Chapter 13

Statistics and Probability

Graphical Representation of Data

OBJECTIVES for Sections 13-1 through 13-3:
1. *To draw and interpret a dot frequency diagram and to determine relative frequencies for a given set of data.*
2. *To draw and interpret a histogram and a frequency polygon for a given set of data.*
3. *To draw and interpret a cumulative frequency polygon for a given set of data.*

13–1 Dot Frequency Diagrams and Relative Frequency

Statistics is the science of organizing and analyzing a set of numerical facts, or *data*, so that probable conclusions can be drawn from the data.

Suppose that 25 thirteen-year-old girls are measured and that their heights to the nearest centimeter are given in the following array:

158	160	155	159	160
162	161	158	163	157
156	165	160	157	165
161	157	163	160	160
158	161	159	161	158

In trying to analyze such a distribution of heights, it may be helpful to display the data as a **dot frequency diagram,** as shown in Figure 1.

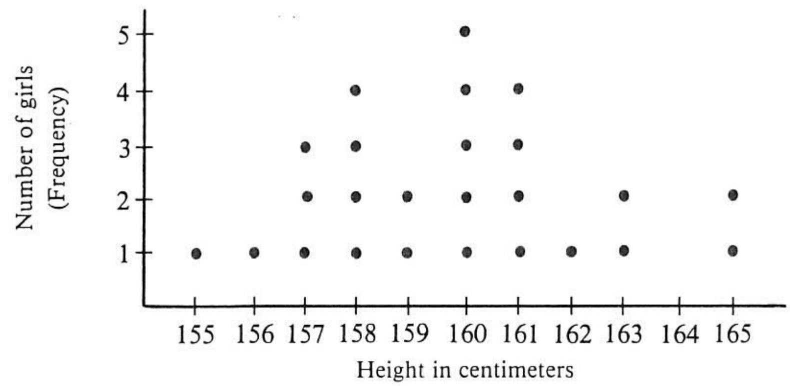


Figure 1

By observing the dots you can quickly see the **frequency,** or number of occurrences, of each measurement. You can also see, for example, that 21 of the 25 students have heights between 157 cm and 163 cm, inclusive.

You can obtain the **relative frequency** of each measurement if you divide its particular frequency by the total number of measurements. Converting these ratios to percents makes the data easier for most people to interpret.

		Relative frequency	
Height	*Frequency*	*Fraction*	*Percent*
155	1	$\frac{1}{25}$	4
156	1	$\frac{1}{25}$	4
157	3	$\frac{3}{25}$	12
158	4	$\frac{4}{25}$	16
159	2	$\frac{2}{25}$	8
160	5	$\frac{5}{25}$	20
161	4	$\frac{4}{25}$	16
162	1	$\frac{1}{25}$	4
163	2	$\frac{2}{25}$	8
165	2	$\frac{2}{25}$	8
Total:	25	$\frac{25}{25} = 1$	100

From this table you can see that the heights of 72%, or about three fourths of the class, are between 157 cm and 161 cm, inclusive.

Oral Exercises

In Exercises 1–3 state the frequency and the relative frequency, as a fraction and a percent, of the given letter from the word REPEATER.

1. T 2. R 3. E

In Exercises 4–8 state the frequency and the relative frequency, as a fraction and a percent, of the given letter from the word INDEPENDENCE.

4. E 5. N 6. P 7. D 8. C

9. Do you think that the set of arm length measurements of all the members of your algebra class would show the same pattern of distribution as the set of heights? That is, would most of the measurements cluster near the middle measurement?
10. Name at least one word in which the letter s has a frequency of 2 and a relative frequency of $\frac{2}{5}$, or 40%.

Written Exercises

Make a dot frequency diagram for the given data. Label the axes "Number" and "Frequency."

A

1.

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

2.

32	30	33	36	30	36
36	34	35	30	30	36
30	31	36	35	32	33

Make a table showing the frequencies and relative frequencies, as fractions and percents, for the data in the given exercise.

3. Exercise 1 4. Exercise 2

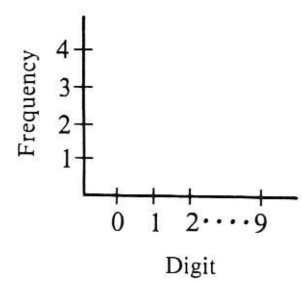
5. Make a dot frequency diagram for the letters in the word ENTERTAINING.
6. Make a dot frequency diagram for the letters in the word MONOTONOUS.

Make a table for the data in the given exercise, showing frequencies and relative frequencies, as fractions and percents.

7. Exercise 5 8. Exercise 6

B

9. In order to express $\frac{1}{17}$ as a decimal, write $\frac{1}{17} = 0.\overline{0588235294117647}$. (Recall that the fraction $\frac{1}{n}$ can have at most $n - 1$ digits in the repeating block of digits in its decimal equivalent.)
   a. Make a table showing the frequency and relative frequency of each of the ten digits.
   b. Make a dot frequency diagram. Label the horizontal axis "Digit" and the vertical axis "Frequency."



10. The birth dates of the first 20 United States presidents are shown.
   a. Make a table showing the frequency and relative frequency, as a fraction and percent, of each month.
   b. Make a dot frequency diagram for the data. Label the horizontal axis "Month" and the vertical axis "Frequency."

No.	*Name*	*Date of Birth*
1	George Washington	Feb. 22, 1732
2	John Adams	Oct. 30, 1735
3	Thomas Jefferson	Apr. 13, 1743
4	James Madison	Mar. 16, 1751
5	James Monroe	Apr. 28, 1753
6	John Quincy Adams	July 11, 1767
7	Andrew Jackson	Mar. 15, 1767
8	Martin Van Buren	Dec. 5, 1782
9	William Henry Harrison	Feb. 9, 1773
10	John Tyler	Mar. 29, 1790
11	James K. Polk	Nov. 2, 1795
12	Zachary Taylor	Nov. 24, 1784
13	Millard Fillmore	Jan. 17, 1800
14	Franklin Pierce	Nov. 23, 1804
15	James Buchanan	Apr. 23, 1791
16	Abraham Lincoln	Feb. 12, 1809
17	Andrew Johnson	Dec. 29, 1808
18	Ulysses S. Grant	Apr. 27, 1822
19	Rutherford B. Hayes	Oct. 4, 1822
20	James A. Garfield	Nov. 19, 1831

C 11. Record the first letter of the last name of all the students in your algebra class. Make a table showing the frequencies and relative frequencies, as fractions and percents, for the data.

13–2 Histograms and Frequency Polygons

For a large set of data with a wide range of values, a dot frequency diagram is not a practical device for visualizing frequencies. Instead you can make a table showing a frequency distribution in which the data are grouped in equal intervals, and the frequency is shown for each interval. From this you can make a type of *bar graph* called a **histogram** to help visualize the distribution.

The following table shows the distribution of the weights in kilograms of 100 high-school sophomore boys.

		Relative frequency	
Interval	*Frequency*	*Fraction*	*Percent*
45–50	2	$\frac{2}{100}$	2
50–55	7	$\frac{7}{100}$	7
55–60	12	$\frac{12}{100}$	12
60–65	28	$\frac{28}{100}$	28
65–70	23	$\frac{23}{100}$	23
70–75	16	$\frac{16}{100}$	16
75–80	8	$\frac{8}{100}$	8
80–85	3	$\frac{3}{100}$	3
85–90	1	$\frac{1}{100}$	1
Total:	100	$\frac{100}{100} = 1$	100

The weights here are grouped into nine intervals, each of length 5. A large collection of data is usually compressed into anywhere from 10 to 20 such *class intervals*, depending on the number and range of the measurements.

A boundary value ordinarily is included in the interval on its *left*. For example, in Figure 2 a weight of 50 kg would be included in the 45–50 interval, *not* in the 50–55 interval.

The histogram of the given distribution (Figure 2) indicates that the greatest clustering of weights occurs between 60 and 70 kg (51%).

