11 Statistics and Probability

Statistics show that very few accidents occur in the early morning, in fog, at speeds in excess of 150 km/h. Does this mean that these are the safest conditions for driving? (See Section 11-4, *Example 1*.)

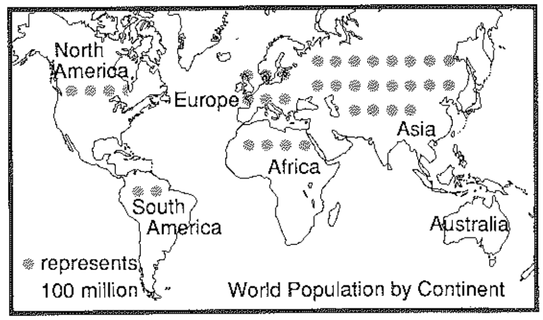
11-1 INTERPRETING GRAPHS

In the fast-moving computer age, we encounter vast quantities of data. *Statistics* is the branch of mathematics that deals with the collection, organization, and interpretation of data. These data are usually organized into tables and/or presented as graphs. Some common types of graphs are shown below. Try to answer the question that accompanies each graph.

Pictograph

The graph shown here uses the symbol ● to represent 100 million people. A graph that uses a symbol to represent a certain amount is called a *pictograph.*

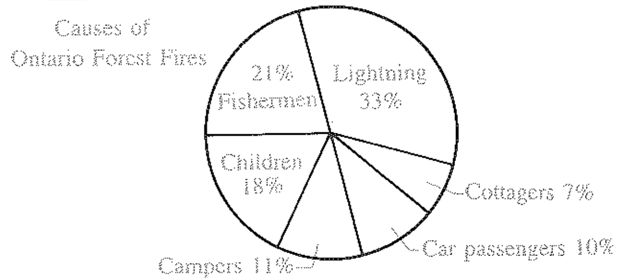
- What are the populations of North America and Europe?



Circle Graph

In a *circle graph,* a complete set of data is presented by the circle. Various parts of the data are represented by the sectors of the circle.

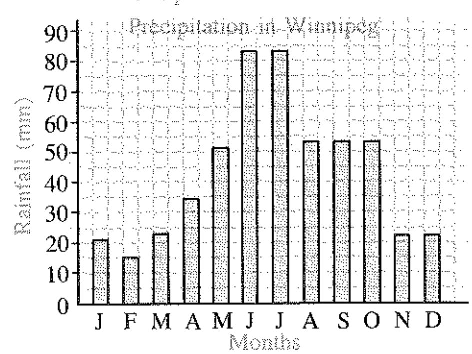
- What percent of Ontario's forest fires are caused by lightning?



Bar Graph

The graph shown here uses a vertical bar to represent the amount of precipitation each month. Graphs of this type are called *bar graphs.* Bar graphs have horizontal or vertical bars.

- What appears to be the rainy season in Winnipeg?



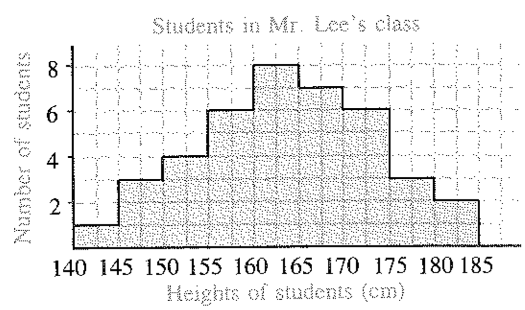
Histogram

A graph that uses bars, where each bar represents a range of values, is called a *histogram.*

The bars on a histogram do not have spaces between them because the data are continuous.

This histogram shows the number of students whose heights fall in each 5 cm interval from 140 cm to 189 cm.

- How many students are at least 170 cm tall?

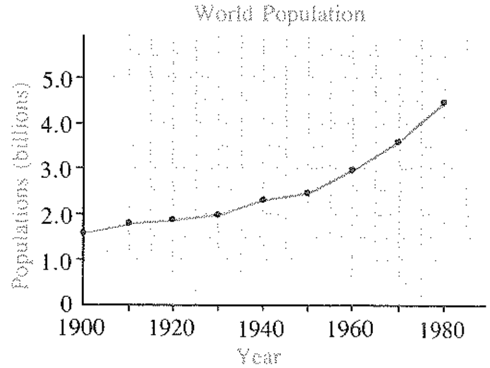


Broken-Line Graph

The graph shown here gives the population of the world at the end of each decade from 1900 to 1980. Since the exact population during each decade is not known, adjacent plotted points are joined by a line segment. Graphs like this are called *broken-line graphs*.

The only points on a broken-line graph that represent data are the endpoints of the segments.

- About how many years did it take the world's population to grow from 3 billion to 4 billion?



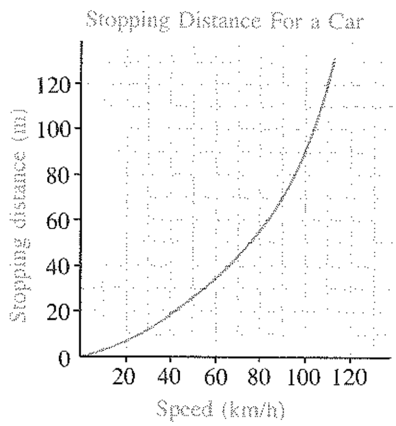
Continuous-Line Graph

A graph that shows the value of one variable, such as speed, corresponding to the value of another variable, such as stopping distance, for all values over a given interval is called a *continuous-line graph.*

This graph shows the distance required to bring a car to rest from the moment the brakes are applied, when the car is travelling at speeds up to 100 km/h.

All the points on a continuous-line graph correspond to data.

- What is the car's stopping distance when it is travelling at 60 km/h?



EXERCISES 11-1

(A)

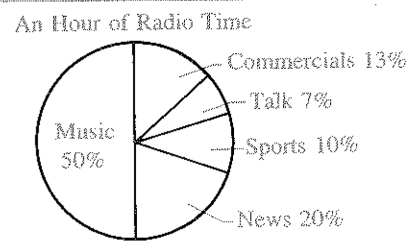
1. Name the type of graph that you think would be most appropriate for displaying each set of data. Explain why you chose that graph.
   a) Kevin's expenditures are divided as follows: 30% for entertainment, 40% for sports equipment, 22% for clothes, and 8% for school supplies.
   b) The approximate populations are given for Canada's 5 most populous provinces.
   c) The average temperature in Vancouver is given for each month of a particular year.
   d) A graph is to be drawn from which temperatures in degrees Celsius can be converted into temperatures in degrees Fahrenheit.
   e) The number of students whose final marks in mathematics were in these intervals: 0-25, 26-50, 51-75, and 76-100.

2. Explain the difference between:
   a) a pictograph and a bar graph
   b) a broken-line graph and a continuous-line graph
   c) a bar graph and a histogram.

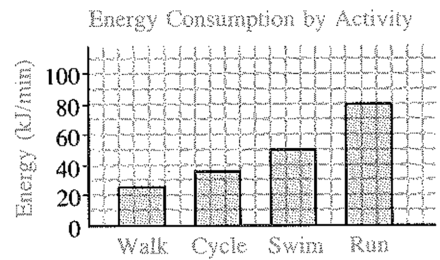
3. What is the sum of the percents shown on a circle graph? Explain your answer.

(B)

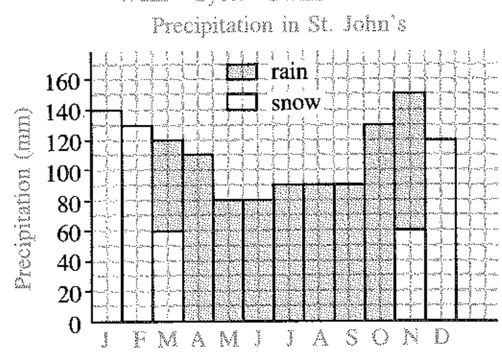
4. The circle graph shows how an hour of radio time is spent.
   a) What percent of each hour is devoted to news?
   b) How many minutes each hour are devoted to commercials?



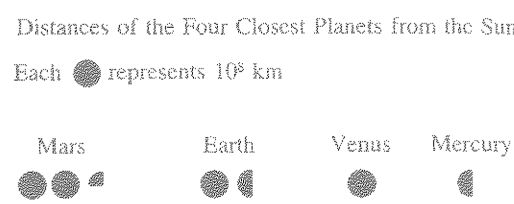
5. The bar graph shows the energy in kilojoules per minute used for different activities.
   a) Which activity burns up energy twice as fast as walking?
   b) If you cycled for 30 min, how much energy would you use?
   c) For how long would you have to run to burn up 2200 kJ of energy from a chocolate milkshake?

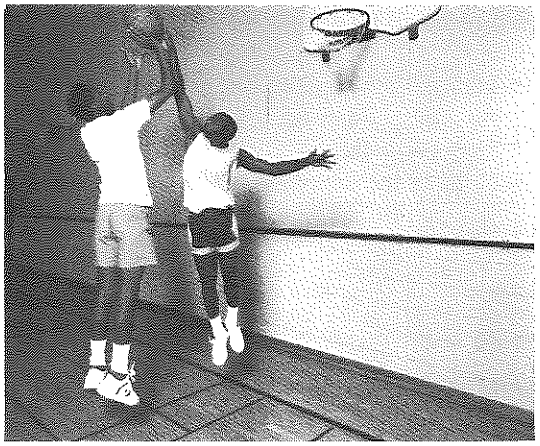


6. The histogram shows the monthly average amount of precipitation in St. John's.
   a) Which month had the most snow?
   b) Which month had the most rain?
   c) During which month did the snowfall equal the rainfall?
   d) What is the average snowfall in a year?



7. The pictograph shows the approximate distances from the sun to the four closest planets.
   a) How far, in kilometres, is each planet from the sun? Write your answers in scientific notation.
   b) About how many times as far from the sun is Mars than Mercury?
   c) Jupiter is about 780 million kilometres from the sun. Write this distance in scientific notation.
   d) What difficulty would you encounter if you tried to show the distances of all the planets from the sun using a pictograph?





11-2 ORGANIZING AND PRESENTING DATA

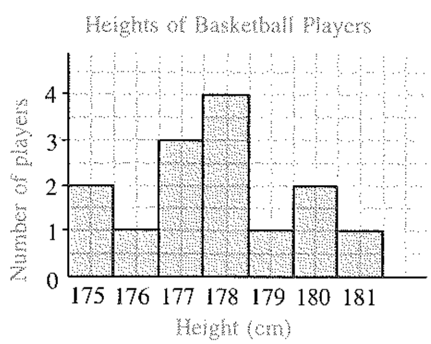
The heights of the players on a school basketball team were recorded by the coach.

Art	178 cm	Jason	177 cm	Neil	175 cm
Brian	181 cm	Joe	176 cm	Paul	178 cm
Bruce	180 cm	John	175 cm	Scott	177 cm
Dick	177 cm	Kevin	178 cm	Terry	178 cm
Gordon	180 cm	Larry	179 cm		

To get a better idea of the distribution of the heights of his players, the coach made a *tally chart* or *frequency table*. The frequency of a measurement is the number of times it occurs.

Using the frequency table, the coach drew a bar graph. Each height is represented by a bar. The length of a bar corresponds to the frequency of that measurement.

Height (cm)	Number of Players	Frequency
175	||	2
176	|	1
177	|||	3
178	||||	4
179	|	1
180	||	2
181	|	1



Often, the number of different data is too great for each measurement to be represented by a bar. Then the information is shown on a histogram.

