CHAPTER

1 Linear Functions

Over short time intervals, many changes in the economy are well modeled by linear functions. In an exercise in the first section of this chapter we will examine a linear model that predicts airline passenger traffic in the year 2005 at some of the fastest-growing airports in the United States. Such predictions are important tools for airline executives and airport planners.

ISBN: 0-536-10718-1

Before using mathematics to solve a real-world problem, we must usually set up a **mathematical model,** a mathematical description of the situation. Constructing such a model requires a solid understanding of the situation to be modeled, as well as familiarity with relevant mathematical ideas and techniques.

Much mathematical theory is available for building models, but the very richness and diversity of contemporary mathematics often prevents people in other fields from finding the mathematical tools they need. There are so many useful parts of mathematics that it can be hard to know which to choose.

To avoid this problem, it is helpful to have a thorough understanding of the most basic and useful mathematical tools that are available for constructing mathematical models. In this chapter we look at some mathematics of *linear* models, which are used for data whose graphs can be approximated by straight lines.

1.1 SLOPES AND EQUATIONS OF LINES

THINK ABOUT IT *How fast has tuition at public colleges been increasing in recent years, and how well can we predict tuition in the future?*

In Example 15 of this section, we will answer these questions using the equation of a line.

There are many everyday situations in which two quantities are related. For example, if a bank account pays 6% simple interest per year, then the interest I that a deposit of P dollars would earn in one year is given by

$$I = .06 \cdot P, \qquad \text{or} \qquad I = .06P.$$

The formula $I = .06P$ describes the relationship between interest and the amount of money deposited.

Using this formula, we see, for example, that if $P = \$100$, then $I = \$6$, and if $I = \$12$, then $P = \$200$. These corresponding pairs of numbers can be written as **ordered pairs,** $(100, 6)$ and $(200, 12)$, pairs of numbers whose order is important. The first number denotes the value of P and the second number the value of I.

Ordered pairs are **graphed** with the perpendicular number lines of a **Cartesian coordinate system,** shown in Figure 1. The horizontal number line, or ***x*-axis,** represents the first components of the ordered pairs, while the vertical or ***y*-axis** represents the second components. The point where the number lines cross is the zero point on both lines; this point is called the **origin.**

The name "Cartesian" honors René Descartes (1596–1650), one of the greatest mathematicians of the seventeenth century. According to legend, Descartes was lying in bed when he noticed an insect crawling on the ceiling and realized that if he could determine the distance from the bug to each of two perpendicular walls, he could describe its position at any given moment. The same idea can be used to locate a point in a plane.

Each point on the xy-plane corresponds to an ordered pair of numbers, where the x-value is written first. From now on, we will refer to the point corresponding to the ordered pair (a, b) as "the point (a, b)."

Locate the point $(-2, 4)$ on the coordinate system by starting at the origin and counting 2 units to the left on the horizontal axis and 4 units upward, parallel to the vertical axis. This point is shown in Figure 1, along with several other sample points. The number -2 is the ***x*-coordinate** and the number 4 is the ***y*-coordinate** of the point $(-2, 4)$.

ISBN: 0-536-10718-1

The x-axis and y-axis divide the plane into four parts, or **quadrants.** For example, quadrant I includes all those points whose x- and y-coordinates are both positive. The quadrants are numbered as shown in Figure 1. The points on the axes themselves belong to no quadrant. The set of points corresponding to the ordered pairs of an equation is the **graph** of the equation.

The x- and y-values of the points where the graph of an equation crosses the axes are called the ***x*-intercept** and ***y*-intercept,** respectively.* See Figure 2.

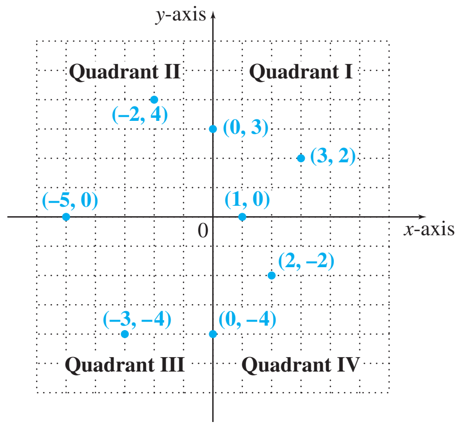


FIGURE 1

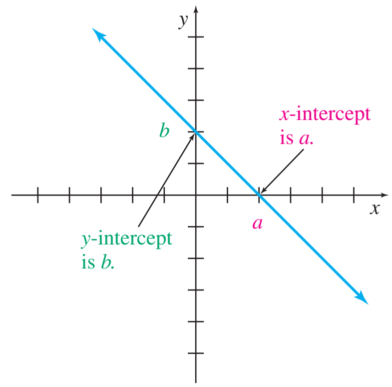


FIGURE 2

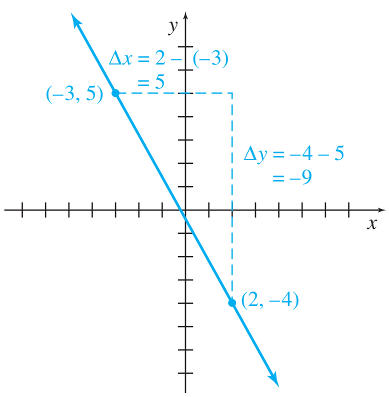


FIGURE 3

Slope of a Line

An important characteristic of a straight line is its *slope,* a number that represents the "steepness" of the line. To see how slope is defined, look at the line in Figure 3. The line goes through the points $(x_1, y_1) = (-3, 5)$ and $(x_2, y_2) = (2, -4)$. The difference in the two x-values,

