouped into class intervals that are each 5 points wide.

X	f
20-24	2
15-19	5
10-14	4
5-9	1

57. A professor records the number of students who are absent each day for the semester. Because this is a discrete variable, a bar graph should be used to show the frequency distribution.

58. A researcher surveys a sample of n = 200 college students and asks each person to identify his or her favorite movie from the past year. If the results are presented in a frequency distribution graph, the researcher should use a bar graph.

59. If it is appropriate to present a distribution of scores in a polygon, then it would also be appropriate to present the scores in a bar graph.

60. A histogram is constructed so that adjacent bars touch.

61. The normal distribution is an example of a symmetrical distribution.

62. In February in New York, the daily high temperatures are typically low with only a few relatively warm days. A frequency distribution showing the daily high temperatures would probably form a negatively skewed distribution.

63. The scores for a very easy exam would probably form a positively skewed distribution.

64. If a set of exam scores forms a negatively skewed distribution, it suggests that the majority of the students did not score well on the exam.

65. A score equal to the 5^{th} percentile is one of the highest scores in the distribution.

66. For the distribution in the following table, the 80^{th} percentile is X = 24.

X	c%
25-29	100%
20-24	80%
15-19	20%

67. For the distribution in the following table, the percentile rank for X = 19.5 is 20%.

X	c%
25-29	100%
20-24	80%
15-19	20%

68. For the distribution in the following table, the 90^{th} percentile is X = 27.5.

X	c%
25-29	100%
20-24	80%
15-19	20%

69. For the distribution in the following table, the percentile rank for X = 25 is 82%.

X	c%
25-29	100%
20-24	80%
15-19	20%

70. A stem and leaf display does not provide enough information to obtain a complete listing of the original set of scores.

Other Exam Items

71. Find each value requested for the set of scores in the following frequency distribution table.
    a. n
    b. ΣX
    c. ΣX^2

X	f
5	1
4	2
3	3
2	5
1	2

72. Briefly explain what information is available in a regular frequency distribution table that is not available in a grouped table.

73. For the following scores
    a. construct a frequency distribution table.
    b. sketch a histogram of the frequency distribution.

    6, 4, 3, 5, 4, 2, 4
    5, 4, 6, 1, 4, 5, 2

74. (www) For the distribution shown in the following table:
    a. Find the percentile rank for X = 14.5.
    b. Find the 60^{th} percentile.
    c. Find the percentile rank for X = 11.
    d. Find the 66^{th} percentile.

X	f	cf	c%
25-29	4	25	100%
20-24	6	21	84%
15-19	7	15	60%
10-14	5	8	32%
5-9	3	3	12

75. Construct a stem and leaf display for the following scores.

    30, 23, 58, 28, 35, 67, 27, 42, 46, 35
    51, 33, 18, 33, 25, 38, 48, 36, 31, 39

Answers for Multiple-Choice Questions (with section and page numbers from the text)

1. c, 2.2, p. 40
2. b, 2.2, p. 42
3. c, 2.2, p. 40
4. c, 2.2, p. 40
5. b, 2.2, p. 40
6. d, 2.2, p. 44
7. b, 2.2, p. 44
8. c, 2.2, p. 44
9. d, 2.2, p. 44
10. d, 2.2, p. 44
11. b, 2.2, p. 44
12. c, 2.2, p. 44
13. a, 2.3, p. 46
14. c, 2.3, p. 45
15. d, 2.3, p. 47
16. d, 2.3, p. 47
17. a, 2.3, p. 47
18. b, 2.3, p. 46
19. c, 2.3, p. 46
20. c, 2.3, p. 45
21. a, 2.3, p. 48
22. b, 2.3, p. 47
23. d, 2.3, p. 49
24. c, 2.4, p. 50
25. b, 2.4, p. 51
26. a, 2.4, p. 51
27. b, 2.4, p. 50
28. c, 2.4, p. 52
29. b, 2.4, p. 51
30. c, 2.4, p. 51
31. b, 2.5, p. 55
32. b, 2.5, p. 55
33. b, 2.5, p. 58
34. d, 2.5, p. 58
35. d, 2.5, p. 55
36. b, 2.5, p. 55
37. c, 2.5, p. 58
38. c, 2.5, p. 58
39. c, 2.6, p. 60
40. c, 2.6, p. 60

Answers for True/False Questions (with section and page numbers from the text)

41. T, 2.2, p. 40
42. F, 2.2, p. 40
43. T, 2.2, p. 39
44. F, 2.2, p. 40
45. F, 2.2, p. 41
46. T, 2.2, p. 40
47. F, 2.2, p. 40
48. F, 2.2, p. 44
49. T, 2.2, p. 44
50. F, 2.2, p. 43
51. F, 2.2, p. 43
52. F, 2.2, p. 43
53. T, 2.2, p. 43
54. T, 2.2, p. 44
55. T, 2.2, p. 44
56. T, 2.2, p. 44
57. F, 2.3, p. 47
58. T, 2.3, p. 47
59. F, 2.3, p. 47
60. T, 2.3, p. 46
61. T, 2.3, p. 50
62. T, 2.4, p. 51
63. F, 2.4, p. 52
64. T, 2.4, p. 52
65. F, 2.5, p. 53
66. F, 2.5, p. 55
67. T, 2.5, p. 55
68. F, 2.5, p. 58
69. T, 2.5, p. 58
70. F, 2.6, p. 60

Answers for Other Exam Items

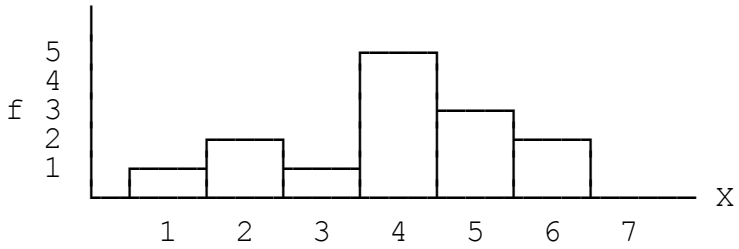
71. a. $n = 13$
    b. $\Sigma X = 34$
    c. $\Sigma X^2 = 106$

72. A regular table identifies each individual score exactly. However, in a grouped table, you simply know that an individual score is located in a particular interval, but you do not know its exact value.

73. a.

X	f
6	2
5	3
4	5
3	1
2	2
1	1

b.



74. a. 32%
b. X = 19.5
c. 18*
d. X = 20.75

75.

6	7
5	18
4	826
3	033586159
2	3857
1	8