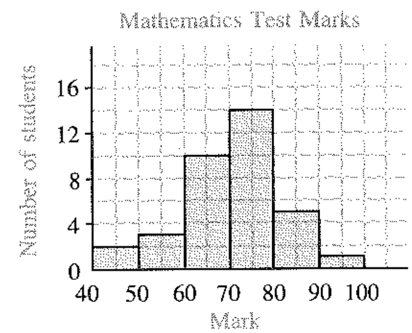
Example 1. Here is a set of marks (out of 100) obtained by a class on a mathematics test.

72	53	73	59	68	83	71	67	77	78	70	67	63
65	56	86	47	78	72	79	67	74	62	84	92	88
71	74	81	70	66	64	75	65	46				

a) Make a tally chart and frequency table for intervals of 10 marks.
b) Draw a histogram.
c) In which interval did most students' marks fall?
d) The pass mark was 50.
   i) How many students passed the test?
   ii) How many students failed the test?

Solution. a) From an inspection of the marks, the lowest mark is 46 and the highest mark is 92.
Choose intervals of 10 marks from 40 to 49, 50 to 59, . . ., 90 to 99. Make a tally chart and frequency table.

Interval	Number of Students	Frequency
40-49	||	2
50-59	|||	3
60-69	~~||||~~ ~~||||~~	10
70-79	~~||||~~ ~~||||~~ ||||	14
80-89	~~||||~~	5
90-99	|	1



b) Draw a set of axes. Label the horizontal axis with the intervals of marks. Because the intervals begin at 40, the axis between 0 and 40 is interrupted. Label the vertical axis with the numbers of students. Give the graph a title.
c) From the histogram, the interval in which most students' marks fall is the longest bar. That is, most students' marks fall between 70 and 79.
d) i) From the histogram, the number of students who passed the test is the total represented by the lengths of the bars for the interval 50 to 99.
   Students who passed: $3 + 10 + 14 + 5 + 1 = 33$
   33 students passed the test.
   ii) From the histogram, the students who failed have marks in the interval 40-49. Two students failed the test.

In *Example 1*, suppose the chosen intervals were 45-54, . . ., 85-94. What part of the example could not have been answered from an inspection of the histogram?

Example 2. The chart shows the number of students in each grade of a high school.

Grade	9	10	11	12
Students	266	248	230	142

a) Show this information on a circle graph.
b) What percent, to the nearest whole number, of the students are in:
i) grade 9 ii) grade 11?

Solution. a) The total student population is 266 + 248 + 230 + 142, or 886.
Each grade will be represented by a sector of the circle.
The sector angle for each grade is proportional to the number of students in that grade.
Express the student population of each grade as a fraction of the total student population, and multiply by 360°.

For grade 9 students, the angle is $\frac{266}{886}(360°)$, or 108.1°.

For grade 10 students, the angle is $\frac{248}{886}(360°)$, or 100.8°.

For grade 11 students, the angle is $\frac{230}{886}(360°)$, or 93.5°.

For grade 12 students, the angle is $\frac{142}{886}(360°)$, or 57.7°.

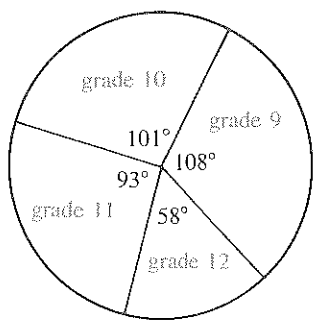
Write each sector angle to the nearest degree. Check to see if the angles add to 360°.

$108° + 101° + 94° + 58° = 361°$

Since several angles were rounded up, the total is 361°. To achieve a total of 360°, round down the angle for grade 11 students to 93°.

Draw a circle, mark the sector angles, and label each sector. Give the graph a title.

Number of Students in Each Grade



b) i) The grade 9 students represent $\frac{266}{886}(100\%)$, or about 30%.

ii) The grade 11 students represent $\frac{230}{886}(100\%)$, or about 24%.

EXERCISES 11-2

(A)

1. A car dealer hired students to determine the ages of cars owned by the residents of the area. The students listed their findings in this table.

Age of car (years)	0	1	2	3	4	5	6	7
Number of cars	25	40	50	65	45	30	20	15

Display the data on a bar graph.

2. The average daily temperature in Ottawa, for each month, is given in this table.

Month	J	F	M	A	M	J	J	A	S	O	N	D
Temperature (°C)	−11	−10	−3	6	13	18	21	19	14	8	1	−8

a) On a set of axes, plot a point for each temperature from January to December.
b) Join adjacent points with a straight-line segment to form a broken-line graph.

3. The composition of a hot dog is given in this table.

Ingredient	Water	Fat	Protein	Other
Mass (g)	20.4	11.0	4.2	2.3

Draw a circle graph to show this information.

(B)

4. The actual mass of a 2 kg box of chocolates was checked by weighing a selection of 345 boxes. Here are the results.

Mass (kg)	1.91-1.94	1.95-1.98	1.99-2.02	2.03-2.06	2.07-2.10
Frequency	15	85	75	150	20

a) Display this information in a histogram.
b) Quality control dictates that any box under 2 kg must be sold as "seconds".
   i) How many of these checked boxes must be sold as seconds?
   ii) What percent of the boxes are sold as seconds?
c) How could the intervals of mass be organized so that it is possible that fewer boxes would be sold as seconds?

5. A school's grade 9 students obtained these marks (out of 100) in an English examination.

55	66	64	98	56	69	68	62	52	69	65	63	51	90	69
68	32	66	72	44	80	61	84	74	66	79	61	89	78	63
66	59	75	53	69	23	92	78	73	67	38	65	67	41	75
63	71	57	77	66	56	63	73	24	56	76	71	61	51	46
84	55	63	68	86	65	69	66	60	62	68	82	73	65	76
79	88	44												

a) Make a tally chart and frequency table, using intervals of 10 marks.
b) Draw a histogram.
c) The pass mark is 60. How many students passed?
d) What percent of students had marks:
i) 60 or greater ii) less than 50?

6. The table shows the amount, as a percent, that a typical family spends in each category.
a) Draw a circle graph to illustrate this information.
b) Suppose the family's net annual income is \$30 000. How much is spent on:
i) recreation ii) transportation
iii) food iv) housing?

Item	Amount
Housing	40%
Food	24%
Transportation	15%
Clothing	9%
Savings	5%
Recreation	5%
Miscellaneous	2%

Ⓒ

7. Use the information in *Example 1*.
a) Draw a histogram with an interval of 5 marks.
b) Compare your histogram with that on page 408. What further information could be determined from your histogram?

8. Two groups of students wrote the same mathematics test. Group A was given notice of the test and was able to prepare for it. Group B had the test sprung upon them. The groups obtained the following marks out of 20.

A 16 17 18 20 11 18 20 19 15 20 15 15 17 12
19 8 13 16 17 14 19 14 20 12 15

B 12 11 15 18 12 6 9 11 5 11 11 14 11 16 17
12 13 9 8 19 10 10 7 11 18

a) Choose a suitable interval for the marks and draw a histogram for each set of data.
b) What conclusions can you draw from your histograms?

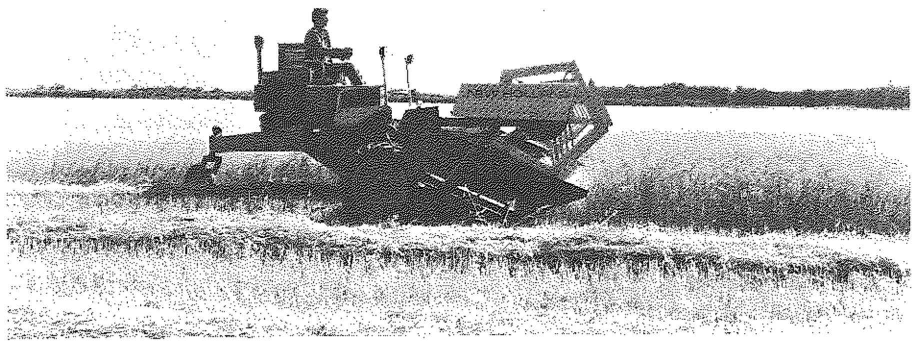
9. This table shows how Canadians are divided among various age groups.

Age group (years)	1-18	19-34	35-64	65 and over
Canadian population as a percent	32%	25%	32%	11%

a) Display this information on a circle graph.
b) Display this information on a histogram.
c) Which graph do you think is more useful? Explain your answer.
d) Assume a Canadian population of 26 000 000.
i) How many Canadians are under 35 years of age?
ii) How many Canadians are under 65 years of age?

MATHEMATICS AROUND US

Saskatchewan's Wheat Production



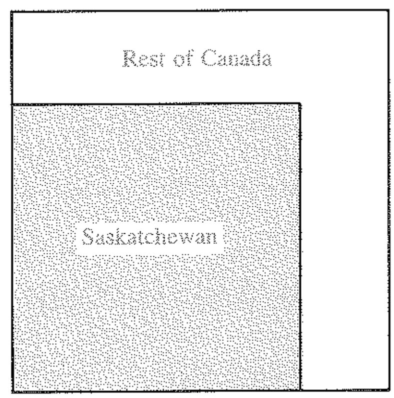
Saskatchewan produces more than half of Canada's wheat crop, as illustrated in the table. The wheat production is expressed as a percent of the total Canadian production.

Suppose Canada's wheat production is represented by the area of a 5 cm square. Then, Saskatchewan's wheat production may be represented by the area of a smaller square inside it.

The area of the large square is 25 cm^2. Therefore, the area of the small square is 0.57(25) cm^2, or 14.25 cm^2. The side length of the smaller square is $\sqrt{14.25}$ cm, or about 3.8 cm.

Saskatchewan's wheat production is represented by the area of a square of side length about 3.8 cm. This type of graph is called a *box graph*.

Provincial Wheat Production	
Saskatchewan	57%
Alberta	25%
Manitoba	12%
Other provinces	6%



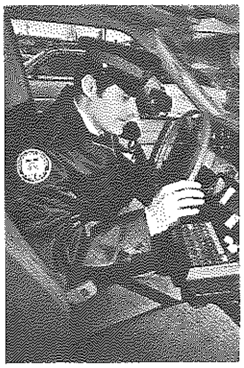
Wheat Production

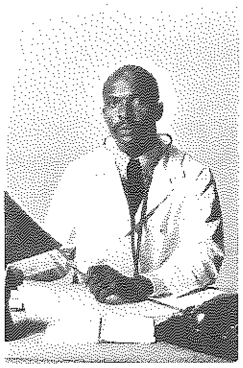
QUESTIONS

1. a) What would be the side length of the square representing the wheat production of:
   i) Alberta   ii) Manitoba?
   b) Draw the box graphs for part a).
2. Use the data in the table to draw a box graph to represent the population of:
   a) Alberta
   b) Manitoba
   c) Saskatchewan

Population of Canada 1986	
Alberta	2 373 000
Manitoba	1 075 000
Saskatchewan	1 020 000
Other provinces	21 034 000
Total	25 502 000

11-3 MEASURES OF CENTRAL TENDENCY


Police officer
$35 000


Doctor
$90 000


Lawyer
$60 000


Pilot
$55 000

The typical salary of each of four occupations is given above. Who earns more, doctors or lawyers?

It appears from the data given that doctors have greater incomes than lawyers. This can be misleading. Many lawyers earn more than $90 000 per year, and many doctors earn less than $60 000.

We want a single number that best represents the income of all doctors or all lawyers. We are looking for an "average" income.

Three of the most commonly used averages are the mean, the median, and the mode.

> The *mean* of a set of numbers is the arithmetical average of the numbers; that is, the sum of all the numbers divided by the number of numbers.