$$x_2 - x_1 = 2 - (-3) = 5$$

in this example, is called the **change in *x*.** The symbol Δx (read "delta x") is used to represent the change in x. In the same way, Δy represents the **change in *y*.** In our example,

$$\begin{aligned}\Delta y &= y_2 - y_1\\ &= -4 - 5\\ &= -9.\end{aligned}$$

These symbols, Δx and Δy, are used in the following definition of slope.

SLOPE OF A LINE

The **slope** of a line is defined as the vertical change (the "rise") over the horizontal change (the "run") as one travels along the line. In symbols, taking two different points (x_1, y_1) and (x_2, y_2) on the line, the slope is

$$m = \frac{\textbf{Change in } y}{\textbf{Change in } x} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1},$$

where $x_1 \neq x_2$.

*Some people prefer to define the intercepts as ordered pairs, rather than as numbers.

ISBN: 0-536-10718-1

By this definition, the slope of the line in Figure 3 is

$$m = \frac{\Delta y}{\Delta x} = \frac{-4 - 5}{2 - (-3)} = -\frac{9}{5}.$$

The slope of a line tells how fast y changes for each unit of change in x.

NOTE Using similar triangles, it can be shown that the slope of a line is independent of the choice of points on the line. That is, the same slope will be obtained for *any* choice of two different points on the line. ■

EXAMPLE 1 *Slope*

Find the slope of the line through each pair of points.

(a) $(-7, 6)$ and $(4, 5)$

Solution Let $(x_1, y_1) = (-7, 6)$ and $(x_2, y_2) = (4, 5)$. Use the definition of slope.

$$m = \frac{\Delta y}{\Delta x} = \frac{5 - 6}{4 - (-7)} = -\frac{1}{11}$$

(b) $(5, -3)$ and $(-2, -3)$

Solution Let $(x_1, y_1) = (5, -3)$ and $(x_2, y_2) = (-2, -3)$. Then

$$m = \frac{-3 - (-3)}{-2 - 5} = \frac{0}{-7} = 0.$$

Lines with zero slope are horizontal (parallel to the x-axis).

(c) $(2, -4)$ and $(2, 3)$

Solution Let $(x_1, y_1) = (2, -4)$ and $(x_2, y_2) = (2, 3)$. Then

$$m = \frac{3 - (-4)}{2 - 2} = \frac{7}{0},$$

which is undefined. This happens when the line is vertical (parallel to the y-axis).

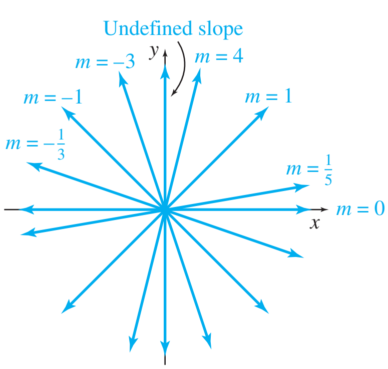


FIGURE 4

CAUTION The phrase "no slope" should be avoided; specify instead whether the slope is zero or undefined. ■

In finding the slope of the line in Example 1(a), we could have let $(x_1, y_1) = (4, 5)$ and $(x_2, y_2) = (-7, 6)$. In that case,

$$m = \frac{6 - 5}{-7 - 4} = \frac{1}{-11} = -\frac{1}{11},$$

the same answer as before. The order in which coordinates are subtracted does not matter, as long as it is done consistently.

Figure 4 shows examples of lines with different slopes. Lines with positive slopes go up from left to right, while lines with negative slopes go down from left to right.

ISBN: 0-536-10718-1

It might help you to compare slope with the percent grade of a hill. If a sign says a hill has a 10% grade uphill, this means the slope is .10, or $1/10$, so the hill rises 1 foot for every 10 feet horizontally. A 15% grade downhill means the slope is $-.15$.

FOR REVIEW

For review on solving a linear equation, see Section R.4.

Equations of a Line An equation in two first-degree variables, such as $4x + 7y = 20$, has a line as its graph, so it is called a **linear equation.** In the rest of this section, we consider various forms of the equation of a line.

EXAMPLE 2 *Equation of a Line*

Find the equation of the line through $(0, -3)$ with slope $3/4$.

Solution We can use the definition of slope, letting $(x_1, y_1) = (0, -3)$ and (x, y) represent another point on the line.

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

$$\frac{3}{4} = \frac{y - (-3)}{x - 0} = \frac{y + 3}{x} \quad \text{Substitute.}$$

$$3x = 4(y + 3) \quad \text{Cross multiply.}$$

$$3x = 4y + 12$$

A generalization of the method of Example 2 can be used to find the equation of any line, given its y-intercept and slope. Assume that a line has y-intercept b, so that it goes through the point $(0, b)$. Let the slope of the line be represented by m. If (x, y) is any point on the line *other* than $(0, b)$, then the definition of slope can be used with the points $(0, b)$ and (x, y) to get

$$m = \frac{y - b}{x - 0}$$

$$m = \frac{y - b}{x}$$

$$mx = y - b$$

$$y = mx + b.$$

This result is called the *slope-intercept form* of the equation of a line, because b is the y-intercept of the graph of the line.