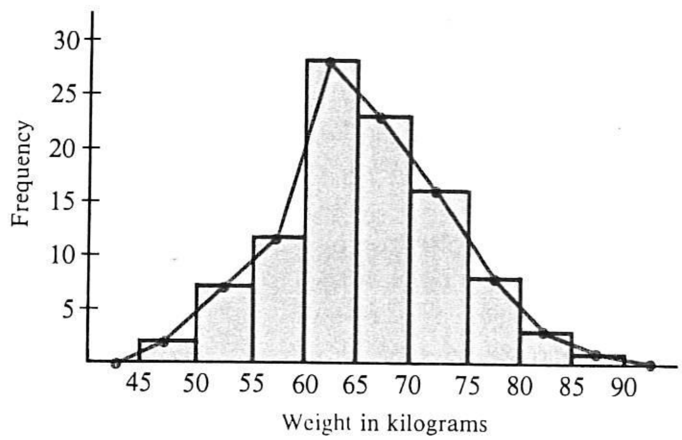


Figure 2

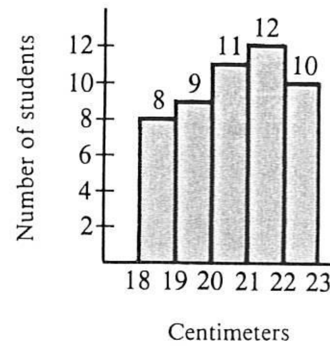
The red broken-line graph joining the midpoints of the intervals is called a **frequency polygon.**

Notice that the frequency polygon extends a half-interval beyond the histogram at each end, starting and ending on the horizontal axis. For a reason why this is done, see Oral Exercises 9 and 10.

A frequency distribution can be displayed by using either a frequency polygon or a histogram.

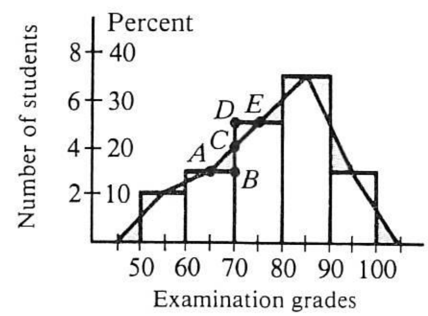
Oral Exercises

The histogram at the right shows a frequency distribution of the head size measurements in two algebra classes.



1. How many had head size measurements between 19 cm and 20 cm?
2. How many measured between 22 cm and 23 cm?
3. How many measured between 20 cm and 22 cm?
4. How many measured greater than 20 cm?
5. How many students were measured?
6. What was the relative frequency (as a fraction) of measurements between 19 cm and 20 cm?
7. What percent of the measurements were between 21 cm and 22 cm?
8. What percent of the measurements were between 18 cm and 20 cm?

Exercises 9 and 10 refer to the figure at the right.



9. Triangles *ABC* and *EDC* are congruent. Do they have equal areas? How many other pairs of triangles in the figure have equal areas?
10. Explain why the area between the horizontal axis and the frequency polygon is equal to the sum of the areas of the rectangles in the histogram.

Written Exercises

A

1. a. Make a frequency distribution table showing the data at the right. Group the data in the intervals 5–35, 35–65, 65–95, and show relative frequencies as fractions and percents.
   b. Make a histogram from the table in part (a).
   c. Draw a frequency polygon for the distribution in part (a), using the histogram in part (b).

42	8	54	93	38
27	61	84	33	30
60	21	24	91	52
31	39	37	68	39
77	64	36	47	44

2. Repeat Exercise 1 using the data at the right.

41	70	27	94	38
8	60	91	52	52
62	14	88	40	32
60	42	48	90	36
55	43	30	51	66

3. Make a histogram and a frequency polygon for the frequency distribution of Scholastic Aptitude Test—Mathematics scores shown below.

Class Interval	*Frequency*
450–500	15
500–550	8
550–600	8
600–650	5
650–700	4
700–750	2

4. The speeds of cars passing a point on a highway were measured for fifteen minutes. Make a histogram and frequency polygon for the frequency distribution of the speeds shown below. Speed measurements are in miles per hour.

Class Interval	*Frequency*
40–45	12
45–50	24
50–55	30
55–60	9
60–65	6

B

5. For the set of measurements shown below:
   a. Make a table, grouping measurements in the intervals 4.5–34.5, 34.5–64.5, 64.5–94.5, and showing frequencies and relative frequencies, as fractions and percents.
   b. Draw a histogram for the data shown in the table of part (a).
   c. Draw a frequency polygon for the data shown in the table of part (a).

48	91	10	36	70
39	51	29	62	57
20	27	42	81	41
25	42	18	56	83

6. Repeat Exercise 5 using these measurements:

59	65	16	39	90
8	35	42	84	36
73	7	34	48	61
26	19	28	43	58
88	43	48	38	11

7. The lengths in kilometers of some of the major rivers of the world are listed below.

River	*Length (km)*	*River*	*Length (km)*
Albany	976	Loire	1014
Amazon	6400	Mekong	4160
Amur	4320	Mississippi	3757
Brahmaputra	2880	Nile	6632
Colorado	2320	Orinoco	2560
Columbia	1989	Ottawa	1264
Congo	4349	Po	648
Dnieper	2272	Rio Grande	3016
Elbe	1158	St. Lawrence	1280
Euphrates	3576	Thames	344
Garonne	571	Yangtze	5440
Indus	2880	Zambezi	2720
Jordan	320		

a. Make a table of frequencies and relative frequencies, as fractions and percents, grouping the lengths in the intervals 0–1000, 1000–2000, 2000–3000, and so on.

b. Draw a histogram from the data in your table.

c. Use your histogram to draw a frequency polygon for the data.

C 8. Listed below are the numbers of immigrants in a recent year entering the United States from selected countries (recorded as the country of birth).

Country	*Number of Immigrants*	*Country*	*Number of Immigrants*
Austria	507	Italy	5,969
Canada	20,181	Japan	4,496
China	2,944	Lithuania	23
Denmark	378	Mexico	52,479
Finland	284	Norway	438
France	2,905	Poland	3,863
Germany	7,166	Spain	3,285
Greece	5,942	Turkey	1,306
Hungary	528	U.S.S.R.	1,919
India	18,625	Wales	128

a. Select appropriate grouping intervals and make a table of frequencies and relative frequencies, as fractions and percents.

b. Draw a histogram from the data in your table.

c. Use your histogram to draw a frequency polygon for the data.

d. Obtain a copy of a recent almanac. Compare the immigration statistics for these countries for the two most recent years given. What changes do you notice?

13–3 Cumulative Frequency

In analyzing a set of numerical data, it is often helpful to tabulate **cumulative frequencies** and **cumulative percents,** that is, the number and percent of measurements that are *less than or equal to* a given value.

The table below shows the frequency distribution for the set of 25 heights given in Section 13-1, along with the *cumulative frequency* and the *cumulative percent*. (The second column gives the frequency as a percent.) Notice that any measurement that falls on an interval boundary is included in the interval to its left; for example, 157 is included in the 155–157 interval.

Interval	*Frequency*	*Percent*	*Cumulative frequency*	*Cumulative percent*
153–155	1	4	1	4
155–157	4	16	5	20
157–159	6	24	11	44
159–161	9	36	20	80
161–163	3	12	23	92
163–165	2	8	25	100

The table tells us, for example, that 80% of the students have heights equal to or less than 161 cm. In the **cumulative frequency polygon** (Figure 3) displaying the facts in the table, the ordinates of the red dots are cumulative frequencies, and the abscissas are the right-hand endpoints of the corresponding intervals.

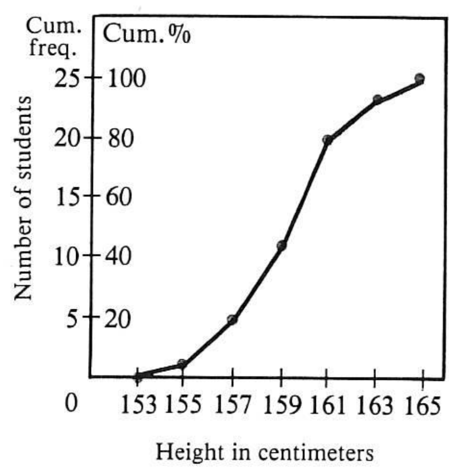


Figure 3

Because of rounding in the process of computation, the final value obtained in using cumulative percents sometimes does not appear to be exactly 100%.

EXAMPLE Make a table showing the frequencies, percents, cumulative frequencies, and cumulative percents for the following set of measures:

7, 8, 8, 6, 7, 8, 7, 8

SOLUTION Count the number of times each measure occurs. Then complete the table.