> The *median* of a set of numbers is the middle number when the numbers are arranged in order. If there is an even number of numbers, the median is the mean of the two middle numbers.

> The *mode* of a set of numbers is the most frequently occurring number. There may be more than one mode, or there may be no mode.

The mean, the median, and the mode of a set of numbers are referred to as *measures of central tendency*.

Example 1. The recorded rainfall, in millimetres, for seven consecutive days in Kitimat, British Columbia, is given. 12, 14, 8, 8, 8, 12, 15
Find.
a) the mean rainfall
b) the median rainfall
c) the mode for the rainfall

Solution. a) For the mean, add the numbers and divide by 7.

$$\frac{12 + 14 + 8 + 8 + 8 + 12 + 15}{7} = \frac{77}{7}$$
$$= 11$$

The mean rainfall is 11 mm.

b) For the median, arrange the numbers in order.
8, 8, 8, 12, 12, 14, 15
The middle value is the fourth value, 12.
The median rainfall is 12 mm.

c) The mode is the value which occurs most often. From the list in part b), 8 occurs most often.
The mode for the rainfall is 8 mm.

Example 2. The annual incomes for the people who work at the Beta Metal Works are shown below.
1 Manager: \$80 000; 1 Supervisor: \$45 000; 3 Mechanics: \$35 000; 5 Laborers: \$25 000
a) Determine the mean, the median, and the mode for the payroll.
b) Which measure could be used to make the salaries look:
i) high ii) low?
c) Which measure most fairly represents the average income in the company?

Solution. a) The mean salary, in dollars, is given by:

$$\frac{80\,000 + 45\,000 + (3 \times 35\,000) + (5 \times 25\,000)}{10}$$
$$= \frac{355\,000}{10}$$
$$= 35\,500$$

For the median, arrange the salaries in order.
80 000, 45 000, 35 000, 35 000, 35 000, 25 000, 25 000, 25 000, 25 000, 25 000
Since there is an even number of salaries, the median is the mean of the fifth and sixth values.

$$\frac{35\,000 + 25\,000}{2} = \frac{60\,000}{2}$$
$$= 30\,000$$

The mode is the salary that occurs most often, \$25 000.
The mean is \$35 500, the median is \$30 000, and the mode is \$25 000.

b) i) To make the salaries look high, the mean value of \$35 500 would be chosen as being representative.
ii) To make the salaries look low, the mode value of \$25 000 would be chosen as being representative.

c) The mean value of \$35 500 is earned by only 2 of the 10 employees. Therefore, it is a high representative value.
Since every employee earns at least the mode value of \$25 000, this is a low representative value.
The median value of \$30 000 probably best represents the average income.

EXERCISES 11-3

(A)

1. Find the mean, the median, and the mode of each set of data. Give the answers to 1 decimal place where necessary.
   a) 10, 12, 8, 9, 12, 14, 11, 15, 9, 12
   b) 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5
   c) 2.3, 4.1, 3.7, 3.2, 2.8, 3.6
   d) 15, 18, 16, 21, 18, 14, 12, 19, 11
   e) 9, 12, 7, 5, 18, 15, 5, 11
   f) $\frac{1}{2}, \frac{1}{4}, \frac{2}{3}, \frac{5}{12}, \frac{3}{4}, \frac{2}{3}, \frac{1}{2}, \frac{7}{12}, \frac{1}{6}, \frac{1}{2}$

(B)

2. Over a period of time, some shares were purchased as follows: 10 shares at \$8 per share; 20 shares at \$9.50 per share; and 15 shares at \$8.50 per share. What was the mean price per share?

3. For the numbers 5, 6, 7, 8, 9, find the effect on the mean:
   a) if each number is increased by 2 b) if each number is doubled.

4. If the mean of the numbers 8, 12, 13, 14, x, is 13, what is the value of x?

5. The mean of seven marks on a mathematics test is 68. However, the correction of an error in marking raises one student's mark by 14. Calculate the new mean.

6. The Cabot Manufacturing Company has the following employees at the rates of pay shown.

Position	Weekly Pay (\$)
1 President	1730
1 Designer	1150
1 Supervisor	770
3 Assemblers	670

Position	Weekly Pay (\$)
1 Secretary	580
1 Typist	480
3 Packers	385
3 Apprentices	290

   a) Find the measures of central tendency.
   b) Which measure of central tendency most fairly represents the pay structure of the company? Give reasons for your answer.

7. Two groups of students wrote the same mathematics test and obtained the following marks out of 20.

A 16 17 18 20 11 18 20 19 15 20 15 15 17 12
19 8 13 16 17 14 19 14 20 12 15

B 12 11 15 18 12 6 9 11 5 11 11 14 11 16 17
12 13 9 8 19 10 10 7 11 18

a) Calculate the three measures of central tendency for each group.
b) Calculate the measures of central tendency for the marks taken together.
c) How do the results of parts a) and b) compare?

8. The number of accidents at a ski resort for five months is given.
Which measure of central tendency best describes this data?

Dec.	Jan.	Feb.	Mar.	Apr.
25	35	40	35	5

9. Which measure of central tendency is the most suitable to describe each number?
a) the average number of children in a Canadian family
b) a person's average weekly salary
c) a class's average mark in a test
d) the average rainfall in Quebec City
e) the average time you spend on homework each night
f) the average number of hours a ten-year-old child spends watching TV shows each week
g) the average price of gasoline in a given area
h) the average size of shoes sold by a store

Ⓒ

10. Find five numbers that have:
a) a mean of 9 and a median of 8
b) a mean of 14 and a median of 8.

11. In a set of data, the smallest number is increased by 5 and the largest number is decreased by 5. What changes occur in the measures of central tendency?

12. Mary scored marks of 86, 82, 93, 97, and 78 on term tests in mathematics. What mark must she achieve on the next term test to have a mean score of 90? Explain your answer if the tests are marked out of 100.

13. Akira found that the mean of 100 numbers was exactly 26. Then he subtracted 26 from each of the 100 numbers and added the differences together. What total did he obtain?

14. Use the computer program on page xx to help you find the mean value of the first 100 unit fractions. Is the mean value greater than or less than the median value?

COMPUTER POWER

Calculating Means and Medians

The program below can be used to determine both the mean and the median of three or more numbers. To determine the mean, the computer adds the numbers as they are entered, and then divides the sum by the number of numbers. Then, to determine the median, the computer arranges the numbers in order. If the number of numbers is odd, the computer determines the middle number, and if it is even, the computer determines the mean of the two middle numbers.

```
100 REM *** CALCULATING MEANS AND MEDIANS ***
110 INPUT "HOW MANY NUMBERS ARE THERE? (MINIMUM 3): ";N
120 DIM X(N+1)
130 M=INT((N+1)/2):SUM=0
140 INPUT "ENTER THE FIRST NUMBER: ";X(1):SUM=X(1)
150 FOR I=2 TO N-1
160     INPUT "ENTER THE NEXT NUMBER: ";X(I)
170     SUM=SUM+X(I)
180 NEXT I
190 INPUT "ENTER THE LAST NUMBER: ";X(N):SUM=SUM+X(N)
200 FOR I=1 TO M+1
210     FOR J=I+1 TO N
220        IF X(I)>=X(J) THEN Y=X(I):X(I)=X(J):X(J)=Y
230     NEXT J
240 NEXT I
250 IF N/2<>INT(N/2) THEN X(M+1)=X(M)
260 PRINT:PRINT "THE MEAN IS: ";SUM/N
270 PRINT "THE MEDIAN IS: ";X(M)/2+X(M+1)/2
280 END
```

1. Use the program to find the mean and the median of each set of numbers.

a)

29	38	26	29	30	41
35	33	29	37	25	29
32	33	32	28	30	31
24	38	28			

b)

238	479	506	384	726
839	448	356	606	928
335	668	937	46	207
199	983	663	585	224
669	532	496	355	

2. The heights, in centimetres, of 32 students in a grade 9 class at Runnymede Collegiate were recorded as follows.

153.6	157.9	155.4	156.0	155.8	154.4	149.8	156.5	155.0
157.1	156.4	153.6	156.8	151.2	166.4	155.3	157.4	152.6
162.5	158.3	152.4	148.7	153.6	154.2	156.2	150.4	158.1
159.2	150.1	152.6	153.7	162.8				

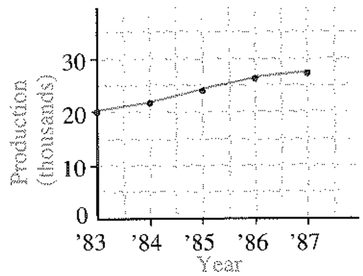
a) What are the mean and the median heights of the students in this class?
b) What is the height of the shortest student in the taller half of the class?

11-4 MISUSING STATISTICS

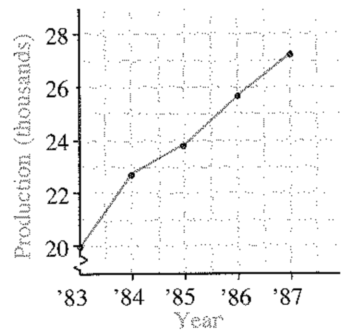
The table shows the number of cars produced by one plant of the Standard Automobile Corporation during a five-year period. These data were used to construct each broken-line graph shown below.

Does each graph convey the same impression?

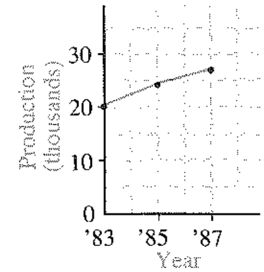
Year	Cars Produced
1983	20 150
1984	22 300
1985	23 850
1986	25 600
1987	27 100



Graph A



Graph B



Graph C

Graph A is an honest graph, which shows that the company is making moderate increases in production.

Graph B is misleading. There appears to be a spectacular increase in production over the five-year period. This occurs because the vertical scale does not begin at 0.

Graph C is misleading. This graph also suggests a greater increase in production than actually took place. This is achieved by making the vertical axis much longer than the horizontal axis. As a general rule, the axes should be approximately the same length.

Not all misuses of statistics involve graphs. The following examples contain statements that are typical of some that can be found in newspapers and magazines.

Example 1. Comment on the reasoning in this sentence.
"As few accidents happen in the early morning, very few of these as a result of fog, and fewer still as a result of travelling faster than 150 km/h, it would be best to travel at high speed on a foggy morning."

Solution. There are few accidents as a result of high speed on a foggy morning because:

- there is little traffic on the roads in the early morning
- many motorists would wait for the fog to clear, and those who do drive would reduce their speeds
- few motorists drive at more than 150 km/h at any time.

Therefore, it would be wrong to conclude from the statistics that it would be best to travel at high speed on a foggy morning.

Example 2. Comment on this newspaper report.

The Gazette

Two Out of Three Have Heart Trouble

In an interview with a Gazette reporter, Dr. Cole of Eastern Hospital stated that she was consulted by 30 patients last week. Of those 30, she found that 20 had had heart trouble. This is an alarmingly high

Solution. The headline is inaccurate because it is based only on the people who visited Dr. Cole. They very likely went to see her because they were not in good health. The 30 patients were not a representative portion of the population. Any representative portion of the population would include many more people than 30, most of whom would be healthy.

EXERCISES 11-4

Ⓑ

1. Say whether or not the following statements are correct interpretations of the statistics. Give reasons.
   a) A cure for the common cold has been found. In a recent test, 300 cold victims took the new wonder drug, Coldgone. After four days, only 7 persons still had colds.
   b) Last year, 50 motorists and 8 bicyclists were killed in traffic accidents. This proves that it is safer to ride a bicycle than drive a car.
   c) The data in the table show that teenagers are more likely to have a skiing accident than persons in other age groups. Also, persons over 50 years of age are the best skiers.