SLOPE-INTERCEPT FORM

If a line has slope m and y-intercept b, then the equation of the line in **slope-intercept form** is

$$\boldsymbol{y = mx + b.}$$

When $b = 0$, we say that y is **proportional** to x.

EXAMPLE 3 *Slope-Intercept Form*

Find the equation of the line in slope-intercept form having y-intercept $7/2$ and slope $-5/2$.

ISBN: 0-536-10718-1

Solution Use the slope-intercept form with $b = 7/2$ and $m = -5/2$.

$$y = mx + b$$
$$y = -\frac{5}{2}x + \frac{7}{2}$$

The slope-intercept form shows that we can find the slope of a line by solving its equation for y. In that form the coefficient of x is the slope and the constant term is the y-intercept. For instance, in Example 2 the slope of the line $3x = 4y + 12$ was given as $3/4$. This slope also could be found by solving the equation for y.

$$4y + 12 = 3x$$
$$4y = 3x - 12$$
$$y = \frac{3}{4}x - 3$$

The coefficient of x, $3/4$, is the slope of the line. The y-intercept is -3.

The slope-intercept form of the equation of a line involves the slope and the y-intercept. Sometimes, however, the slope of a line is known, together with one point (perhaps *not* the y-intercept) that the line goes through. The *point-slope form* of the equation of a line is used to find the equation in this case. Let (x_1, y_1) be any fixed point on the line and let (x, y) represent any other point on the line. If m is the slope of the line, then by the definition of slope,

$$\frac{y - y_1}{x - x_1} = m,$$

or

$$y - y_1 = m(x - x_1). \quad \text{Multiply both sides by } x - x_1.$$

POINT-SLOPE FORM

If a line has slope m and passes through the point (x_1, y_1), then an equation of the line is given by

$$\boldsymbol{y - y_1 = m(x - x_1),}$$

the **point-slope form** of the equation of a line.

EXAMPLE 4 *Point-Slope Form*

Find an equation of the line that passes through the point $(3, -7)$ and has slope $m = 5/4$.

Solution Use the point-slope form.

$$y - y_1 = m(x - x_1)$$
$$y - (-7) = \frac{5}{4}(x - 3) \quad y_1 = -7,\ m = \tfrac{5}{4},\ x_1 = 3$$
$$y + 7 = \frac{5}{4}(x - 3)$$
$$4y + 28 = 5(x - 3) \quad \text{Multiply both sides by 4.}$$
$$4y + 28 = 5x - 15$$
$$4y = 5x - 43 \quad \text{Combine constants.}$$

FOR REVIEW

See Section R.4 for details on eliminating denominators in an equation.

ISBN: 0-536-10718-1

The equation of the same line can be given in many forms. To avoid confusion, the linear equations used in the rest of this section will be written in slope-intercept form, $y = mx + b$, which is often the most useful form.

The point-slope form also can be useful to find an equation of a line if we know two different points that the line goes through. The procedure for doing this is shown in the next example.

EXAMPLE 5 *Using Point-Slope Form to Find an Equation*

Find an equation of the line through $(5, 4)$ and $(-10, -2)$.

Solution Begin by using the definition of slope to find the slope of the line that passes through the given points.

$$\text{Slope} = m = \frac{-2 - 4}{-10 - 5} = \frac{-6}{-15} = \frac{2}{5}$$

Either $(5, 4)$ or $(-10, -2)$ can be used in the point-slope form with $m = 2/5$. If $(x_1, y_1) = (5, 4)$, then

$$\begin{aligned}
y - y_1 &= m(x - x_1) \\
y - 4 &= \frac{2}{5}(x - 5) && y_1 = 4, m = \tfrac{2}{5}, x_1 = 5. \\
5y - 20 &= 2(x - 5) && \text{Multiply both sides by 5.} \\
5y - 20 &= 2x - 10 && \text{Distributive property} \\
5y &= 2x + 10 && \text{Add 20 to both sides.} \\
y &= \frac{2}{5}x + 2 && \text{Divide by 5 to put in slope-intercept form.}
\end{aligned}$$

Check that the same result is found if $(x_1, y_1) = (-10, -2)$.

EXAMPLE 6 *Horizontal Line*

Find an equation of the line through $(8, -4)$ and $(-2, -4)$.

Solution Find the slope.

$$m = \frac{-4 - (-4)}{-2 - 8} = \frac{0}{-10} = 0$$

Choose, say, $(8, -4)$ as (x_1, y_1).

$$\begin{aligned}
y - y_1 &= m(x - x_1) \\
y - (-4) &= 0(x - 8) && y_1 = -4, m = 0, x_1 = 8 \\
y + 4 &= 0 && 0(x - 8) = 0 \\
y &= -4
\end{aligned}$$

Plotting the given ordered pairs and drawing a line through the points, show that the equation $y = -4$ represents a horizontal line. See Figure 5a on the next page. Every horizontal line has a slope of zero and an equation of the form $y = k$, where k is the y-value of all ordered pairs on the line.