Measure	*Frequency*	*Percent*	*Cumulative frequency*	*Cumulative percent*
6	1	$12\frac{1}{2}$	1	$12\frac{1}{2}$
7	3	$37\frac{1}{2}$	4	50
8	4	50	8	100
	Total: 8	100		

Oral Exercises

Exercises 1–9 refer to the table of heights at the right.

Height (m)	*Frequency*	*Cumulative frequency*
1.50	3	3
1.51	2	a
1.52	4	9
1.53	*b*	15
1.54	4	19
1.55	9	28
1.56	7	c
1.57	*d*	40
1.58	5	45
1.59	3	48
1.60	2	50

1. What number does the entry *a* represent?
2. Explain how the third entry in the cumulative frequency column, 9, was determined, and then interpret what it means.
3. What number does *b* represent?
4. What number does *c* represent?
5. What number does *d* represent?
6. How many people were 1.55 m tall?
7. How many were at most 1.57 m tall?
8. How many people were measured in all?
9. How many people were more than 1.54 m but less than 1.57 m tall? State two ways of obtaining the answers to this question.

Written Exercises

Make a table showing the frequencies, percents, cumulative frequencies, and cumulative percents for each set of data.

A 1. 7, 9, 2, 3, 4, 3, 6, 3, 8, 4, 8, 6

2. 23, 21, 25, 22, 23, 23, 25, 21, 23, 22, 25, 23, 21, 23, 25

3. Copy and complete the table.

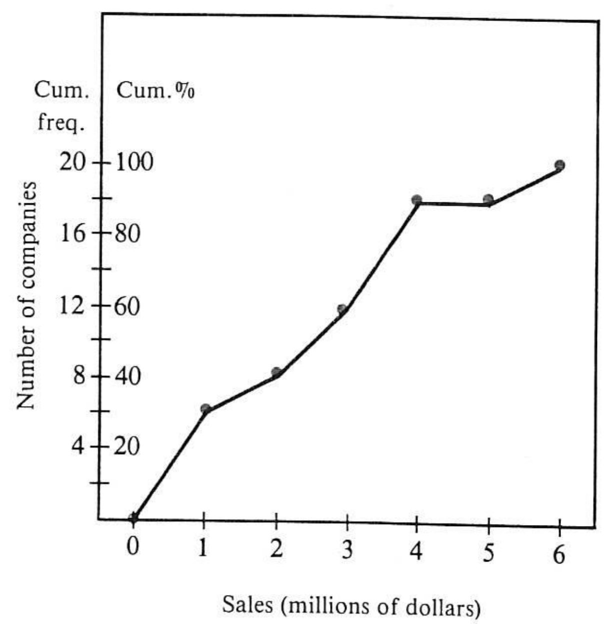
Cost in dollars	*Frequency*	*Percent*	*Cumulative frequency*	*Cumulative percent*
20	2	8	?	?
21	3	?	?	?
22	4	?	?	?
23	9	?	?	?
24	6	?	?	?
25	?	?	?	?
Total:	?	?		

4. Copy and complete the table.

Ages of students	*Frequency*	*Percent*	*Cumulative frequency*	*Cumulative percent*
13	3	6	?	?
14	8	?	?	?
15	12	?	?	?
16	9	?	?	?
17	10	?	?	?
18	6	?	?	?
19	?	?	?	?
Total:	?	?		

The cumulative frequency polygon shown at the right concerns the annual sales in each of 20 companies during a recent year.

5. How many companies had sales no greater than $1,000,000?
6. How many companies had sales up to $3,000,000, inclusive?
7. How many companies had sales between $4,000,000 and $5,000,000?
8. What percent of the companies had sales of at most $4,000,000?
9. What percent of the companies had sales over $3,000,000?



Draw a cumulative frequency polygon for the data in the given exercise.

10. Exercise 3

11. Exercise 4

For Exercises 12 and 13, copy and complete the table. Then draw a cumulative frequency polygon.

B 12.

Interval	*Frequency*	*Percent*	*Cumulative frequency*	*Cumulative percent*
115–130	2	?	?	?
130–145	0	?	?	?
145–160	4	?	?	?
160–175	4	?	?	?
175–190	7	?	?	?
190–205	8	?	?	?
Total:	?	?		

13.

Interval	*Frequency*	*Percent*	*Cumulative frequency*	*Cumulative percent*
0–100	7	14	?	?
100–200	10	?	?	?
200–300	12	?	?	?
300–400	?	?	?	?
400–500	6	?	?	?
Total:	?	?		

14. Following is a table showing the number of annual tax returns filed in a certain year with adjusted gross incomes of $30,000 or less. The number of returns in each interval has been recorded to the nearest million. Copy and complete the table, then draw a cumulative frequency polygon. Express all percents to the nearest tenth of a percent.

Adjusted gross income	*Frequency*	*Percent*	*Cumulative frequency*	*Cumulative percent*
0–$5000	18	?	?	?
$5000–$10,000	22	?	?	?
$10,000–$15,000	23	?	?	?
$15,000–$20,000	25	?	?	?
$20,000–$25,000	30	?	?	?
$25,000–$30,000	32	?	?	?

C 15. Toss two dice 100 times and record each total. Make a table showing the frequencies, percents, cumulative frequencies, and cumulative percents. Which total occurred most often? least often?

16. A sociology professor found the following family incomes in a survey of an urban neighborhood.

$36,600	$18,000
$34,800	$21,000
$49,000	$12,000
$3,980	$43,900
$18,500	$37,540
$7,400	$53,270
$29,150	$13,700
$32,200	$4,500
$9,680	$24,430
$10,710	$26,000
$9,400	$4,780
$20,750	$8,670
$64,000	$56,350
$21,650	$15,880
$29,000	$11,940

Divide the incomes into ranges of $5000, with the final range being "over $45,000." Show them in a table with frequencies, percents, cumulative frequencies, and cumulative percents.

Computer Exercises For students with computer experience

1. The following data represent the number of household members in a sample of 20 households.

5	1	3	4	4	2	3	7	2	1
2	4	5	1	3	8	4	3	8	4

Write a program that will read these numbers from a DATA list and will display in table form each number of household members from 1 through 8 and its frequency in the sample.

2. Modify the program that you wrote for Exercise 1 so that it also displays the *relative* frequency of each number of household members, both as a fraction and as a percent.

3. Modify the program that you wrote for Exercise 1 so that it displays only the *cumulative* frequency of each number of household members. (*Hint*: Test whether or not each data value is equal to 8. If it is, branch to the first of a series of steps that increment the counters of the frequencies. If it is not, test whether or not it is equal to 7. If it is, branch to the *second* in a series of such steps. If it is not, continue testing against 6, 5, 4, 3, 2, and 1 in a similar manner.)

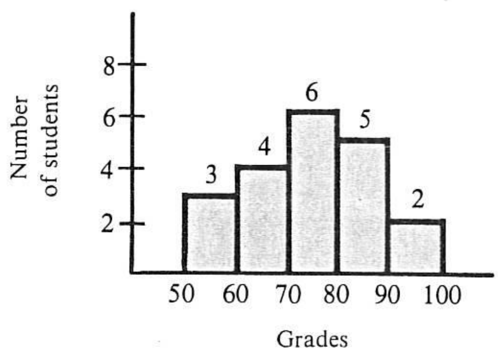
Self-Test 1

VOCABULARY
dot frequency diagram (p. 624)
frequency of a measurement (p. 624)
relative frequency of a measurement (p. 624)
histogram (p. 626)
frequency polygon (p. 627)
cumulative frequency (p. 631)
cumulative percent (p. 631)
cumulative frequency polygon (p. 631)

1. Make a dot frequency diagram for the given data and find the relative frequency of the measurement 5. *Obj. 1, p. 623*

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

2. Complete: In the histogram shown below, _?_ students received grades between 60 and 90, and _?_ students in all took the test. *Obj. 2, p. 623*



3. Make a table showing the frequency, percent, cumulative frequency, and cumulative percent for the following set of data: *Obj. 3, p. 623*

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

Then draw a cumulative frequency polygon for these data.

Check your answers with those at the back of the book.

Arithmetical Description of Data

OBJECTIVES for Sections 13-4 and 13-5:
1. To find the median, mode, and arithmetic mean of a frequency distribution.
2. To find the range, variance, and standard deviation of a frequency distribution.