Age of Skier (years)	Percent of Skiing Accidents
under 10	10
10-19	61
20-29	19
30-39	5
40-49	3
over 50	2

   d) Almost 48% of injuries in downhill skiing are to skiers who have never had lessons. This means that skiers who take lessons are involved in more skiing accidents than those who do not. It follows that it is better not to take lessons.
   e) Last year, there were 45 000 job vacancies in Canada. There were also 500 000 unemployed workers. This shows that the unemployed are not interested in working.

2. You are the Public Relations Manager for an insurance company. A case is being made for cheaper insurance rates for persons under 25 years of age based on these statistics. What points would you make in reply?

Age Group	Number of Drivers in Fatal Accidents	Percent of Total Fatalities
under 25	98	39.2
over 25	152	60.8

3. Printed on the wrapper of a stick of gum is the statement, "4 out of 5 dentists recommend sugarless gum for their patients who chew gum." Comment on this statement.

4. The table records a company's annual profits for the years 1980 to 1987.
   a) Draw an honest graph to represent the data.
   b) Draw a graph on which the annual profit does not appear to change very much.
   c) Draw a graph that exaggerates the increase in profit.

Year	**Profit ($1 000 000s)**
1980	5.0
1981	5.2
1982	5.4
1983	5.8
1984	6.5
1985	7.0
1986	7.4
1987	8.5

5. The following table gives a company's monthly sales for a year.

Month	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Sales ($1000s)	95	98	98	95	90	85	80	78	75	72	68	64

   a) Draw an honest graph to represent the data.
   b) Draw a graph on which the monthly sales do not appear to change very much.
   c) Draw a graph that exaggerates the decrease in sales.

Ⓒ

6. Obtain three examples of statistical data from newspapers and magazines, or from radio and television commercials, and comment on the statements made.

7. The table shows the hitting records of two baseball players, Pops and Rooky, against left-handed and right-handed pitchers.

	Pops	Rooky
Against left-handed pitchers	180 hits out of 540 at bats	20 hits out of 100 at bats
Against right-handed pitchers	230 hits out of 460 at bats	420 hits out of 900 at bats

Each player was at bat 1000 times.
   a) Calculate the batting averages of Pops and Rooky. Whose average is greater?
   b) Calculate the batting averages of Pops and Rooky against left-handed pitchers. Whose average is greater?
   c) Calculate the batting averages of Pops and Rooky against right-handed pitchers. Whose average is greater?
   d) Compare the answers to parts a), b), and c). What do you notice?

MATHEMATICS AROUND US

Counting Large Populations

In 1986 it was reported that the Earth's population became 5 billion. But, as this newspaper article reveals, a question such as this can never be answered accurately. It takes so long to collect the data that they will be obsolete by the time they are all collected and analysed.

Earth's population estimated to hit 5 billion today

WASHINGTON (AP) — Somewhere on Earth today, the world's five billionth person will be born, say experts at the U.S.-based Population Institute.

If true, the new milestone will come just 10 to 12 years after the four billionth person checked in. But not all experts agree today is the day.

Carl Haub, a demographer at the private Population Reference Bureau, noted United Nations estimates indicate the five billion mark won't be reached until about next March. Other statisticians have said the milestone may have already quietly passed, since many countries simply do not keep very good track of their populations.

QUESTIONS

1. From an almanac or some other source, obtain the most recent value of the population of the Earth. How accurately is the value given?
2. Do you think there was actually a 5 billionth person? Discuss.
3. Suppose you were counting the people in your community.
   a) Do you think you could do this correct to the nearest person?
   b) Give as many reasons as you can why you might not be able to determine the population correct to the nearest person.
   c) How accurately do you think you could determine the population?
4. Try to determine the population of your school correct to the nearest person. Include both students and staff.

11-5 SAMPLING A POPULATION

What television program is the most popular of those on the air at a particular time?

It is clearly impossible to poll the entire population to find out who is watching what. Instead, a representative portion of the population, called a *sample*, is polled. If the sample is carefully chosen, the viewing preferences of the sample will accurately reflect the preferences of the entire population. For example, the Nielsen ratings which rank the popularity of television shows in North America are determined by a survey which samples less than one home in 10 000, that is, less than 0.01% of the North American population! Gallup polls assess the political preferences of the entire population of Canadians by surveying fewer than 2000 people.

If a sample is truly representative of the population from which it is drawn, then conclusions made about the population are likely to be valid. In statistics, the *population* is the whole of anything of which a sample is being taken.

A sample that is chosen in such a way that it is typical of the population it represents is called a *random sample*. It is very important that a sampling process be purely random in that all members of the population must share an equal chance of being selected.

To obtain information about a population, follow these steps.

- Decide on a sample size.
- Choose a device for selecting a random sample of that size.
- Collect the data from the sample.
- Organize and interpret the data.
- Make inferences about the characteristics of the population.

There are several ways of collecting data.

- Personal interviews — door-to-door, at shopping centres, by telephone
  For example, a roller-skate manufacturer needs to know how many roller skates of each size to make. The manufacturer would arrange to have personal interviews conducted at selected rinks.

- Questionnaires — by mail, with a purchased article, in newspapers
  For example, a politician wants to know how her or his constituents feel about an environmental issue. The politician would send questionnaires by mail to the constituents.
- Tests and measurements — recording instruments, quality control, time study
  For example, a quality-control engineer for a light bulb manufacturer wants to know the life of the light bulbs, and how many are defective. The engineer would conduct tests on a sample of light bulbs chosen at random from the production line.

Example. A Vancouver company is hired by a television station to conduct a poll to predict the outcome of a national election. To gather this information, the company considers sampling Canadian voters in one of the following ways.

a) Interview 100 people at random.
b) Poll a random sample of 1000 people in British Columbia.
c) Put an advertisement in all major newspapers asking people to tell their political preferences.
d) Send 10 questionnaires to all major businesses to be completed by anyone selected at random.

Describe the main weakness(es) of each method.

Solution.

a) The sample is probably too small to be reliable.
b) Political preferences are often regional in nature. A sample of voters in British Columbia is not likely to be representative of the political opinions of all Canadians.
c) Generally, only people with very strong political views will take the trouble to respond to an advertisement. The sample will not be random.
d) This sample tends to exclude such voting groups as students, farmers, homemakers, and senior citizens, and is therefore not a random sample.

EXERCISES 11-5

(A)

1. A student visited every household within three blocks of her home and recorded the number of persons in each household.

Number of persons in household	1	2	3	4	5	6	7
Number of households	3	7	12	21	14	8	5

a) How many households did she visit?
b) How many persons live within three blocks of the student?

2. The make of every third car in a full parking lot is noted and the number of each make is recorded.

Make of car	Ford	General Motors	Chrysler	American Motors	Foreign
Number of each make	47	64	29	25	82

a) Which is the most popular car?
b) How many cars were in the lot?
c) If the parking fee is \$3.50 per car, what are the total receipts for these cars?

B

3. How would you collect data to find the following information? Give reasons.
a) The popularity of a TV program
b) The most popular breakfast cereal
c) The average number of children in a family
d) The number of occupants in cars in rush hours
e) The most popular recording artist
f) The average family's food budget

4. How would you collect data to find the following information? What kind of people or items would be in your sample?
a) The player most likely to be voted "outstanding rookie"
b) The top 10 movies of the year
c) The life of flashlight bulbs
d) The amount spent by car owners on repairs each month
e) The time required to eat lunch in a cafeteria
f) The most popular soft drink

5. Explain why data are collected from a sample and not a population, for each situation.
a) The quality control in the manufacture of flash cubes
b) The number of pets per family
c) The purity of processed food
d) The strength of aluminum extension ladders
e) The cost of ski equipment
f) The percent of the population with the various blood types

6. Decide what kind of a sample you need, then work singly, in pairs, or in groups to collect the following data.
a) The age and the height of the students in your class
b) The number of persons in cars in the rush hour
c) The amount spent on lunch in the cafeteria
d) The most popular musical group
e) The time spent waiting in line in the cafeteria
f) The weekly earnings of students
g) The percent of times a thumbtack lands point up when dropped from a height of 25 cm
h) The number of letters in English words

11-6 SOURCES OF DATA

What percent of Canadian families have exactly two children?

There are basically two ways to answer this question. We could collect the data directly, or we could seek the data from some other source.

Collecting Data Directly

Since it is impractical to consider every Canadian family, it would be necessary to conduct a survey. We would identify a sample of Canadian families, and determine the number of children in each family. For the sample to be representative of the Canadian population, it would have to be determined randomly. The sample should contain families from all parts of Canada, and families with no children.

Data Accepted From Other Sources

An almanac is likely to contain population data which could be used to answer the question. If the data are based on census figures, they would have been determined by counting all the families in Canada, and not just a sample. Since a census is taken only once every ten years, the data may be somewhat out of date.

Possible sources of data

- almanac
- encyclopedia
- newspaper files
- computer data base
- library
- Statistics Canada publications
- publications of provincial or municipal governments

EXERCISES 11-6

(A)

1. Suppose you were estimating the percent of Canadian families with two children. What are the advantages and disadvantages of obtaining the data:
   a) directly
   b) from some other source?

B

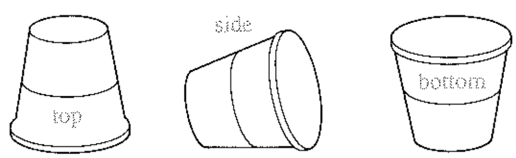
2. Conduct a survey in your class or your school to determine what percent of the students have at least one brother or sister.
   a) Do you think this would be a reasonable estimate of the percent of Canadian families with two children? Explain.
   b) Is the actual percent of Canadian families with two children likely to be higher or lower than your result in part a)?
   c) Check your prediction in part b) by consulting an almanac.

3. To answer each question, would you seek data from another source? If so, indicate what source you would consult. Or, would you collect the data directly? If so, indicate how you might collect the data.
   a) What fraction of Canadians are under 18 years of age?
   b) What fraction of students in your class have a job during the school year?
   c) What is the world record for the women's high jump?
   d) How much would it cost to buy $100 in U.S. funds?
   e) What percent of the meals eaten by Canadian teenagers are not eaten at home?
   f) What percent of the cars on the road have personalized licence plates?
   g) How many telephone calls does the average person make in one year?
   h) How many television stations can be received in your community?
   i) What percent of the Canadian population was born in Canada?
   j) How many people in your community support the political candidate you are campaigning for?

4. Obtain the following data.
   a) What are the dimensions of a Canadian two-dollar bill?
   b) What is the population of Canada?
   c) What is the current price of gold?
   d) What time does the sun set in your community today?
   e) How many radio stations are there in your city or the city closest to your home?
   f) How many traffic signals are there within 1 km of your school?

INVESTIGATE

1. Toss a paper cup 20 times.
2. Record the number of times it lands on its top, its bottom, and its side.



3. Add the results of everyone in the class.
4. Write the number of times each outcome occurs as a fraction of the total number of tosses. This fraction is called the *relative frequency* of each outcome.
5. Discuss the results.

11-7 PREDICTING RESULTS

One of the principal uses of statistics is in predictions.

By studying samples of voter opinions, we can forecast election outcomes. These cannot be controlled. We cannot say for certain what will happen. Nevertheless, sampling enables us to assess the likelihood that a particular outcome will occur.

Consider the previous *INVESTIGATE*, where a paper cup was tossed 20 times. Suppose the cup landed on its side 12 times. We say that the relative frequency of the outcome "landing on its side" is $\frac{12}{20}$, which simplifies to $\frac{3}{5}$, or 0.6.

$$\text{Relative frequency of an outcome} = \frac{\text{Number of times the outcome occurs}}{\text{Total number of outcomes}}$$

Example. In an experiment, a paper cup is tossed 400 times. Here are the results.