ISBN: 0-536-10718-1

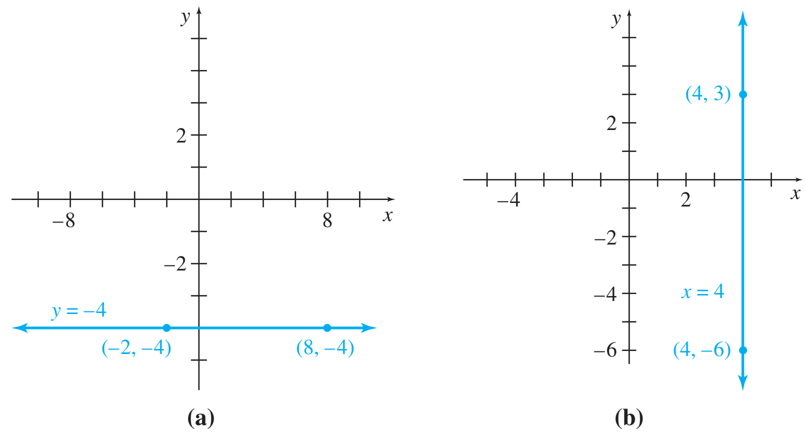


FIGURE 5

EXAMPLE 7 *Vertical Line*

Find an equation of the line through $(4, 3)$ and $(4, -6)$.

Solution The slope of the line is

$$m = \frac{-6 - 3}{4 - 4} = \frac{-9}{0},$$

which is undefined. Since both ordered pairs have x-coordinate 4, the equation is $x = 4$. Because the slope is undefined, the equation of this line cannot be written in the slope-intercept form.

Again, plotting the given ordered pairs and drawing a line through them show that the graph of $x = 4$ is a vertical line. See Figure 5(b).

The slope of a horizontal line is 0.
The slope of a vertical line is undefined.

The different forms of linear equations discussed in this section are summarized below. The slope-intercept and point-slope forms are equivalent ways to express the equation of a nonvertical line. The slope-intercept form is simpler for a final answer, but you may find the point-slope form easier to use when you know the slope of a line and a point through which the line passes.

EQUATIONS OF LINES

Equation	Description
$y = mx + b$	**Slope-intercept form:** slope m, y-intercept b
$y - y_1 = m(x - x_1)$	**Point-slope form:** slope m, line passes through (x_1, y_1)
$x = k$	**Vertical line:** x-intercept k, no y-intercept (except when $k = 0$), undefined slope
$y = k$	**Horizontal line:** y-intercept k, no x-intercept (except when $k = 0$), slope 0

ISBN: 0-536-10718-1

Parallel and Perpendicular Lines One application of slope involves deciding whether two lines are parallel. Since two parallel lines are equally "steep," they should have the same slope. Also, two lines with the same "steepness" are parallel.

PARALLEL LINES

Two lines are **parallel** if and only if they have the same slope, or if they are both vertical.

EXAMPLE 8 *Parallel Line*

Find the equation of the line that passes through the point $(3, 5)$ and is parallel to the line $2x + 5y = 4$.

Solution The slope of $2x + 5y = 4$ can be found by writing the equation in slope-intercept form.

$$2x + 5y = 4$$
$$y = -\frac{2}{5}x + \frac{4}{5}$$

This result shows that the slope is $-2/5$. Since the lines are parallel, $-2/5$ is also the slope of the line whose equation we want. This line passes through $(3, 5)$. Substituting $m = -2/5$, $x_1 = 3$, and $y_1 = 5$ into the point-slope form gives

$$y - y_1 = m(x - x_1)$$
$$y - 5 = -\frac{2}{5}x + \frac{6}{5}$$
$$y = -\frac{2}{5}x + \frac{6}{5} + 5$$
$$y = -\frac{2}{5}x + \frac{31}{5}$$

As already mentioned, two nonvertical lines are parallel if and only if they have the same slope. Two lines having slopes with a product of -1 are perpendicular. A proof of this fact, which depends on similar triangles from geometry, is given as Exercise 43 in this section.

PERPENDICULAR LINES

Two lines are **perpendicular** if and only if the product of their slopes is -1, or if one is vertical and the other horizontal.

EXAMPLE 9 *Perpendicular Line*

Find the slope of the line L perpendicular to the line having the equation $5x - y = 4$.

Solution To find the slope, write $5x - y = 4$ in slope-intercept form:

$$y = 5x - 4.$$

ISBN: 0-536-10718-1

The slope is 5. Since the lines are perpendicular, if line L has slope m, then

$$5m = -1$$
$$m = -\frac{1}{5}.$$

The next two examples use different forms of the equation of a line to analyze real-world data.

EXAMPLE 10 *Workforce*

In recent decades, the percentage of the U.S. civilian population age 16 and over that is in the workforce has risen at a roughly constant rate, from 59.4% in 1960 to 66.6% in 2002.* Find the equation describing this linear relationship.

Solution For this example, let x represent time in years, with $x = 0$ for 1960. Such rescaling of a variable is often used to simplify the arithmetic, although computers and calculators have made rescaling less important than in the past. Here it allows us to work with smaller numbers, and, as you will see, find the y-intercept of the line more easily. We will use such rescaling on many examples throughout this book. When we do, it is important to be consistent. In this example, if we want to refer to the year 1975, we must let $x = 15$, and not $x = 1975$. Let y represent the percent of the population in the workforce.