13–4 Statistical Averages

You have seen how a graph can give a quick visual summary of the distribution of values in a large collection of data. A compact description of how the data are "centered" can be given by obtaining certain "averages" from the data: the *median*, the *mode*, and the *arithmetic mean*.

To obtain the median from a set of data, such as the 25 heights in Section 13-1, the data must first be arranged in order, as at the right.

The **median** in an ordered set of n values is the *middle* entry if n is odd. If n is even, then there are two middle measurements, and the median is half their sum. In this example, the median is 160.

The **mode** is the value that occurs with the *greatest frequency*. In this example, the mode happens to be the same as the median, 160. There may be more than one mode in a list of data. For example, in 1, 3, 3, 4, 4, 6, there are *two* modes: 3 and 4.

The **arithmetic mean,** often called the **average,** or simply the **mean,** is the sum of the n values divided by n. Here the mean is $\frac{3994}{25}$, or 159.76.

	155	
	156	
	157	
	157	
	157	
	158	
	158	
	158	
	158	
	159	
	159	
Mode	160	
Mode	160	← Median
Mode	160	
Mode	160	
Mode	160	
	161	
	161	
	161	
	161	
	162	
	163	
	163	
	165	
	165	
Sum:	3994	

In some cases the median may give a more accurate picture of a distribution than the mean because one or two extreme values can greatly affect the mean. For example, the annual income of the owner of a small business might be \$35,000, while the earnings of the four employees might be \$9000, \$9000, \$9000, and \$10,000. The owner would perhaps want to point to the *mean income* of the entire group, which is

$$\frac{\$35{,}000 + \$9000 + \$9000 + \$9000 + \$10{,}000}{5} = \frac{\$72{,}000}{5}, \text{ or } \$14{,}400.$$

The employees, on the other hand, would probably feel that the *median income*,

$$\$9000,$$

is a more meaningful figure as far as they are concerned.

In most cases, however, the arithmetic mean is the most reliable measure, and also the most useful one for computational work. The mode is of very limited value.

EXAMPLE 1 For the list of data 1, 7, 8, 2, 3, 6, 8, state:

a. the mode(s) b. the median c. the mean

SOLUTION Order the data from least to greatest:

1, 2, 3, 6, 7, 8, 8

a. There is one mode, 8.

b. The number of values is odd, so the median is the middle number, 6.

c. $\frac{1 + 2 + 3 + 6 + 7 + 8 + 8}{7} = \frac{35}{7} = 5$

$\therefore$ the mean is 5.

EXAMPLE 2 Find the mean of the 25 heights in the table in Section 13-1 by using their frequency count.

SOLUTION

Height	*Frequency*
155	1
156	1
157	3
158	4
159	2
160	5
161	4
162	1
163	2
165	2

To find the sum, multiply each height by its frequency. Then add the products.

$$155(1) + 156(1) + 157(3) + 158(4) + 159(2) + 160(5) + 161(4) + 162(1) + 163(2) + 165(2)$$
$$= 155 + 156 + 471 + 632 + 318 + 800 + 644 + 162 + 326 + 330$$
$$= 3994$$

$$\text{Mean} = \frac{3994}{25} = 159.76 \approx 159.8$$

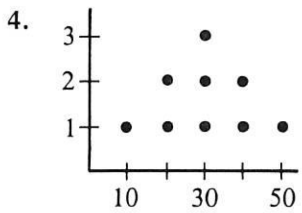
Oral Exercises

For the list of data 4, 4, 5, 6, 6, 7, 8, 8, 8, 9, state the following.

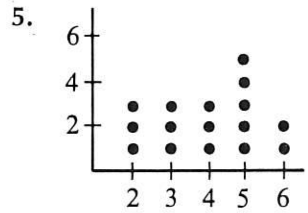
1. the mode(s)
2. the median
3. the mean

For the data in each dot frequency diagram, state the mode, median, and mean.

4.



5.



6. Does the median have to be a member of a set of data? Does the mode?
7. Does every set of data have a mode? Explain.
8. Suppose you received the following grades on six quizzes: 78, 91, 79, 41, 84, 83. Which average, the mean or the median, appears to provide the better description of your typical performance? Why?

Written Exercises

In Exercises 1–4, find the following for each list of data.

a. the mode(s) **b. the median** **c. the mean**

A

1. 4, 5, 5, 6, 7, 10, 12
2. 15, 16, 19, 19, 20, 21, 21, 21
3. 9, 13, 13, 23, 25, 12, 13, 18, 24, 25
4. 14, 26, 28, 17, 31, 23, 26, 21, 28, 26

In Exercises 5 and 6, the data represent scores on an algebra test given to two classes. Find the mean from each table of scores. If necessary, round to the nearest tenth.

5.

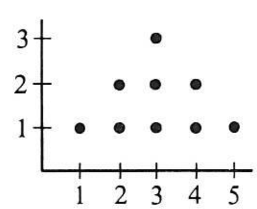
Score	*Frequency*
90	8
80	14
70	11
60	8
50	8
40	7

6.

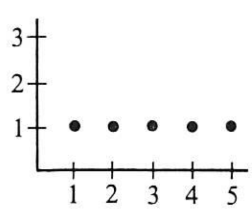
Score	*Frequency*
90	2
80	12
70	20
60	7
50	6
40	5

B

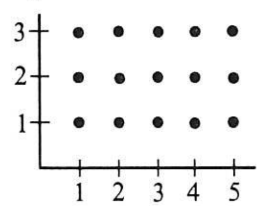
7. Find the median in Exercise 5.
8. Find the median in Exercise 6.
9. Determine the mode, median, and mean of the following four distributions of measurements. What do they have in common?



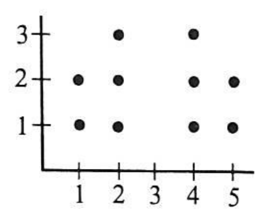
A



B



C



D

10. When the owner of a shoe store is putting in an order for more shoes, which descriptive measure of the foot sizes of the regular customers would be most useful: the mean, median, or mode? Explain.

11. If $r_1, r_2, r_3, \ldots, r_n$ is a list of n measurements, write a formula for their mean, M.

12. If the list in Exercise 11 is ordered from least to greatest, and $n = 25$, what symbol would you use to denote the median measurement?

13. If the sum of n measurements is 410 and their mean is 5, how many measurements are there?

14. If the mean of 6, 22, 40, x, 53, 14, 26, and 31 is 26, find the value of x.

15. The mean of 12 items is 7. If the mean of the first four items is 4, the mean of the next three items is 11, and the mean of the remaining 5 items is x, find the value of x.

16. Sean has scored an average of 172 after 27 games of bowling. What average must he achieve over the next 9 games to finish the season with a 173 average?

C 17. Show that if x and y are integers, and one is odd and one is even, then their mean is not an integer.

18. Show that if $x < y$, and if the mean of x and y is z, then $x < z < y$.

19. a. If there are an equal number of measurements one unit less than some value n as there are one unit greater than n, show that the mean of the set of measurements is n.

b. If the measurements are as described in part (a), and there are also an equal number of measurements two units less than n as there are two units greater than n, show that the mean is n.

c. What is the generalization of this pattern?

20. Show that if y^2 is the mean of x^2 and z^2 and $x^2 \neq z^2$, then $\frac{1}{z + x}$ is the mean of $\frac{1}{y + z}$ and $\frac{1}{x + y}$.

13–5 Measures of Variation

The **range** of a collection of data is the difference between the greatest and least values in the set. It is the simplest *measure of the variation* in the data. The range tells us nothing, however, about how the data are scattered or clustered together, in particular, around the mean. For example, consider these two lists of data, A and B.

A: 2, 5, 6, 7, 10
B: 2, 2, 6, 10, 10

They have the same range, 8, and the same mean, 6. But in list A the numbers are clustered much more closely about the mean than in list B.

You can compare the degree of scattering in two data samples by finding the *variance* of each one. The **variance** is the average of the squares of the *deviations,* or differences, of *all* the measurements from their mean.

EXAMPLE 1 Find the variance for each of lists A and B on page 640, and interpret the results.

SOLUTION For list A, the deviations from the mean are: $2 - 6$, or -4; $5 - 6$, or -1; $6 - 6$, or 0; $7 - 6$, or 1; $10 - 6$, or 4. Then the variance is:

$$\frac{(-4)^2 + (-1)^2 + 0^2 + 1^2 + 4^2}{5} = \frac{34}{5} = 6.8$$

For list B, the deviations are: $2 - 6$, or -4; $2 - 6$, or -4; $6 - 6$, or 0; $10 - 6$, or 4; $10 - 6$, or 4.