Outcome	top	side	bottom
Frequency	106	246	48

a) Find the relative frequency of each outcome.
b) Predict how many times the cup would land on the bottom in 1000 tosses.

Solution. a) Relative frequency of an outcome $= \frac{\text{Number of times the outcome occurs}}{\text{Total number of outcomes}}$

$$\text{Relative frequency of landing on the top} = \frac{106}{400} = 0.265$$

$$\text{Relative frequency of landing on the side} = \frac{246}{400} = 0.615$$

$$\text{Relative frequency of landing on the bottom} = \frac{48}{400} = 0.12$$

b) To predict the number of times a particular outcome will occur, multiply its relative frequency by the number of tosses.
The number of times a cup will land on the bottom in 1000 tosses is 1000(0.12), or 120.

Jane, the owner of an art shop, decided to make and sell sheets of adhesive letters used for notices and posters. She needed to know how many of each letter to put on a sheet of 500 letters.

To find the frequency with which each letter occurs in the English language, she examined a large sample of poetry. Jane chose the first three verses of *The Tiger* by William Blake.

Tiger, tiger, burning bright
In the forests of the night,
What immortal hand or eye
Could frame thy fearful symmetry?

In what distant deeps or skies
Burnt the fire of thine eyes?
On what wings dare he aspire?
What the hand dare seize the fire?

And what shoulder and what art
Could twist the sinews of thy heart?
And, when thy heart began to beat,
What dread hand and what dread feet?

Jane made a tally chart and frequency table for the letters. She counted the total number of letters in the sample. There are 301 letters. Here is the tally chart for 4 letters.

Letter	a	e	n	s
Tally	卌 卌 卌 卌 卌 ||||	卌 卌 卌 卌 卌 卌 卌 ||||	卌 卌 卌 卌	卌 卌 卌
Frequency	29	39	20	15

Jane found the percent of each letter by multiplying its relative frequency by 100%. For example,

Percent of letter a is $\frac{29}{301}(100\%)$, or about 10%.

Hence, letter a should be 10% of the sheet.

$$\begin{aligned} 10\% \text{ of } 500 &= 0.10(500) \\ &= 50 \end{aligned}$$

There should be about 50 of letter a on the sheet of letters.

Jane repeated these calculations for each of the letters in the sample poetry.

Would a different poem, or a sample of prose give a different result?

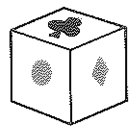
Some letters of the alphabet did not appear in this sample of poetry. Should these letters be included on the sheet?

EXERCISES 11-7

Ⓐ

1. A computer simulated the toss of a penny 319 020 times. Heads occured 160 136 times. What was the relative frequency of heads, to 3 decimal places?

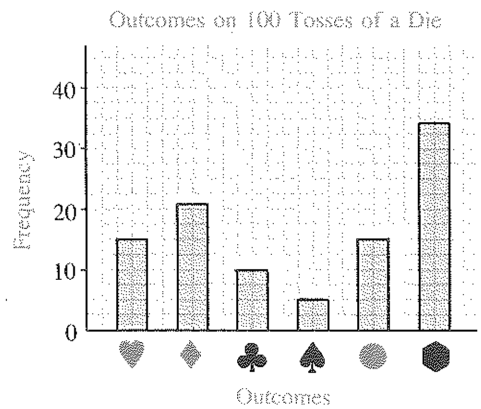
2. A die has these faces.



The die was rolled 100 times. The frequency of each outcome is shown on the graph.

a) Find the relative frequency of each face shown below.

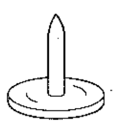
i) ii) iii)



b) Do you think it is a "fair" die? Explain your answer.

Ⓑ

3. When a thumbtack is tossed, there are two possible outcomes.


point up


point down

a) Toss 10 thumbtacks onto a desk. Record the number that land point up, and calculate the relative frequency.
b) Combine the results of everyone in the class.
c) What is the relative frequency for point up after tossing 300 thumbtacks?

4. A thumbtack is tossed 400 times and lands point up 250 times. About how many times should it land point up if it is tossed 5000 times?

5. a) Toss a coin the number of times indicated and record the frequency of heads.
i) 10 times ii) 20 times iii) 30 times
b) Calculate the relative frequency of heads in each case.
c) Combine your results with those of other students to obtain the relative frequency of heads for a greater number of tosses.
d) How does the relative frequency of heads compare with 0.5 for a greater number of tosses?

6. If, in a coin-tossing experiment, you calculated the relative frequency of heads to be 0.47, what should be the relative frequency of tails?

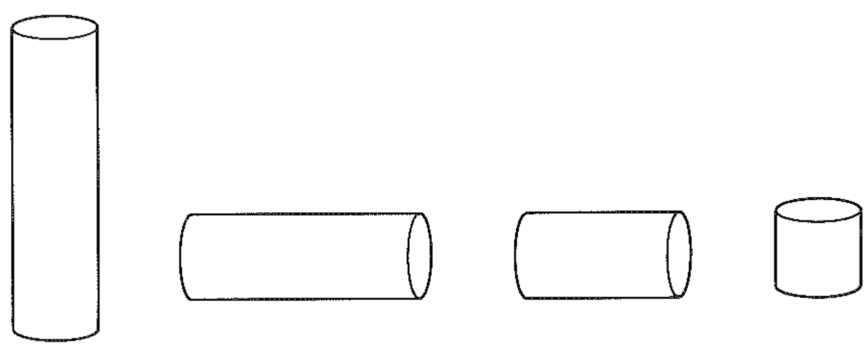
7. Choose 200 lines from a magazine story or newspaper article. Count the number of complete sentences and the number of words in each sentence. What is the relative frequency of sentences containing:
   a) fewer than 9 words b) more than 12 words?

8. a) Toss two coins 30 times and record the number of times they show:
      i) two heads ii) two tails iii) one head, one tail.
   b) Calculate the relative frequency in each case.
   c) Combine your results with those of other students to find the relative frequencies for a greater number of tosses.
   d) If two coins were tossed 5000 times, about how many times would they show:
      i) two heads ii) two tails iii) one head, one tail?

Ⓒ

9. By the end of the second week of the baseball season, a player has had 9 hits out of 20 official times at bat.
   a) Calculate the player's batting average.
   b) In the next game, the player gets 0 hits out of 3 times at bat. Calculate his batting average after this game.
   c) By the final month of the baseball season, the player has had 106 hits out of 425 times at bat. Calculate his batting average now.
   d) The player gets 0 hits out of his next 3 times at bat. What does this make his average?
   e) Why did a game with 0 hits out of 3 times at bat make less difference to the player's batting average at the end of the season than at the beginning?

10. When a cylinder is tossed there are two possible outcomes; it can land on an end or on its side. From a broom handle, cut cylinders 1 cm, 2 cm, 3 cm, and 4 cm long. Record the outcomes of 50 tosses for the four cylinders.



   a) What is the relative frequency of the cylinder landing on an end?
   b) What is the effect of the length to diameter ratio on the way a cylinder lands?

11. You intend to toss a coin 100 times to determine the relative frequency of heads. Investigate whether it makes a significant difference if you toss:
   a) 1 coin, 100 times b) 2 coins, 50 times
   c) 4 coins, 25 times d) 10 coins, 10 times.

COMPUTER POWER

Counting Characters

Since compiling frequencies and relative frequencies is tedious and time consuming, the computer is now used for this purpose. The following program in BASIC will do this for ten or fewer letters in any passage.

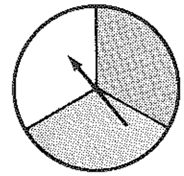
```
100 REM *** COUNTING CHARACTERS ***
110 INPUT "HOW MANY LINES IN YOUR PASSAGE OF TEXT? ";A
120 DIM Q$(10),X$(A),Y$(A,80),R(A),Q(A,10)
130 PRINT "HOW MANY DIFFERENT LETTERS DO YOU WISH"
140 INPUT "TO COUNT? ";N
150 FOR F=1 TO N
160     PRINT "ENTER LETTER NUMBER ";F;" TO BE COUNTED"
170     INPUT Q$(F)
180 NEXT F
190 FOR K=1 TO A
200     PRINT "ENTER LINE #";K
210     INPUT X$(K)
220     FOR J=1 TO LEN(X$(K))
230         Y$(K,J)=MID$(X$(K),J,1)
240         IF Y$(K,J)=" " THEN GOTO 290
250         R(K)=R(K)+1
260         FOR I=1 TO N
270             IF Y$(K,J)=Q$(I) THEN Q(K,I)=Q(K,I)+1
280         NEXT I
290     NEXT J
300     R=R+R(K)
310     FOR L=1 TO N
320         S(L)=S(L)+Q(K,L)
330     NEXT L
340 NEXT K
350 PRINT:PRINT "REL. FREQ.","FREQ.","LETTER"
360 FOR M=1 TO N
370     PRINT S(M)/R,S(M),Q$(M)
380 NEXT M
390 END
```

This program counts all characters, but not spaces. Do not enter punctuation marks or the results may be inaccurate. After entering each line of text, press $\boxed{\text{RETURN}}$ and wait for the computer to ask for the next line.

1. Use the program to obtain the frequency and the relative frequency of each letter a, e, i, n, o, s, and t in:
   a) the three verses of "The Tiger" on page 428 (do not enter any commas)
   b) the first paragraph of this feature.

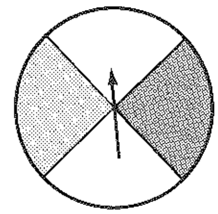
   Explain the results.

11-8 PROBABILITY



When spun, the pointer on this wheel may stop on white, black, or red. Since the three portions are the same size, we say that each outcome is *equally likely*.

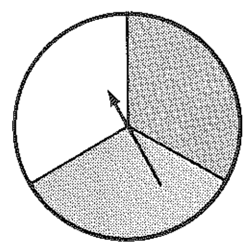
Example 1. For each experiment
i) List the outcomes.
ii) State whether the outcomes are equally likely.
a) A penny is tossed.
b) A penny and a dime are tossed.
c) The pointer on the wheel shown is spun.



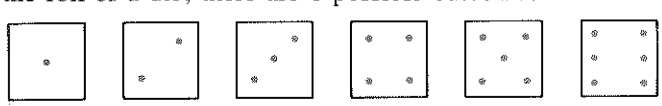
Solution. a) i) The outcomes are head (H) and tail (T).
ii) They are equally likely.
b) i) The outcomes are: H H, H T, T H, and T T.

ii) They are equally likely.
c) i) The outcomes are: red, black, and white.
ii) Since there are two white portions and only one black and one red, the pointer is more likely to stop on white. The outcomes are not equally likely.

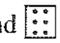
For the wheel shown, since each outcome is equally likely, there is one chance in three that the pointer will stop on black.
We say that the probability that the pointer will stop on black is $\frac{1}{3}$, and we write: P(black) $= \frac{1}{3}$.



When a die is rolled, the chances that it will show ⚁ are 1 in 6.
We say that the probability that the die will show a 2 is $\frac{1}{6}$, and we write: $P(2) = \frac{1}{6}$.
Any set of outcomes of an experiment is called an *event*.
On the roll of a die, there are 6 possible outcomes.



Let A denote the event that we get an even number when we roll a die.
Event A occurs if the outcome of a toss is ⚁ or ⚃ or ⚅.

We say that the outcomes , , and  are *favorable* to event A, because if any one of them occurs then event A occurs.

On the toss of two coins, there are 4 possible outcomes.