With 1960 corresponding to $x = 0$, the year 2002 corresponds to $x = 2002 - 1960 = 42$. The two ordered pairs representing the given information are $(0, 59.4)$ and $(42, 66.6)$. The slope of the line through these points is

$$m = \frac{66.6 - 59.4}{42 - 0} = \frac{7.2}{42} = .1714286 \approx .171.^{\dagger}$$

This means that, on average, the percent of the population in the workforce has gone up by about .17% per year.

Using $m = .171$ and $(x_1, y_1) = (0, 59.4)$ in the point-slope form gives the required equation,

$$y - 59.4 = .171(x - 0)$$
$$y = .171x + 59.4.$$

This result could also have been obtained by observing that $(0, 59.4)$ is the y-intercept.

Notice that if this formula is valid for all nonnegative x, then eventually y becomes 100:

$$.171x + 59.4 = 100$$
$$.171x = 40.6 \qquad \text{Subtract 59.4 from both sides.}$$
$$x = 40.6/.171 \approx 237, \qquad \text{Divide both sides by .171.}$$

which indicates that 237 years from 1960 (in the year 2197), 100% of the population will be in the workforce.

*U.S. Department of Labor, *Bureau of Labor Statistics Data*.
†The symbol ≈ means "is approximately equal to."

ISBN: 0-536-10718-1

In Example 10, of course, it is still possible that in 2197 there will be people who are not in the workforce; the trend of recent decades may not continue. Most equations are valid for some specific set of numbers. It is highly speculative to extrapolate beyond those values. On the other hand, people in business and government often need to make some prediction about what will happen in the future, so a tentative conclusion based on past trends may be better than no conclusion at all. There are also circumstances, particularly in the physical sciences, in which theoretical reasons imply that the trend will continue.

EXAMPLE 11 *Antibiotic Resistance*

The linear equation $y = 4.9x - 9783.9$, where x represents the year, can be used to estimate the percent of the gonorrhea cases diagnosed in Hawaii from 1997 to 2001 that are resistant to the commonly prescribed antibiotic ciprofloxacin.*

(a) Determine this percent in 2003.

Solution Substitute 2003 for x in the equation.

$$\begin{aligned} y &= 4.9x - 9783.9 \\ &= 4.9(2003) - 9783.9 \\ &\approx 30.8 \end{aligned}$$

This means that about 30.8% of gonorrhea cases diagnosed in Hawaii in 2003 were resistant to the antibiotic ciprofloxacin.

(b) Find and interpret the slope of the line.

Solution The equation is given in slope-intercept form, so the slope is the coefficient of x, which is 4.9. Since

$$m = 4.9 = \frac{\text{change in } y}{\text{change in } x} = \frac{4.9}{1},$$

the slope indicates the change in the percent of ciprofloxacin-resistant cases in Hawaii per year from 1997 to 2001. Because the slope is positive, the percent of ciprofloxacin-resistant cases in Hawaii increased by 4.9% per year.

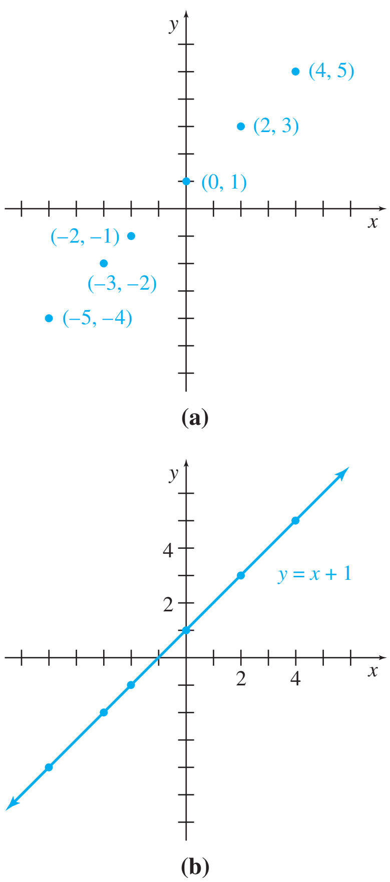


FIGURE 6

Graph of a Line

We can graph the linear equation defined by $y = x + 1$ by finding several ordered pairs. For example, if $x = 2$, then $y = 2 + 1 = 3$, giving the ordered pair $(2, 3)$. Also, $(0, 1)$, $(4, 5)$, $(-2, -1)$, $(-5, -4)$, $(-3, -2)$, among many others, are ordered pairs that satisfy the equation.

To graph $y = x + 1$, we begin by locating the ordered pairs obtained above, as shown in Figure 6(a). All the points of this graph appear to lie on a straight line, as in Figure 6(b). This straight line is the graph of $y = x + 1$.

It can be shown that every equation of the form $ax + by = c$ has a straight line as its graph. Although just two points are needed to determine a line, it is a good idea to plot a third point as a check. It is often convenient to use the x- and y-intercepts as the two points, as in the following example.

*U.S. Department of Defense, Gonococcal Isolate Surveillance Project, April 2002.

ISBN: 0-536-10718-1

EXAMPLE 12 *Graph of a Line*

Graph $3x + 2y = 12$.