The variance is:

$$\frac{(-4)^2 + (-4)^2 + 0^2 + (4)^2 + (4)^2}{5} = \frac{64}{5} = 12.8$$

Since the variance in A is 6.8 whereas in B it is 12.8, the data in B are considerably more scattered from their mean than are the data in A.

If the observations are measured in units, such as meters, then the variance is given in terms of square units. When you take the *principal square root* of the variance, you then have a measure of variation that is in the same units as those of the given data. This measure is called the **standard deviation,** usually denoted by s.

EXAMPLE 2 Find the standard deviation s for lists A and B.

SOLUTION For A, the variance s^2 is 6.8, so $s = \sqrt{6.8} \approx 2.6$
For B, the variance s^2 is 12.8, so $s = \sqrt{12.8} \approx 3.6$

Oral Exercises

For Exercises 1–3, use the list of data 3, 4, 6, 7, 10, and find the following.

1. the range
2. the mean
3. the deviation of each value from the mean
4. Find the variance for the list 3, 4, 6, 7, 10.
5. The mean of each of the following lists of data is 8. Which one do you think has the greater variance?
   a. 5, 7, 8, 9, 11
   b. 4, 5, 5, 6, 7, 9, 10, 11, 15
6. Check your answer to Exercise 5 by finding the variance of each list.

Written Exercises

Find the range for each list of data.

A 1. 3, 14, 16, 29, 32, 50

2. 2, 3, 13, 14, 14, 15, 20, 23

3. For the list 2, 4, 7, 8, 9, find to the nearest tenth:
   a. the variance, s^2
   b. the standard deviation, s

4. For the list 10, 11, 13, 14, 17, find to the nearest tenth:
   a. the variance, s^2
   b. the standard deviation, s

In Exercises 5 and 6, the data represent hourly wages of employees in two small businesses. Use the frequency count and find the following to the nearest tenth.

a. the mean **b. the variance** **c. the standard deviation**

B 5.

Hourly wage	*Frequency*
\$ 6	2
\$ 9	2
\$12	2
\$14	1
\$20	1

6.

Hourly wage	*Frequency*
\$ 8	1
\$10	4
\$14	2
\$17	2
\$20	1

7. Refer to the data lists A and B on page 640. Find the sum of the deviations for each list. How do the two sums compare? Do they tell you anything about how the data are scattered?

If $r_1, r_2, r_3, \ldots, r_n$ denotes a list of n observations and m denotes their mean, write an expression for each of the following.

8. the deviation of r_1 from the mean
9. the average of the set of all the deviations from the mean
10. the variance, s^2
11. the standard deviation, s

C 12. a. If the deviations of data from the mean of the data are added, positive and negative quantities will offset each other, resulting in a sum of zero. An early statistical attempt to overcome this effect was the introduction of the absolute values of deviations in the addition process. Thus the *average deviation* from the mean was defined to be the following:

$$\frac{|r_1 - m| + |r_2 - m| + \cdots + |r_n - m|}{n}$$

Calculate this quantity for the data in Exercises 3 and 4. Compare this description of "scattering" with the standard deviation s found in part (b) of Exercises 3 and 4. Do the average deviations also appear to indicate which data are more scattered?

b. The average deviation is seldom used except as a percent of the mean. For the data in Exercises 3 and 4, express the average deviation as a percent of the mean.

Computer Exercises

For students with computer experience

1. Write a program that will allow you to input a list of positive numbers and will compute the arithmetic mean of the numbers. The program should accept as many positive numbers as you want to input and should have you signal the end of your list by inputting one negative number. (Be careful that the program does not use this negative number in computing the mean.)
2. Modify the program that you wrote for Exercise 1 so that it will allow you to input a list of positive numbers *each followed by a frequency*.
3. Write a program that will allow you to input a set of numerical data and will compute the variance and the standard deviation of the data. The program should first ask that you input the *number* of data values to be entered, then ask you to input the data values themselves. (*Note*: This program requires the use of an *array* of values, and you will need to work with *subscripted variables*.)
4. Suppose that you are given a set of 25 quiz grades consisting of whole numbers from 1 through 10. Write a program that will compute the mode(s) of this set of data. (*Hint*: This program requires the use of an array of values. For each value of N from 1 through 10, find the frequency of the grade N and store this number as the value of G(N). Then, for each value of N, test whether G(N) is greater than or equal to *all* the values G(J), where J is a whole number from 1 through 10. If it is, then N is a mode of the set of data.)

Self-Test 2

VOCABULARY

median (p. 637)
mode (p. 637)
arithmetic mean (p. 637)
range (p. 640)
variance (p. 641)
standard deviation (p. 641)

1. Find the median, mode, and mean of the following data: 7, 2, 2, 6, 4, 2, 7, 2, 4, 14. *Obj. 1, p. 637*
2. Find the range, variance, and standard deviation of the data in Exercise 1. *Obj. 2, p. 637*

Check your answers with those at the back of the book.

PROGRAMMING IN BASIC

The following program illustrates how a computer can be used in statistical analysis. Note that the program uses the READ and DATA statements in lines 60 and 410–450.

```
10   PRINT "ANALYSIS OF A SET OF DATA:"
20   LET S = 0
30   LET A = 0
40   LET B = 0
50   LET C = 0
60   READ M
70   IF M = -1 THEN 170
80   LET S = S + M
90   IF M < 40 THEN 150
100  IF M < 70 THEN 130
110  LET A = A + 1
120  GOTO 60
130  LET B = B + 1
140  GOTO 60
150  LET C = C + 1
160  GOTO 60
170  LET N = A + B + C
180  PRINT
190  PRINT "N = ";N,"AVERAGE = ";S/N
200  PRINT "INTERVAL     FREQ.      REL. FREQ."
210  PRINT "10-39          ";C;"          ";C/N
220  PRINT "40-69          ";B;"          ";B/N
230  PRINT "70-99          ";A;"          ";A/N
240  PRINT
250  PRINT "10-39:     ";
260  LET X = C
270  GOSUB 360
280  PRINT "40-69:     ";
290  LET X = B
300  GOSUB 360
310  PRINT "70-99:     ";
320  LET X = A
330  GOSUB 360
340  GOTO 460
350  REM *SUBROUTINE
360  FOR I = 1 TO X
370  PRINT "*";
380  NEXT I
390  PRINT
400  RETURN
```

```
410  DATA 29,72,17,53,84,55,14,93
420  DATA 65,29,25,82,16,84,92,39
430  DATA 48,72,71,35,52,68,44,16
440  DATA 64,86,67,86,46,67,86,63
450  DATA 46,48,67,25,45,72,88,91,-1
460  END
```

Exercises

1. Type in and RUN the program as given.
2. Modify the program so that the intervals are 0–19, 20–39, 40–59, 60–79, and 80–99. Then RUN the revised program.

Probability

OBJECTIVES for Sections 13-6 and 13-7:
1. *To find the probability of a specified event in an experiment with equally likely outcomes.*
2. *To calculate an experimental probability.*

13–6 The Probability of an Event

The main use for statistical theory is to help people make decisions on the basis of incomplete information. For example, business people and politicians often use marketing surveys and opinion polls to assess the attitude of the *population* (the customers or voters) by studying a *sample* of it. In choosing a sample of data and then trying to draw conclusions about the population, the statistician must consider the theory of *probability*.

Simple games of chance offer ideal situations for understanding the notion of probability. Suppose that you perform the experiment of tossing a die and observing the number of dots on the top face. The six possible outcomes are the following:

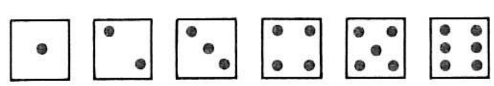


Figure 4

If the die is not loaded you can assume that each outcome is equally likely. That is, each outcome has an equal chance of occurring, namely, 1 out of 6. Then the *measure* of chance, or **probability**, that a particular outcome will occur is $\frac{1}{6}$.

> If an experiment has n possible, equally likely outcomes, then the probability that any one of them will occur is $\frac{1}{n}$.