Let B denote the event that we obtain one head and one tail. Then, the outcomes H T and T H are favorable to event B.

The greater the number of equally likely outcomes favorable to an event, the more likely it is that the event will occur.

> If the outcomes of an experiment are equally likely, then the probability of an event A is given by:
>
> $$P(A) = \frac{\text{Number of outcomes favorable to A}}{\text{Total number of outcomes}}$$

Example 2. A lottery issued 1000 tickets, which were all sold. What is the probability of your winning if you hold:

a) 1 ticket b) 17 tickets c) 100 tickets?

Solution. Each ticket has an equal chance of being drawn. That is, each outcome is equally likely.

a) $P(\text{win}) = \frac{1}{1000}$ b) $P(\text{win}) = \frac{17}{1000}$ c) $P(\text{win}) = \frac{100}{1000}$, or $\frac{1}{10}$

Example 3. A jar contains 3 black balls, 4 white balls, and 5 striped balls. If a ball is picked at random, what is the probability that it is:

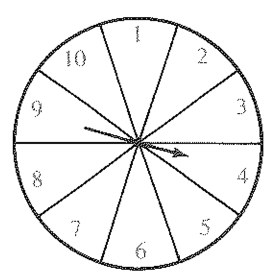
a) black b) white c) striped?

Solution. There are 12 balls and each one has an equal chance of being picked. Each outcome is equally likely.

a) $P(\text{black}) = \frac{3}{12}$ b) $P(\text{white}) = \frac{4}{12}$ c) $P(\text{striped}) = \frac{5}{12}$

Example 4. For the wheel shown, determine the probability of each event.

a) Event A: Wheel stops on a number equal to or less than 3.
b) Event B: Wheel stops on a number greater than 6.
c) Event C: Wheel stops on an even number.



Solution. There are ten equally likely outcomes.

a) Event A has three favorable outcomes. These are landing on 1, 2, or 3.

$$P(A) = \frac{3}{10}$$

b) Event B has four favorable outcomes. These are landing on 7, 8, 9, or 10.

$$P(B) = \frac{4}{10}$$

c) Event C has five favorable outcomes. These are landing on 2, 4, 6, 8, or 10.

$$P(C) = \frac{5}{10}$$

Example 5. For the wheel in *Example 4*

a) Event D is that the wheel stops on a number from 1 to 10. What is P(D)?

b) Event E is that the wheel stops on a number greater than 10. What is P(E)?

Solution. a) Every possible outcome is favorable to event D.

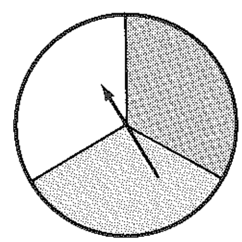
Therefore, $P(D) = \frac{10}{10}$, or 1

b) No outcome is favorable to event E.

Therefore, $P(E) = \frac{0}{10}$, or 0

Probability and relative frequency are closely linked. The probability of an event indicates what the relative frequency should be if the experiment is performed many times.

For the wheel shown, we know that $P(\text{black}) = \frac{1}{3}$. This does not mean that in 30 spins the pointer will stop on black exactly 10 times. It means that it is more probable that it will stop on black 10 times, than any other number of times.



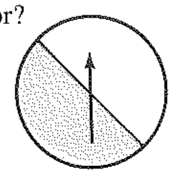
Example 6. For the wheel above, what is the probable number of times that the pointer should stop on black in 2000 spins?

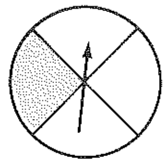
Solution. Since $P(\text{black}) = \frac{1}{3}$, the spinner should probably stop on black about $\frac{1}{3}(2000)$ times, or about 667 times.

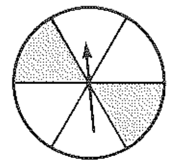
EXERCISES 11-8

(A)

1. When the pointer is spun, what is the probability that it will stop in the colored sector?

a) 

b) 

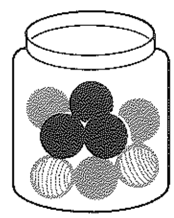
c) 

2. In the SCRABBLE™ Brand crossword game, the letters of the alphabet are distributed over the 100 tiles as shown in the table.

   From a full bag of tiles, what is the probability of randomly selecting:
   a) B b) E c) S?

Distribution of Tiles					
A-9	F-2	K-1	P-2	U-4	Z-1
B-2	G-3	L-4	Q-1	V-2	Blank-2
C-2	H-2	M-2	R-6	W-2	
D-4	I-9	N-6	S-4	X-1	
E-12	J-1	O-8	T-6	Y-2	

3. What is the probability that a ball chosen at random from the jar is:
   a) black b) striped
   c) colored d) not black
   e) either black or colored f) neither black nor colored



(B)

4. Cathy and Trevor have birthdays in June. Let A denote the event that Cathy's birthday is a multiple of 5. Let B denote the event that Trevor's birthday is on Wednesday. Let C denote the event that A and B are both true and both birthdays occur on the same day.
   a) If all days are equally likely, what outcomes are favorable to:
      i) event A ii) event B iii) event C?
   b) Find the probabilities of event A, event B, and event C.

JUNE						
Sun	Mon	Tues	Wed	Thurs	Fri	Sat
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

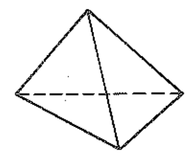
5. For each experiment
   i) List the outcomes.
   ii) State whether the outcomes are equally likely.
      a) A ball is drawn from a bag containing a red ball, a white ball, and a green ball.
      b) A quarter and a nickel are tossed.
      c) A letter is picked at random from any page of a book printed in English.
      d) A ball is drawn from a bag containing 2 white balls, 3 blue balls, and 5 red balls.
      e) A wheel containing the letters A to H is spun.

6. Five hundred tickets are printed for a lottery. Carla bought 7 tickets. What is the probability of her winning if:
   a) all the tickets were sold
   b) 370 tickets were sold and the rest destroyed?

7. A traffic light is red for 30 s, green for 25 s, and orange for 5 s in every minute. What is the probability that the light is orange when you first see it?

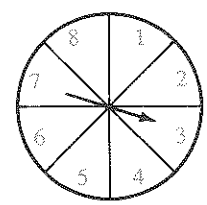
8. A pair of opposite faces of a white die are colored red. If the die is tossed, what is the probability that the top is:
   a) white
   b) red?

9. Calculate the probability of tossing a regular tetrahedron so that it lands with the 4 face down. The numbers on the faces are:
   a) 2, 4, 6, 8
   b) 1, 4, 4, 7
   c) 1, 3, 5, 7.



10. What is the probability of receiving a $3 bill in change when groceries are purchased at a store?

11. When the pointer on this wheel is spun, what is the probability that it will stop on :
   a) an odd number
   b) an even number
   c) a one-digit number
   d) a two-digit number?

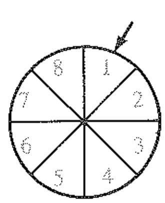
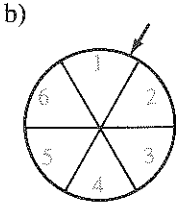
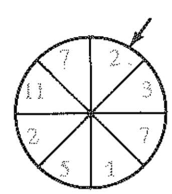
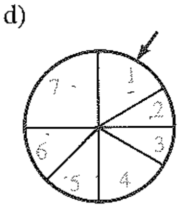


12. What is the probability of a regular die, when tossed, showing:
   a) 5
   b) an odd number
   c) a prime number
   d) a number less than 3
   e) a one-digit number
   f) a two-digit number?

13. What is the probable number of times that a coin should show heads if it is tossed:
   a) 25 times
   b) 100 times
   c) 1000 times?

14. What is the probable number of times that a die should show 5 if it is tossed:
   a) 25 times
   b) 100 times
   c) 1000 times?

15. Each wheel is spun 50 times. What is the probable number of times that the wheel should come to rest with the arrow pointing to 2?

a)

b)

c)

d)


16. What is the probability that all the students in your class are older than 10 years of age?

17. You can pick one marble from any of the three bags. You win a prize if you pick a red marble. Which bag should you choose to have the best chance of winning?
Bag A contains 3 red and 7 white.
Bag B contains 2 red and 3 white.
Bag C contains 4 red and 11 white.

18. The words STATISTICS AND PROBABILITY are spelled out with SCRABBLE™ Brand crossword tiles. Then, these tiles are put in a bag. What is the probability that a tile drawn from the bag at random will be:
a) a vowel
b) a consonant
c) one of the first 10 letters of the alphabet?

19. The table lists the number of cars in a parking lot by their ages.

Car's age (years)	0	1	2	3	4	5	6	7
Number	25	40	50	70	45	35	20	15

Calculate the probability that the age of a car selected at random will be:
a) 2 years
b) greater than 4 years
c) less than 3 years
d) 3 to 5 years.

20. What is the probability that a card drawn at random from a deck of 52 cards will be:
a) red
b) a spade
c) a black 7
d) a face card (Jack, Queen, or King)?

Ⓒ

21. A die is loaded so that the outcomes have the relative frequencies shown in this table.

Outcome	1	2	3	4	5	6
Relative Frequency	0.12	0.17	0.17	0.08	0.35	0.11

What is the probability of throwing:
a) a number less than 3
b) an even number?

22. Life insurance companies use birth and death statistics in calculating the premiums for their policies. The table shows how many of 100 000 people at age 10 are still living at ages 30, 50, 70, and 90.

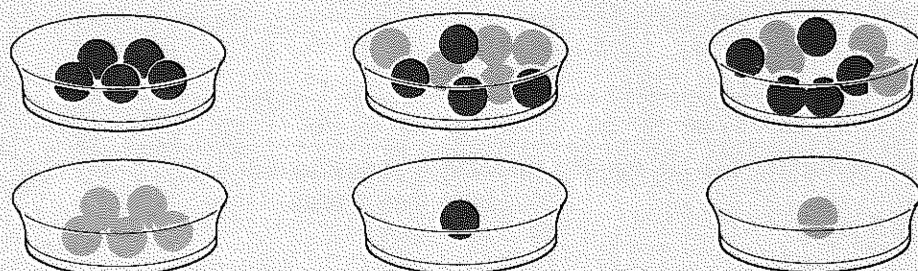
Age (years)	10	30	50	70	90
Number of People Living	100 000	95 144	83 443	46 774	2220

a) What is the probability that a 10-year-old child will live:
i) to age 50 ii) to age 70 iii) to age 50 but not 70?
b) What is the probability that a 30-year-old person will live to age 90?

PROBLEM SOLVING

Conduct An Experiment

"Each of you", said the judge to the three prisoners before her, "will be given 5 black balls and 5 colored balls. You may distribute these any way you please between two pans. Then you will be blindfolded while the 2 pans are moved around and then you must select a ball from one of the pans. If the ball you select is colored you will be freed. However, if it is black you will be returned to prison."



Estimate, for each distribution shown above, the probability that a colored ball will be chosen.

Understand the problem

- Is the probability that a colored ball will be chosen the same for all the distributions?

Think of a strategy

- Try conducting an experiment with the balls placed as illustrated above.

Carry out the strategy

- Get 5 yellow tennis balls and 5 white tennis balls.
- Duplicate the first distribution shown above.
- Conduct the experiment described by the judge. Repeat the selection many times and estimate the probability of selecting a colored ball.
- Repeat the procedure for the other two distributions shown above.
- For which distribution is it most likely that a colored ball will be chosen?

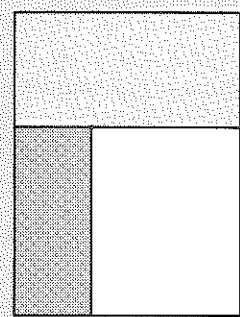
Look back

- Is there a distribution different from those shown above for which it is more likely that a colored ball will be chosen?