Solution To find the y-intercept, let $x = 0$.

$$3(0) + 2y = 12$$
$$2y = 12 \quad \text{Divide both sides by 2.}$$
$$y = 6$$

Similarly, find the x-intercept by letting $y = 0$, which gives $x = 4$. Verify that when $x = 2$, the result is $y = 3$. These three points are plotted in Figure 7(a). A line is drawn through them in Figure 7(b).

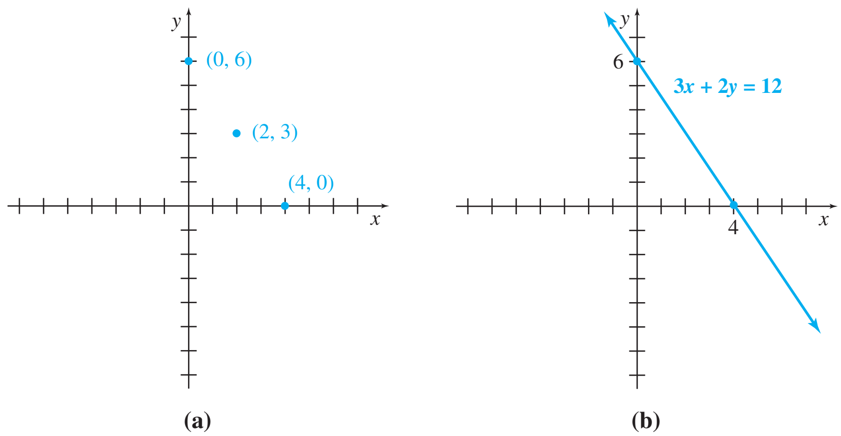


FIGURE 7

Not every line has two distinct intercepts; the graph in the next example does not cross the x-axis, and so it has no x-intercept.

EXAMPLE 13 *Graph of a Horizontal Line*

Graph $y = -3$.

Solution The equation $y = -3$, or equivalently, $y = 0x - 3$, always gives the same y-value, -3, for any value of x. Therefore, no value of x will make $y = 0$, so the graph has no x-intercept. As we saw in Example 6, the graph of such an equation is a horizontal line parallel to the x-axis. In this case the y-intercept is -3, as shown in Figure 8.

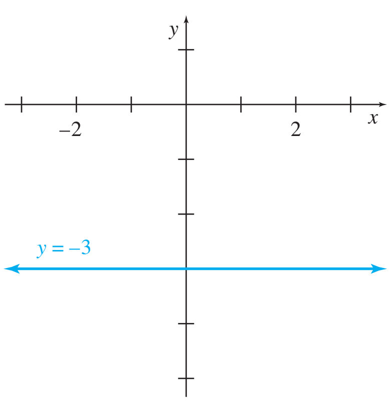


FIGURE 8

In general, the graph of $y = k$, where k is a real number, is the horizontal line having y-intercept k.

The graph in Example 13 had only one intercept. Another type of linear equation with coinciding intercepts is graphed in Example 14.

ISBN: 0-536-10718-1

EXAMPLE 14 Graph of a Line Through the Origin

Graph $y = -3x$.

Solution Begin by looking for the x-intercept. If $y = 0$, then

$$\begin{aligned} y &= -3x \\ 0 &= -3x \qquad \text{Let } y = 0. \\ 0 &= x. \qquad \text{Divide both sides by } -3. \end{aligned}$$

We have the ordered pair $(0, 0)$. Starting with $x = 0$ gives exactly the same ordered pair, $(0, 0)$. Two points are needed to determine a straight line, and the intercepts have led to only one point. To get a second point, we choose some other value of x (or y). For example, if $x = 2$, then

$$y = -3x = -3(2) = -6, \qquad \text{Let } x = 2.$$

giving the ordered pair $(2, -6)$. These two ordered pairs, $(0, 0)$ and $(2, -6)$, were used to get the graph shown in Figure 9.

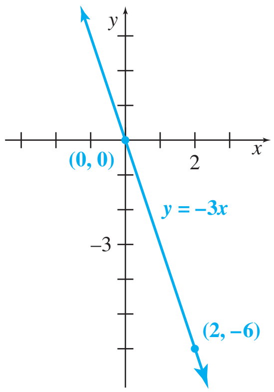


FIGURE 9

Linear equations allow us to set up simple mathematical models for real-life situations. In almost every case, linear (or any other reasonably simple) equations provide only approximations to real-world situations. Nevertheless, these are often remarkably useful approximations.

EXAMPLE 15 Tuition

The table lists the average annual cost (in dollars) of tuition and fees at public four-year colleges for selected years.*

Year	Tuition and Fees
1990	1908
1992	2334
1994	2705
1996	2975
1998	3243
2000	3510
2002	4081

(a) Plot the cost of public colleges by letting $x = 90$ correspond to 1990. Is the data *exactly* linear? Could the data be *approximated* by a linear equation?

Solution Use a scale from 1500 to 4500 on the y-axis. The graph is shown in Figure 10(a) in a figure known as a **scatterplot.** Although it is not exactly linear, it is approximately linear and could be approximated by a linear equation.