An event is a *specified subset* of the set of all possible outcomes in an experiment. When you toss a die, the probability P of the event "the top face shows *either* 3 *or* 4 dots" is $\frac{2}{6}$. That is, there are 2 out of 6 chances that either one or the other of these outcomes will occur. If any *one* of the outcomes in an event occurs, you say that the event *occurs.*

> If an experiment has n possible, equally likely outcomes and an event E consists of e of these outcomes, then the probability P that E will occur is $\frac{e}{n}$. That is:
>
> $$P(E) = \frac{\text{number of outcomes in } E}{\text{number of possible outcomes}} = \frac{e}{n}$$

EXAMPLE 1 There are 4 jacks in a standard bridge deck of 52 cards. What is the probability of the event that a card drawn at random is a jack?

SOLUTION Since there are 4 jacks and 52 cards in the deck, $e = 4$ and $n = 52$. Then

$$P(E) = \frac{e}{n} = \frac{4}{52} = \frac{1}{13}.$$

EXAMPLE 2 A jar contains 2 red, 4 blue, and 5 white marbles. If a marble is drawn from the jar at random, what is the probability of the event "the marble is not red"?

SOLUTION Of the 11 marbles in the jar, 9 are not red. Hence

$$P(E) = \frac{9}{11}.$$

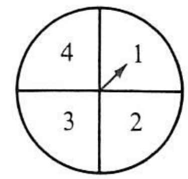
All the experiments described in this chapter are assumed to be *random* ones, that is, experiments conducted in such a way that the outcomes are strictly a matter of chance.

Oral Exercises

1. When a die is tossed, what is the probability that the side with just two dots will be on top?
2. When you toss a coin, what is the probability that it will land with the head showing?

3. When you draw a marble, while blindfolded, from a jar containing 4 marbles—1 red, 1 blue, 1 white, 1 green—what is the probability of drawing a blue marble?
4. In Exercise 3, what is the probability of the event "either a red or a green marble is drawn"?
5. When a coin is tossed, what is the probability of the event "it will land either heads or tails"?

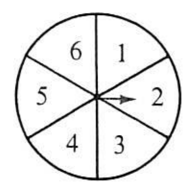
Exercises 6-8 refer to the spinner at the right. Assume that the arrow will neither stop on a dividing line nor favor a particular numbered region.



6. What are the possible outcomes of a spin? Are they all equally likely?
7. What is the probability that the arrow will stop on the region labeled 1? 2? 4?
8. What is the probability that the arrow will not stop on the region labeled 3?
9. Suppose that in the city raffle to benefit the library, 5000 tickets are sold and you buy a single ticket.
   a. What is the probability that your ticket wins the grand prize, if only one grand prize winner is drawn at random?
   b. If your friend also has a single ticket, how does your chance of winning the grand prize compare to your friend's?
   c. What can you do to "double your chances"? If you do this, what is the probability that you will be a grand prize winner?

Written Exercises

Exercises 1–14 refer to the spinner at the right. Assume that the arrow will neither stop on a dividing line nor favor a particular numbered region.



A 1. List the possible outcomes of a spin.
2. List the outcomes that are prime numbers.

What is the probability that after a spin the arrow will stop on:

3. region 3?
4. region 4?
5. region 1 or 2?
6. region 2 or 6?
7. an even-numbered region?
8. a region with a prime number?
9. a region with a number less than 3?
10. a region with a number greater than 2?
11. a region with a number greater than 3?
12. a region with a number less than 2?
13. a region with a number greater than 6?
14. a region with a number less than 7?

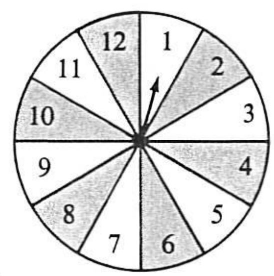
A jar contains 2 white, 3 red, 1 green, and 4 blue marbles. If a marble is drawn at random from the jar, what is the probability that the marble is:

15. white?
16. red?
17. green?
18. blue?
19. not white?
20. not blue?
21. white or red?
22. red or green?
23. red, green, or white?
24. red, green, or blue?
25. yellow?
26. not yellow?

In a standard deck of 52 cards, there are four of each kind of card: ace, king, queen, jack, ten, nine, . . . , two. Half the cards are black and half are red, with two black and two red aces, two black and two red kings, and so on. If a card is drawn at random from the deck, what is the probability that it will be:

27. a black card?
28. a ten?
29. a king?
30. a red card?
31. a four?
32. an ace?
33. a red jack?
34. a black ace?
35. not a red queen?
36. not an ace?
37. not a queen?
38. not a black ace?

What is the probability that after a spin the arrow will stop on:

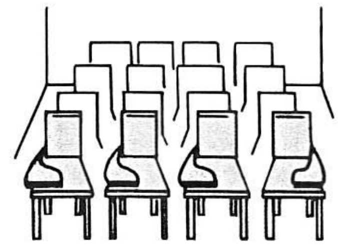


B 39. an unshaded region?
40. a region shaded red?
41. a region corresponding to a multiple of 3?
42. a region corresponding to a multiple of 2?
43. a region corresponding to a factor of 12?
44. a region corresponding to a factor of 10?
45. List the set of possible outcomes when a penny and a nickel are tossed, using the symbols H and T for heads and tails. (For example, for the outcome when the penny lands heads and the nickel tails, you would write HT; but for the outcome when the penny lands tails and the nickel heads, you would write TH.)
46. In Exercise 45 do the events "they both land heads" and "they do not both land heads" have the same probability? What is the probability of each of these events?

C 47. A jar contains four white tickets numbered 1–4, and three red tickets numbered 1–3. What is the probability that one ticket drawn randomly has either an even number or is red?
48. What is the probability that a card drawn at random from a standard deck of 52 cards is:
   a. either a king or a red card?
   b. neither a king nor a red card?

13–7 Experimental Probability

Suppose that in a particular year, a school with 1000 students has 40 who are left-handed. Then there are 40 chances in 1000, or 1 in 25, that a student chosen at random will be left-handed. If there are 25 seats in each classroom, would you therefore conclude that every classroom should have exactly one seat with an arm for left-handed writers? Your intuition and experience tell you, "Of course not! For example there might be three left-handed students in one room and none in another."



What is the probability that exactly 40 of the school's 1000 students the next year will be left-handed? Again, experience based on repeated observations would indicate a relatively small probability of this event. Nevertheless, in the absence of further information regarding the percent of left-handed people in the population, you would have to give $\frac{1}{25}$ as your best estimate of the probability that a randomly chosen student in that school next year will be left-handed.

A certain baseball player may have a *batting average* (that is, a ratio of safe hits to times at bat) of 0.400 for a particular season. Then the **experimental probability** of the player making a safe hit the next time at bat is $\frac{4}{10}$, or $\frac{2}{5}$. This does not mean, however, that the player will *surely* make 2 hits in any given 5 times at bat.

In many real-life situations such as these, you have only experimental probabilities from which to predict future occurrences of events.

If an experiment is conducted n times, and an event E occurs e of these times, then the experimental probability P that E will occur in another trial is $\frac{e}{n}$. That is:

$$P(E) = \frac{\text{number of occurrences of } E}{\text{number of trials}} = \frac{e}{n}$$

EXAMPLE A thumbtack was tossed 100 times. It landed "point up" 60 times and "point down" 40 times.

a. What is the experimental probability that on the next toss it will land point up?

b. If it were tossed 50 additional times, about how many times would you expect it to land point down?

c. If on 50 additional tosses it actually landed point down 25 times, what would be the experimental probability, based on all 150 tosses, that on the next toss it would land point down?

SOLUTION a. For the event E that the thumbtack lands point up, you have $e = 60$ and $n = 100$. Therefore,

$$P(E) = \frac{60}{100} = 0.6.$$

b. Since the experimental probability that the thumbtack will land point down is $\frac{40}{100}$, or 0.4, you would expect it to land point down about 50×0.4, or 20, times out of the additional 50 tosses.

c. If the thumbtack lands point down 40 times out of the first 100 throws, and 25 times out of the next 50 throws, then in all you have $e = 40 + 25 = 65$ and $n = 100 + 50 = 150$. Therefore the experimental probability that the thumbtack will land point down is now

$$P(E) = \frac{e}{n} = \frac{65}{150} \approx 0.43.$$

Oral Exercises

1. A baseball player's batting average is 0.250.
   a. What is the experimental probability that the player will get a hit the next time at bat?
   b. What is the experimental probability that the player will not get a hit?
   c. In the next 8 times at bat, how many times would you expect the player to get a hit?
2. If the experimental probability is $\frac{5}{8}$ that a softball player will get a hit the next time at bat, does this mean that the player is sure to get 5 hits in the next 8 times at bat?
3. The dietician in a company cafeteria made an informal survey one day, finding that 130 employees bought lunch in the cafeteria, 50 brought their lunches with them and ate in the cafeteria, and 20 did not come to the cafeteria at all. What is the experimental probability that an employee will buy lunch in the cafeteria? What is the experimental probability that an employee will not bring a lunch to the cafeteria?
4. If two carriers of a certain trait marry, the experimental probability of any child of theirs having the trait is $\frac{1}{4}$. If they have 4 children, is it true that:
   a. one of them is certain to have the trait?
   b. none of them might have the trait?
   c. all of them might have the trait?