Solve each problem

1. Twelve pennies are placed in a row on a table top. Then every second coin is replaced with a nickel. Every third coin is then replaced with a dime. Finally every fourth coin is replaced with a quarter. What is the total value of the 12 coins on the table?

2. Divide a sheet of paper into 3 rectangles of different sizes. Color each rectangle a different color. Toss a penny many times from a distance and record the frequency of landing in each rectangle. Use these frequencies to estimate the area of each rectangle as a fraction of the area of the sheet of paper. Measure the dimensions and calculate the areas to check your answers.

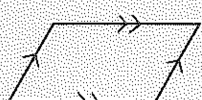


3. Estimate the probability that two cards drawn randomly from a deck of 52 playing cards will both be spades.

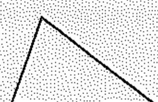
4. Estimate the probability that the sum of the numbers obtained on two rolls of a pair of dice exceeds 8.

5. Which of the following figures form a tesselation of the plane; that is, a covering of the plane without gaps or overlapping?

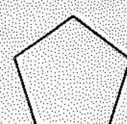
a)



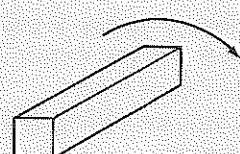
b)



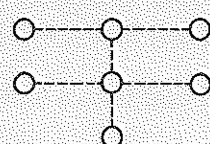
c)



6. A wooden beam with a square cross-section is rolled along the floor. Sketch the path of one corner of the beam.



7. Seven pennies are arranged in 3 lines, with 3 pennies in each line.
Sketch a diagram to show how 2 of these pennies can be moved so that there are 6 lines with exactly 3 pennies in each line.



THE MATHEMATICAL MIND

Games of Chance

We know from dice found in the tombs of ancient Greeks and Egyptians that games of chance have been played for thousands of years. However, it was not until the sixteenth and seventeenth centuries that a serious attempt was made to study games of chance using mathematics.

Chevalier de Méré, a professional gambler and amateur mathematician, had many questions about dice probabilities. He turned to the great mathematician Blaise Pascal for the answers. Pascal, with his friend Pierre de Fermat, began a systematic study of games of chance. The theory of probability was founded.

One of de Méré's questions was, "What is the probability of throwing two dice and *not* getting a 1 or 6?"

Pascal answered, "For each die, the probability is $\frac{4}{6}$, or $\frac{2}{3}$. For both, the probability is $\frac{2}{3} \times \frac{2}{3}$, or $\frac{4}{9}$ — about 0.44."

With this information, de Méré offered the equivalent of this gamble.

> Bet $1. Throw 2 dice. If ⚀ or ⚅ do NOT show, you win $2.

He now knew that for every 100 people who played the game, about 44 would win. That meant he would take in $100 and pay out $88. He could expect to win about $12 every time 100 people played. Now that you know this, would you spend $1 to play this game?

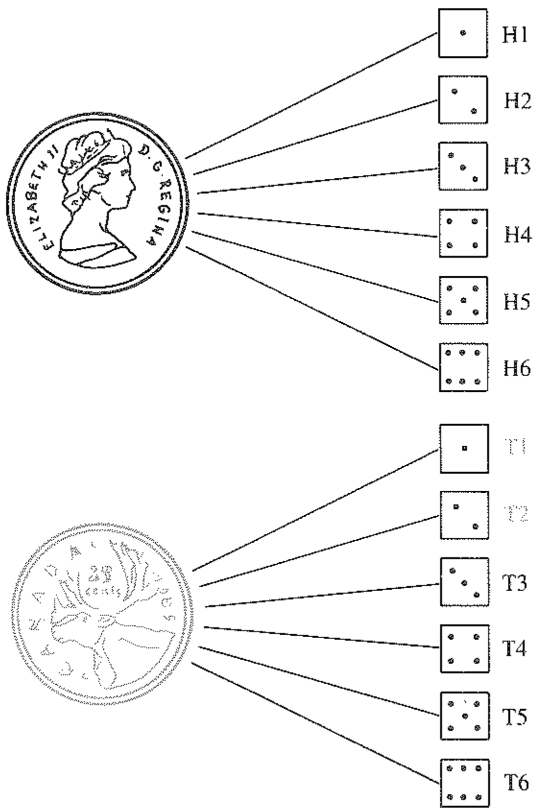
QUESTIONS

1. For each of the following games of chance
   a) Determine whether you can expect to win, or lose, money if you play the game a great number of times.
   b) Decide whether you are willing to play the game.
   c) Explain your decision.
      i) Bet $1. Toss a coin. If it shows a head, you win $2.
      ii) Bet $1. Draw a card from a well-shuffled deck. If it shows a spade, you win $5.
      iii) Bet $1. Draw a card from a well-shuffled deck. If it shows an ace, you win $10.
      iv) Bet $1. Toss two coins. If they show two heads, you win $3.

11-9 THE PROBABILITY OF SUCCESSIVE EVENTS

Suppose a coin is tossed and a die is rolled. What is the probability that the coin shows tails *and* the die shows a number less than 3?

To answer this question, we can draw a tree diagram. This illustrates all the possible outcomes of tossing a coin and rolling a die. That is, for each side of the coin, there are 6 faces of the die.



The 12 outcomes are listed at the side of the tree diagram. The outcomes that show tails with a number less than 3 are T1 and T2. These are the favorable outcomes.

The probability of a coin showing tails and a die showing a number less than 3 is

$$\text{P(tails and less than 3)} = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$$

$$= \frac{2}{12}$$

$$= \frac{1}{6}$$

Another way of finding P(tails and less than 3) is to notice that the number of outcomes involving tails is $\frac{6}{12}$, or $\frac{1}{2}$ the total number of outcomes. Also, the number of outcomes involving the die showing less than 3 is $\frac{4}{12}$, or $\frac{1}{3}$ the total number of outcomes. Hence, the outcomes for tails *and* the die showing a number less than 3 are $\frac{1}{2}$ of $\frac{1}{3}$ the total outcomes.

$$\begin{aligned} P(\text{tails and less than 3}) &= \frac{1}{2} \times \frac{1}{3} \\ &= \frac{1}{6} \end{aligned}$$

> The probability of two (or more) events happening in succession is the product of the probability of each event.
> P(A and B) = P(A) × P(B)

Example 1. A coin is tossed three times. What is the probability that it shows a head each time?

Solution. For each toss, P(head) $= \frac{1}{2}$

$$\begin{aligned} P(3\text{ heads}) &= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \\ &= \frac{1}{8} \end{aligned}$$

The rule for successive events must be used carefully, because the probability of the second event may depend on the first event.

Example 2. A bag contains 2 black balls and 2 red balls. Find the probability of drawing 2 red balls in succession if:

a) the first ball is replaced before drawing the second ball
b) the first ball is not replaced.

Solution. a) On the first draw, there are 4 balls of which 2 are red.

$$\begin{aligned} P(\text{red, on first draw}) &= \frac{2}{4} \\ &= \frac{1}{2} \end{aligned}$$

On the second draw, there are 4 balls of which 2 are red.

$$P(\text{red, on second draw}) = \frac{1}{2}$$

$$P(\text{2 reds in succession}) = \frac{1}{2} \times \frac{1}{2}$$
$$= \frac{1}{4}$$

Since the first ball was replaced before the second ball was drawn, the second event is independent of the first event.

b) $P(\text{red, on first draw}) = \frac{1}{2}$

On the second draw, there are 3 balls of which 1 is red.

$$P(\text{red, on second draw}) = \frac{1}{3}$$

$$P(\text{2 reds in succession}) = \frac{1}{2} \times \frac{1}{3}$$
$$= \frac{1}{6}$$

Since the first ball was not replaced before the second ball was drawn, the second event is dependent upon the first event.

EXERCISES 11-9

(A)

1. A coin and a regular tetrahedron (with faces marked 1, 2, 3, 4) are tossed. Draw a tree diagram to find the probability of getting:
   a) a head and a 1
   b) a tail and an even number.
2. It is equally likely that a child be born a girl or a boy. Draw a tree diagram to find the probability that:
   a) a family of two children will be both girls
   b) a family of 5 children will be all boys.
3. a) Draw a tree diagram to show the result of tossing a coin five times.
   b) What is the probability of tossing a coin five times and getting tails each time?
4. A True-False test has 6 questions. If all the questions are attempted by guessing, what is the probability of getting all 6 right?
5. What is the probability of rolling three consecutive sixes with one die?

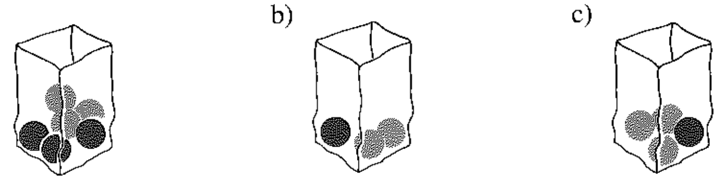
(B)

6. If a thumbtack is tossed, the probability that it lands with the point up is 0.6. What is the probability of tossing a thumbtack four times and having it land with the point up each time?
7. Two people are selected at random. What is the probability that they both have birthdays in September? (Assume a year has 365 days.)

8. A box of 100 flash cubes contains 3 defective ones. If 2 cubes are taken simultaneously from the box, what is the probability that both are defective?

9. Three bags contain black balls and red balls in the numbers shown.

   a) b) c)



   From each bag, find the probability of drawing 2 red balls in succession if:
   i) the first ball is replaced before the second ball is drawn
   ii) the first ball is not replaced.

10. Find the probability of drawing 3 red balls in succession from a bag containing 3 red balls and 3 black balls if:
    a) each ball is replaced before the next ball is drawn
    b) the balls are not replaced after drawing.

11. A card is drawn from each of two well-shuffled decks. Find the probability of drawing two cards that are both:
    a) spades b) red c) aces d) the ace of spades.

12. Find the probability of drawing 4 aces from a deck of cards:
    a) if there is replacement and shuffling after each draw
    b) if there is no replacement of the cards drawn.

13. A die and two coins are tossed. Find the probability of getting:
    a) a 2 and two heads b) a head, a tail, and an odd number.

Ⓒ

14. Suppose a die is tossed until a 6 appears. Find the probability that the throw on which it appears is:
    a) the second b) the third c) the tenth.

15. Two tetrahedrons with faces labelled 1, 2, 3, 4, are tossed. Calculate the probability that they show:
    a) two ones b) anything but two ones
    c) a sum of 5 d) a sum other than 5.

16. When two dice are tossed, what is the probability that they show:
    a) anything other than two sixes
    b) a sum of 7
    c) a sum of 11
    d) a sum not equal to 7 or 11
    e) at least one 3?

17. On a certain day, the probability of precipitation was 40% in Thunder Bay, 70% in London, and 20% in Ottawa. What was the probability of precipitation in all three cities on that day?

11-10 SIMULATING EXPERIMENTS

Adam, Bradley, Courtney, Dana, Erin, and Franco all volunteered to serve on the school organizing committee for the spring dance. The school charter says that the committee must have exactly 4 members. Mrs. Goreski, the staff supervisor, said she would choose a committee of 4 from the 6 volunteers by random selection. What is the probability that Adam is chosen?

In Section 11-8, we learned that to find the probability that Adam is chosen, we would have to count the number of committees of 4 people of which Adam is a member and divide that number by the total number of possible committees. This would be a long and tedious process.

We can find an approximate value of the probability by *simulating* the random selection as follows.

- First, label 6 identical cards with the first letter of each student's name.