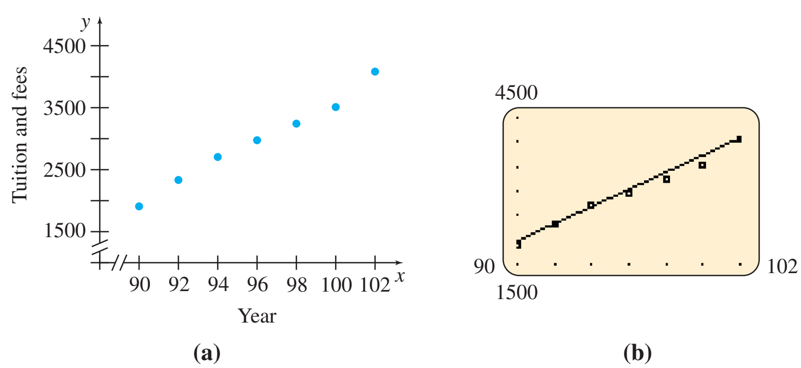


FIGURE 10

*The College Board.

ISBN: 0-536-10718-1

(b) Use the points $(92, 2334)$ and $(102, 4081)$ to determine an equation that models the data.

Solution After finding the slope, we substitute one of the given ordered pairs into a form of the equation of a line. Using $y = mx + b$, we proceed as follows.

The slope is

$$m = \frac{4081 - 2334}{102 - 92} = \frac{1747}{10} = 174.7.$$

From $y = mx + b$, with $x = 92$ and $y = 2334$,

$$\begin{aligned} 2334 &= 174.7(92) + b = 16{,}072.4 + b \\ b &= -13{,}738.4, \end{aligned}$$

so the equation is $y = 174.7x - 13{,}738.4$. (We could have used the other ordered pair for the values of x and y. Verify that we get the same equation.)

(c) Discuss the accuracy of using this equation to estimate the cost of public colleges in the year 2020.

Solution The year 2020 corresponds to $x = 120$, for which the equation predicts a cost of $174.7(120) - 13{,}738.4 = 7225.6$, or about \$7226. The year 2020 is many years in the future, however. Many factors could affect the tuition, and the actual figure for 2020 could be very different from our prediction.

You can plot data with a TI-83/84 Plus graphing calculator using the following steps.

1. Store the data in lists.
2. Define the stat plot.
3. Turn off $Y =$ functions (unless you also want to graph a function).
4. Turn on the plot you want to display.
5. Define the viewing window.
6. Display the graph.

Consult the calculator's instruction booklet or *The Graphing Calculator Manual* that is available with this book for specific instructions. See the calculator-generated graph in Figure 10(b), which includes the points and line from Example 15. Notice how the line closely approximates the data.

1.1 EXERCISES

Find the slope of each line that has a slope.

1. Through $(4, 5)$ and $(-1, 2)$

2. Through $(5, -4)$ and $(1, 3)$

ISBN: 0-536-10718-1

3. Through $(8, 4)$ and $(8, -7)$

4. Through $(1, 5)$ and $(-2, 5)$

5. $y = 2x$

6. $y = 3x - 2$

7. $5x - 9y = 11$

8. $4x + 7y = 1$

9. $x = -6$

10. The x-axis

11. $y = 8$

12. $y = -4$

13. A line parallel to $2y - 4x = 7$

14. A line perpendicular to $6x = y - 3$

Find an equation in slope-intercept form (where possible) for each line in Exercises 15–34.

15. Through $(1, 3)$, $m = -2$

16. Through $(2, 4)$, $m = -1$

17. Through $(6, 1)$, $m = 0$

18. Through $(-8, 1)$, with undefined slope

19. Through $(4, 2)$ and $(1, 3)$

20. Through $(8, -1)$ and $(4, 3)$

21. Through $(1/2, 5/3)$ and $(3, 1/6)$

22. Through $(-2, 3/4)$ and $(2/3, 5/2)$

23. Through $(-8, 4)$ and $(-8, 6)$

24. Through $(-1, 3)$ and $(0, 3)$

25. x-intercept 3, y-intercept -2

26. x-intercept -2, y-intercept 4

27. Vertical, through $(-6, 5)$

28. Horizontal, through $(8, 7)$

29. Through $(-1, 4)$, parallel to $x + 3y = 5$

30. Through $(2, -5)$, parallel to $y - 4 = 2x$

31. Through $(3, -4)$, perpendicular to $x + y = 4$

32. Through $(-2, 6)$, perpendicular to $2x - 3y = 5$

33. The line with y-intercept 2 and perpendicular to $3x + 2y = 6$

34. The line with x-intercept $-2/3$ and perpendicular to $2x - y = 4$

35. Do the points $(4, 3)$, $(2, 0)$, and $(-18, -12)$ lie on the same line? (*Hint:* Find the slopes between the points.)

36. Find k so that the line through $(4, -1)$ and $(k, 2)$ is

a. parallel to $2x + 3y = 6$,

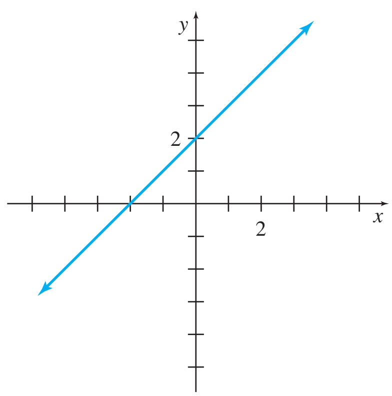
b. perpendicular to $5x - 2y = -1$.