Written Exercises

A 1. If a football quarterback is completing passes at the rate of 55%, what is the experimental probability that:

a. the quarterback will complete the next pass?

b. the quarterback will not complete the next pass?

2. A dart player has hit the bull's eye 8 times out of 40 throws. What is the experimental probability that:

a. the player will hit the bull's eye in the next throw?

b. the player will not hit the bull's eye in the next throw?

3. During a baseball season, a baseball player gets 10 hits out of 25 times at bat.

a. What is the experimental probability that the player will get a hit the next time at bat?

b. About how many hits would you expect the player to get the next 15 times at bat?

c. Suppose that the player actually gets 2 hits the next 15 times at bat. What would be the experimental probability that the player would get a hit the next time at bat?

4. In a basketball game, a player makes the basket on 8 free throws out of 12.

a. What is the experimental probability that the player will make a basket on the next free throw?

b. How many times would you expect the player to make the basket on the next 3 free throws?

c. If the player actually makes the basket all 3 times in the next 3 free throws, what is the experimental probability that the player will make a basket the next time?

Suppose that the results of 36 successive random drawings of one marble from a jar containing an unknown number of colored marbles are as follows: 4 blue, 14 red, 12 yellow, 6 green. What is the experimental probability that the next marble drawn will be:

5. red?
6. blue?
7. green?
8. yellow?
9. yellow or green?
10. red or blue?
11. not purple
12. purple

B 13. Refer to the drawing described in Exercises 5–12. Assume that the ratios of colored marbles found in the drawing are representative of the total number of marbles that were in the jar. Using that fact, can you tell how many marbles were in the jar? Explain.

14. If it is found that there are 90 yellow marbles in the jar described in Exercises 5–12, about how many green ones would you expect to find? About how many blue ones would you expect to find?

C 15. United States mortality tables show that about 95% of people born now can expect to be alive at age 40, but only about 75% of those born now will still be alive at age 65. What is the experimental probability that a person born now who reaches age 40 will still be alive at 65? (Assume that life expectancies do not change in the next 65 years.)

Computer Exercises

For students with computer experience

You will need to use the computer's random-number function for the following exercises. The format of this function varies from computer to computer, so you may need to consult your teacher or the computer manual before you proceed. If you are programming in BASIC, on many computers either RND(1) or RND(0) will generate a random number greater than or equal to 0 and less than 1.

1. Write a program that will generate a random *integer* from 1 through 10 one hundred times and will determine how many times the integer 3 was generated.
2. Modify the program that you wrote for Exercise 1 so that it will determine how many times an integer greater than 7 was generated.
3. Modify the program that you wrote for Exercise 1 so that it will generate a random integer from 1 through 100 one hundred times and will determine how many times an integer divisible by 5 was generated.
4. Write a program that will simulate one hundred rolls of two dice and will determine how many times the total number of dots on the two faces that turn up is 7.
5. Modify the program that you wrote for Exercise 4 so that it will determine how many times the number of dots on each of the two faces that turn up is the same.

Self-Test 3

VOCABULARY probability (p. 645), event (p. 646), experimental probability (p. 649)

When you toss a die, what is the probability that:

1. the top face will show 4 dots? *Obj. 1, p. 645*
2. the top face will show 7 dots?

When you toss a die, what is the probability that:

3. the top face will show an odd number of dots?
4. the top face will show either an even number or an odd number of dots?
5. Assume that it has rained on Thanksgiving Day in New York City for 35 out of the past 65 years. What would be the experimental probability that it will rain there next Thanksgiving Day? *Obj. 2, p. 645*

Suppose the results, when you drew 40 slips of paper at random from a hat containing 100 slips, were as follows: 6 purple, 10 red, 8 yellow, 16 white. What is the experimental probability that the next slip you draw will be:

6. purple?
7. red or yellow?
8. yellow or white?
9. How many red slips would you expect to find in the hat?
10. How many purple slips would you expect to find in the hat?

Check your answers with those at the back of the book.

Chapter Summary

1. *Statistics* is the science of organizing and analyzing a set of facts, or data, so that probable conclusions can be drawn from the data.
2. The number of occurrences of a particular measurement is called the *frequency* of the measurement. The *relative frequency* is the ratio of the frequency of a particular measurement to the total number of measurements.
3. *Histograms* and *frequency polygons* are used to help visualize frequency distributions for large sets of data.
4. The number and percent of measurements that are less than or equal to a given value are called the *cumulative frequency* and the *cumulative percent*.
5. The *median*, the *mode*, and the *arithmetic mean* are three descriptions of the "center" of a set of data.
6. The *range*, the *variance*, and the *standard deviation* are three measures of the variation in a set of data.
7. If an experiment has n possible, equally likely outcomes, then the *probability* that any one of them will occur is $\frac{1}{n}$.
8. An *event* is a specified subset of the set of all possible outcomes in an experiment.

Chapter Review

Write the letter of the correct answer.

1. What is the frequency of the letter I in the word IMITATION? *13-1*

   a. 2 b. 3 c. $\frac{1}{3}$ d. $\frac{1}{6}$

2. What is the relative frequency of the letter N in the word DINNER?

   a. $\frac{2}{5}$ b. 2 c. $\frac{1}{3}$ d. $\frac{1}{6}$

3. For the list of data 46, 5, 14, 16, 21, 9, 19, 23, 31, 13, what is the relative frequency of the data in the interval 15–20? *13-2*

   a. $\frac{1}{10}$ b. $\frac{2}{10}$ c. $\frac{3}{10}$ d. $\frac{4}{10}$

4. What percent of the data in Exercise 3 is in the interval 10–15?

   a. 10% b. 20% c. 30% d. 40%

5. What is the cumulative frequency for the interval 45–50, given the following list of data: 26, 63, 34, 51, 40, 19, 62, 88, 53, 54, 28, 48? *13-3*

   a. 1 b. 3 c. 5 d. 6

6. Find the mean for the following list of data: 16, 33, 25, 40, 38, 42, 19, 19. *13-4*

   a. 19 b. 25 c. 29 d. 30

7. Find the median for the data in Exercise 6.

   a. 19 b. 25 c. 26 d. 29

8. Find the variance for the following list of data: 7, 9, 11, 15, 18. *13-5*

   a. 4 b. 12 c. 16 d. 80

9. Find the standard deviation for the following list of data: 5, 6, 8, 9, 9, 11.

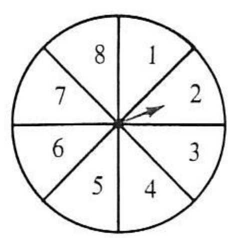
   a. 1.41 b. 2.00 c. 2.43 d. 4.90

10. A jar contains 5 white, 3 red, and 4 blue marbles. If a marble is drawn at random from the jar, what is the probability that it is not blue? *13-6*

    a. $\frac{1}{4}$ b. $\frac{1}{3}$ c. $\frac{2}{3}$ d. $\frac{5}{12}$

11. What is the probability that after a spin the arrow will stop on region 3 or 8?

    a. $\frac{1}{4}$ b. $\frac{1}{2}$ c. $\frac{3}{4}$ d. $\frac{3}{8}$



12. Refer to the spinner in Exercise 11. What is the probability that after a spin the arrow will stop on a region corresponding to a prime number?

a. $\frac{1}{4}$ b. $\frac{1}{2}$ c. $\frac{3}{4}$ d. $\frac{5}{8}$

13. A lacrosse player has hit the goal 7 out of 55 times. Approximately what is the experimental probability that the next attempt will be successful? *13-7*

a. 0.13 b. 0.07 c. 0.14 d. 0.21

14. A baseball player's batting average is 0.167. In the next 12 times at bat, how many times would you expect the player to get a hit?