A	B	C	D	E	F

- Then, shuffle the cards and choose any 4 of the 6 cards. These 4 cards name the committee members. For example, the selection of cards **B**, **C**, **E**, and **F** would indicate that the committee was to consist of Bradley, Courtney, Erin, and Franco.

- Do the second step 30 times. The results of one experiment are listed below.

C F E A	**C F E D**	A D B F	**B E F D**	B A E D
B F A C	**D E F B**	B E C A	A C E F	**E C B D**
F A B E	**E F B C**	**F C D B**	F D B A	**B D E C**
F C A E	A D F E	E D A C	A F D C	B A D E
B F D C	C B A D	A B F D	C B F A	A C B D
A B E F	D A C E	F A C B	E A D F	**C D E B**

- From the list of results, the number of times that **A** was chosen is 21.

$$\text{Then, P(Adam being chosen)} = \frac{\text{Number of times A occurred}}{\text{Total number of draws}}$$

$$= \frac{21}{30}$$

$$= 0.7$$

The probability that Adam is chosen is 0.7.

The more we repeat step 2 of this simulation, the more likely our estimate will be a close approximation of the true probability.

We use simulations to approximate probabilities when the calculation of a true probability is difficult or when there are insufficient data.

The following example shows how we can simulate an experiment by tossing several coins.

Example. Use a simulation to estimate the probability that there will be exactly 3 girls in a family of 5 children.

Solution. We use the fact that the probability that a randomly selected child is a girl is $\frac{1}{2}$. Therefore, we can simulate the selection of a girl with the outcome of "heads" on the toss of a coin.

Toss 5 coins 40 times and record the results. They are listed below and those outcomes in which exactly 3 heads appeared are highlighted.

H H T H H	T T H T H	H T H H H	T H T H T	T T H T H
H T T H T	H H H T T	T H H T H	H T T H T	T H T T H
T T H H T	H H T T T	H T T H H	H H H H T	T T T T H
H T T H H	T H H T T	H T T T T	H T T H H	T H H T T
T T T H T	H H T T H	H T H H H	T H T T T	H H H H H
T T H T T	H H T T H	T H H H H	H T T T H	T T H H T
T H T T H	H T H T T	T H T H H	T H H T T	H T H H H
H H H T T	T H T H H	T T H T T	T H H H H	H T H T H

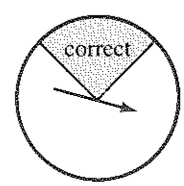
Exactly 3 heads occurred 11 times on 40 tosses so the estimated probability is $\frac{11}{40}$, or 0.275.

By increasing the number of tosses, we increase the likelihood of a more accurate estimate.

EXERCISES 11-10

Ⓑ

1. Explain how this spinner could be used to estimate the probability of guessing at least 3 answers on a 10-question multiple-choice test. Each question offers 4 answers.



2. Describe a simulation you could conduct to estimate each probability.
   a) The probability that in a family of 4 children there are exactly 3 boys.
   b) The probability that you will guess correctly more than 5 answers on an 8-question true-false test.
   c) The probability that three people born in April are all born on an even-numbered day.
   d) The probability that the last two digits in a randomly selected telephone number are both even.

3. Conduct a simulation to estimate the probability that a family of 4 children has 2 girls and 2 boys. Use at least 36 trials.

4. Conduct a simulation to estimate the probability of guessing at least 3 correct answers on an 8-question true-false test. Assume the probability of a correct guess is $\frac{1}{2}$.

5. Use dice to conduct a simulation which you can use to estimate the probability of guessing at least 2 correct answers on an 8-question multiple-choice test. Each question offers 6 answers.

6. Use a telephone book to estimate the probability that the last two digits of a telephone number are both even.

C

7. Use a coin and a die to estimate the probability that at least 2 people in any group of 5 are born in the same month. Assume all months are equally probable. Use at least 15 trials.

INVESTIGATE

Random Numbers

Numbers that are selected at random so that all numbers and all digits are equally likely are called *random numbers*.

Tables of random numbers have been printed, which can be used to estimate the probabilities of certain events.

Here is a list of 200 random numbers.

7813	7191	6347	5646	7021	3575	5608	3257	7225	7593
5149	6646	6674	7952	6267	3078	5721	3502	3224	5082
5166	3831	6934	6965	3025	7346	5883	5451	3482	6223
3256	5295	6413	5325	3557	6079	0148	5742	6781	6540
7707	5662	5186	6524	7383	3965	3718	3287	7075	6541

Use the list the answer these questions.

1. Find the probability that exactly 2 digits out of 4 randomly chosen digits are even. To do this, count how many 4-digit numbers have exactly 2 even digits. Express this number as a fraction of the total, 50.

2. Use the result of *Question 1* to estimate the probability that a family of 4 children has exactly 2 girls and 2 boys.

3. Find the probability that a family of 8 children has exactly 5 boys.

4. Find the probability of guessing at least 4 correct answers on a 10-question true-false test.

5. Find the probability that at least 2 of the 3 digits on a licence plate are the same.

COMPUTER POWER

Simulating Experiments

When a coin is tossed it is just as likely that it will land heads up as tails up. However, when a coin is tossed 1000 times, would we expect it to land heads up exactly 500 times? Is it reasonable to expect the coin to land heads up fewer than 400 times on 1000 tosses?

To obtain answers to such questions we could toss a coin 1000 times and record the results. We might then repeat this experiment many times to verify our results. Tossing a coin a large number of times is a long, tedious process as well as being impractical. It is more appropriate to imitate or *simulate* such experiments.

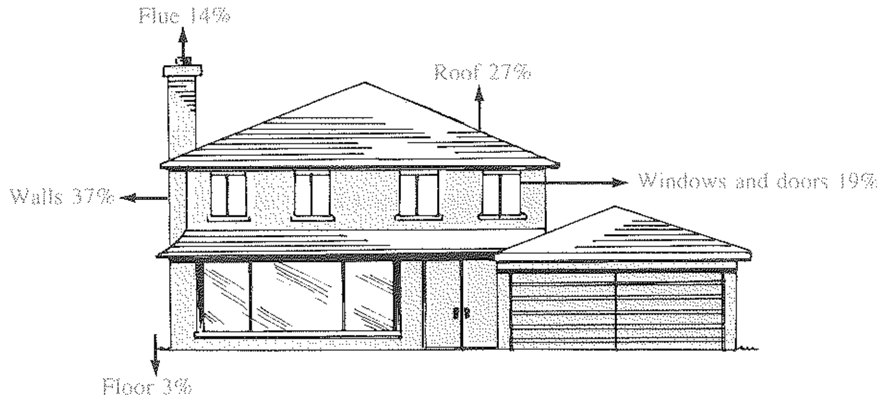
To simulate a coin toss, we could choose a digit from a table of random numbers. However, the computer has made the use of a random numbers table unnecessary. Microcomputers are equipped with a random number generator. This feature makes the microcomputer an extremely useful device for simulating probability experiments.

The program below selects the numbers 0 and 1 at random and identifies the selection of 0 with a head and the selection of 1 with a tail. The number of heads and tails are tallied and displayed when the program is run.

```
100 REM *** SIMULATING COIN TOSSES ***
110 INPUT "HOW MANY TOSSES DO YOU WISH? ";N
120 H=0
130 FOR K=1 TO N
140     X=INT(2*RND(1))
150     IF X=0 THEN H=H+1:PRINT K, "H"
160     IF X=1 THEN PRINT K, "T"
170 NEXT K
180 PRINT:PRINT "TOSSES","HEADS","TAILS"
190 PRINT N,H,N-H
200 END
```

1. Use the program to determine the number of heads in:
   a) 100 tosses of a coin   b) 500 tosses of a coin   c) 1000 tosses of a coin.
2. What percent of the tosses (simulated in *Question 1*) yielded heads out of:
   a) 100 tosses   b) 500 tosses   c) 1000 tosses?
3. What happens to the percent of tosses which yields heads, as the number of tosses becomes very large?
4. Is it likely that fewer than 400 heads will result on 1000 tosses of a coin? Explain your answer.

1. The diagram shows how heat is lost from a typical two-storey home. Draw a circle graph to show this information.



2. The workers in a small factory receive these salaries.

\$10 000	10 000	10 400	10 800	13 200
10 800	11 200	12 000	12 000	12 400
12 400	12 400	12 400	12 800	10 800
14 000	14 000	14 400	14 800	14 800
15 200	16 000	15 600	15 200	16 000

Display the data on a histogram.

3. The table lists the defence spending of some nations of the North Atlantic Treaty Organization (NATO) as a percent of their gross national product (G.N.P.). Show this information on a suitable graph.

Country	Defence Spending as a Percent of G.N.P.
Britain	4.7
Canada	1.8
Denmark	2.4
France	3.3
Norway	3.2
U.S.A.	5.0
West Germany	3.4

4. Calculate the measures of central tendency for the salaries given in *Exercise 2*.

5. If the mean of the numbers 9, 10, 21, 27, 29, 25, 19, 13, x is 21, what is x?

6. Write nine natural numbers that have a median of 25 and a mean of 21.

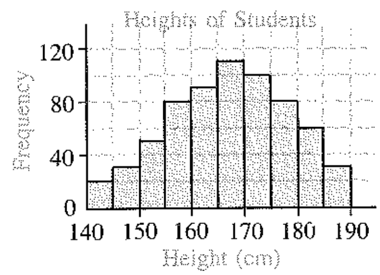
7. How would you collect data to determine the following information?
   a) The extent of mercury poisoning in fish in the Great Lakes
   b) The political party most likely to win the next provincial election
   c) The food-purchasing habits of single males
   d) The force required to break a certain gauge of fishing line

8. Shake 5 coins in a paper cup and empty them onto your desk. Record the frequency of heads. Repeat this procedure 24 times. From your results, if you did this a total of 300 times, with what frequency would you expect 5 coins to show:
   a) 3 heads b) 4 heads c) no heads?

9. A manufacturer of widgets has maintained a minimum standard of 95% dependability over the years.
   a) Three widgets in a batch of 75 are found to be defective. Does the batch meet the minimum standard?
   b) How many defective widgets are permissible in a batch of 250?
   c) Workmanship and materials are improved so that only 4 defective widgets are being found in every 250. What is the probability that a widget selected at random is not defective?

10. A ball is selected at random from 15 balls numbered from 1 to 15. What is the probability that the number is:
   a) even b) prime c) a multiple of 5 d) a 2-digit number?

11. The bar graph shows the distribution of the heights of students at Montcalm Secondary School. What is the probability that a student selected at random will be:
   a) between 150 cm and 165 cm tall
   b) taller than 175 cm
   c) shorter than 155 cm?



12. A box of coins contains 36 quarters, 45 dimes, 25 nickels, and 62 pennies. What is the probability that a coin drawn at random will be:
   a) a quarter b) a nickel or a dime
   c) a quarter or a nickel d) other than a penny?

13. The faces of two regular tetrahedrons are numbered 1 to 4. If they are tossed, what is the probability of getting:
   a) two numbers the same b) a total of 5 c) a difference of 1?

14. An aviary has parakeets of four different colors. There are 10 green, 7 blue, 2 yellow, and 1 white. If two birds escape, what is the probability that they will both be:
   a) green b) blue c) yellow d) white?

15. A cafeteria offers a number of choices for lunch.
   3 appetizers: soup, juice, or salad
   4 main courses: beef, chicken, pork, or fish
   2 desserts: pie or ice cream
   If a three-course meal is selected at random, calculate the probability of getting:
   a) soup b) soup and beef c) juice, fish, and pie.

16. Conduct a simulation to estimate the probability that a family of 4 children has at least 3 boys. Use at least 36 trials.