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

Chapter Test

1. Make a table showing the frequencies and relative frequencies, as fractions and percents for the following set of data: 22, 20, 23, 21, 25, 21, 23, 25, 24, 23. *13-1*

2. Make a histogram and a frequency polygon for the data at the right. Group them in the intervals 10–15, 15–20, 20–25, 25–30. *13-2*

13	18	23	27	29
24	16	14	12	19
21	30	19	24	22

3. Make a table showing the frequencies, percents, cumulative frequencies, and cumulative percents for the data in Exercise 2. *13-3*

4. Find the median, mode, and mean for the following data: 9, 11, 13, 6, 22, 10, 8, 11, 16, 11, 9, 11, 10, 15, 18. *13-4*

For the list of data 15, 8, 18, 9, 20, find to the nearest tenth:

5. the variance

6. the standard deviation *13-5*

7. What is the probability of drawing at random a red ace from a standard deck of 52 cards? *13-6*

8. A box contains three red tickets numbered 1–3, five green tickets numbered 4–8, and twelve blue tickets numbered 9–20. If a ticket is drawn at random from the box, what is the probability that it will be an even-numbered blue ticket?

9. Claire predicted correctly 13 out of 35 times the day for each week's surprise algebra quiz. What is the experimental probability that her next prediction will be correct? *13-7*

10. A thumbtack was tossed 200 times. It landed "point up" 140 times and "point down" 60 times. If it is tossed 30 additional times, about how many times would you expect it to land point up?

Cumulative Review

Chapter 9

Solve.

1. $\frac{3x}{2} + \frac{8-4x}{7} = 3$
2. $\frac{a+1}{4} - \frac{3}{2} \geq \frac{2a-9}{10}$
3. $\frac{m}{m+2} = \frac{3}{5}$
4. $\frac{6}{n-8} - 1 = \frac{2}{n-3}$
5. 74 is 5% of what number?
6. 24 is what percent of 64?
7. What number is $\frac{3}{4}$% of 500?
8. If y varies directly as x^2, and if $y = 36$ when $x = 2$, find y when $x = 10$.
9. If c varies inversely as d, and if $c = 12$ when $d = \frac{1}{4}$, find c when $d = 9$.
10. If y varies jointly as x and z, and if $y = 24$ when $x = 2$ and $z = 3$, find y when $x = 3$ and $z = 4$.
11. One pipe can empty a tank in 5 h, while a second pipe can empty the same tank in 7 h. How long would it take to empty the tank if both pipes are used at the same time?
12. The interest rate required to yield a given income is inversely proportional to the amount of money invested. Bonnie receives income from $20,000 that she has invested at an annual interest rate of 5.5%. How much money must Carl invest at an annual interest rate of 8% in order to receive the same income as Bonnie?

Chapter 10

13. Express $\frac{25}{16}$ as a terminating or repeating decimal.
14. Express $4.\overline{09}$ as a fraction in simplest form.

Simplify.

15. $-\sqrt{(35)^2}$
16. $\sqrt{120}$
17. $5\sqrt{12} - 7\sqrt{27}$
18. $\sqrt{14} \cdot \sqrt{21}$
19. $(4\sqrt{7})(2\sqrt{5})$
20. $(3\sqrt{2} - 1)(\sqrt{2} + 4)$
21. $\frac{\sqrt{40}}{\sqrt{5}}$
22. $\sqrt{\frac{1}{12}}$
23. $\frac{3}{\sqrt{2} + 1}$
24. $\sqrt[3]{-32}$
25. $\sqrt{100x^2y^4}$
26. $\sqrt[5]{a^5b^{10}c^{15}}$

Solve.

27. $5 - 3x^2 = -22$
28. $4y^2 - 9 = 127$
29. $\sqrt{m} - 5 = 2$
30. $\sqrt[3]{x} - 5 = 3$

Chapter 11

Solve. Express irrational solutions in simplest form.

31. $m^2 - 6m + 8 = 0$
32. $n^2 + 6n + 4 = 0$
33. $j^2 + 7j = -2$
34. $k^2 = 20k - 19$
35. $2x^2 = 3x - 1$
36. $8y + 2 = 5y^2$

For each function, determine its zeros, if any, and its maximum or minimum value.

37. $f: x \longrightarrow -2x^2$
38. $g: x \longrightarrow \frac{1}{3}x^2$
39. $G: x \longrightarrow 2x + x^2$
40. $H: x \longrightarrow 3 - x^2$

Sketch the graph of each function and estimate its zeros, if any.

41. $g: x \longrightarrow x^3 - 2$
42. $h: x \longrightarrow -x^4$

Chapter 12

43. If A is a quadrantal angle and $-350° \leq m\angle A \leq -250°$, find $m\angle A$.
44. If $\angle A$ is in standard position, $90° \leq m\angle A \leq 180°$, and $\angle A$ is coterminal with an angle of $-585°$, find $m\angle A$.
45. Use the table on page 684 to find sin 76°.
46. If $0° \leq m\angle A \leq 90°$ and $\tan A = 0.2250$, find $m\angle A$ to the nearest degree.
47. In right triangle ABC, $c = 10$ and $a = 8$. Find $m\angle B$ to the nearest degree.
48. In right triangle ABC, $c = 10$ and $m\angle A = 21°$. Find b to the nearest tenth of a unit.
49. If a vector in standard position has its terminal point at $(-2, 5)$, find the norm and direction of the vector.
50. If vectors $\mathbf{s}$ and $\mathbf{v}$ are in standard position with terminal points $(-1, 5)$ and $(2, 7)$, respectively, and $\mathbf{v} = \mathbf{s} + \mathbf{t}$, find $\|\mathbf{t}\|$.

Chapter 13

51. Make a table that shows the frequencies, the relative frequencies as fractions, and the relative frequencies as percents for the following set of data: 8, 9, 5, 7, 7, 10, 4, 8, 7, 9.
52. Find the median, mode, and mean of the following set of data: 25, 20, 22, 25, 27, 21, 25, 22, 16, 24.
53. When you toss a die, what is the probability that the top face will show 6 dots?
54. An archer hits the bull's eye in 6 out of 50 tries. What is the experimental probability that the next arrow this archer shoots will miss the bull's eye?

Estimating Wildlife Populations

Ecologists often need to know how many animals of a certain species live in a particular area. It may not be possible to get an exact count. There is no practical way to be sure that, for example, all the whales in an ocean have been counted.

One way to estimate the number of animals is to catch some animals, tag them, and release them. Call the number of tagged animals t_1.

After the tagged animals have had time to mix with the rest of the population, more animals are caught or observed. Call this number c. Of these animals, a number t_2 have tags.

Assume that the ratio $\frac{t_2}{c}$ is equal to the fraction of all the animals that are tagged. If the total number of animals is n, then the following proportion can be used to calculate n.

$$\frac{t_2}{c} = \frac{t_1}{n}$$

$$t_2 n = c t_1$$

$$n = \frac{c t_1}{t_2}$$

For example, suppose that 50 tigers are tagged within a particular area. Later, 100 tigers are observed within the same area. Of these, 10 have tags. Then the total number of tigers can be estimated as follows.

$$\begin{aligned} n &= \frac{c t_1}{t_2} \\ &= \frac{100 \times 50}{10} \\ &= 500 \end{aligned}$$

There are several possible sources of error in this procedure. Some animals' tags might come off. The tagged animals might not mix completely with the rest of the population. Animals may join or leave the population before the second capture. However, other counting methods such as aerial surveys confirm that this procedure produces good results.

Exercises

1. Ecologists tag 1000 whales. Later they observe 2500 whales. Of these, 400 have tags. Estimate the total number of whales.
2. To count fur seal pups, 5000 pups are tagged. Two weeks later, 1000 pups are caught. Of these, 200 have tags. Estimate the total number of fur seal pups.
3. To count bass in a lake, 215 bass are tagged. Later 99 bass are caught, of which 15 have tags. About how many bass are in the lake?
4. In the same lake 240 pickerel are tagged. Later 342 pickerel are caught, of which 18 have tags. About how many pickerel are in the lake?
5. Observers in airplanes estimate that there are 1500 moose in a park. To check this estimate, 100 moose are tagged. Later 180 moose are observed. How many of these would be expected to have tags?