37. Use slopes to show that the quadrilateral with vertices at $(1, 3)$, $(-5/2, 2)$, $(-7/2, 4)$, and $(2, 1)$ is a parallelogram.

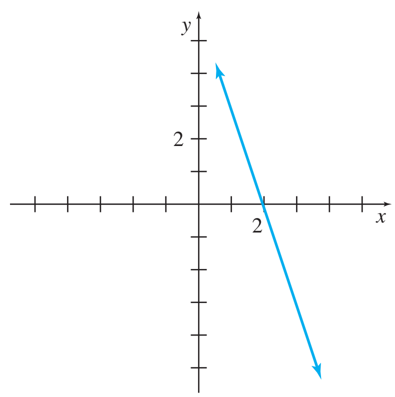
38. Use slopes to show that the square with vertices at $(-2, 5)$, $(4, 5)$, $(4, -1)$, and $(-2, -1)$ has diagonals that are perpendicular.

For the lines in Exercises 39 and 40, which of the following is closest to the slope of the line? (a) 1 (b) 2 (c) 3 (d) 21 (e) 22 (f) -3

39.



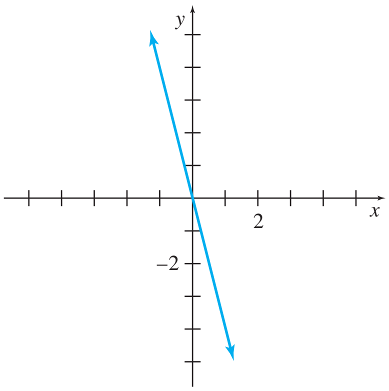
40.



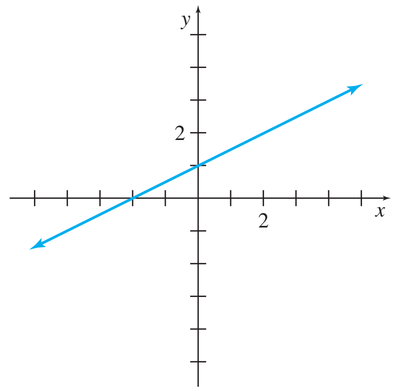
ISBN: 0-536-10718-1

Estimate the slope of the lines in Exercises 41 and 42.

41.



42.



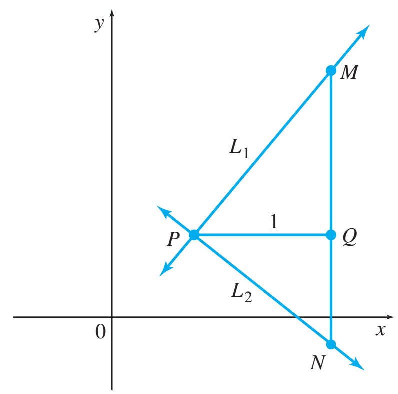
43. To show that two perpendicular lines, neither of which is vertical, have slopes with a product of -1, go through the following steps. Let line L_1 have equation $y = m_1x + b_1$, and let L_2 have equation $y = m_2x + b_2$. Assume that L_1 and L_2 are perpendicular, and use right triangle MPN shown in the figure. Prove each of the following statements.

a. MQ has length m_1.

b. QN has length $-m_2$.

c. Triangles MPQ and PNQ are similar.

d. $m_1/1 = 1/(-m_2)$ and $m_1m_2 = -1$



Graph each equation.

44. $y = x - 1$ **45.** $y = 2x + 3$ **46.** $y = -4x + 9$ **47.** $y = -6x + 12$

48. $2x - 3y = 12$ **49.** $3x - y = -9$ **50.** $3y + 4x = 12$ **51.** $4y + 5x = 10$

52. $y = -2$ **53.** $x = 4$ **54.** $x + 5 = 0$ **55.** $y - 4 = 0$

56. $y = 2x$ **57.** $y = -5x$ **58.** $x + 4y = 0$ **59.** $x - 3y = 0$

Applications

BUSINESS AND ECONOMICS

60. ***Sales*** The sales of a small company were \$27,000 in its second year of operation and \$63,000 in its fifth year. Let y represent sales in the xth year of operation. Assume that the data can be approximated by a straight line.

a. Find the slope of the sales line, and give an equation for the line in the form $y = mx + b$.

b. Use your answer from part a to find out how many years must pass before the sales surpass \$100,000.

61. ***Federal Debt*** The table lists the total federal debt (in trillions of dollars) from 1991 to 2001.*

Year	Federal Debt
1991	3.599
1992	4.002
1993	4.351
1994	4.644
1995	4.921
1996	5.182
1997	5.370
1998	5.479
1999	5.606
2000	5.629
2001	5.770

*Economic Research, Federal Reserve Bank of St. Louis, Fiscal Year 2002.

ISBN: 0-536-10718-1

a. Plot the data by letting $x = 0$ correspond to 1990. Discuss any trends of the federal debt over this time period.

b. Find a linear equation that approximates the data, using the points $(1, 3.599)$ and $(11, 5.770)$. What does the slope of the graph represent? Graph the line and the data on the same coordinate axes.

c. Use the equation from part b to predict the federal debt in the year 2002. Compare your result to the actual value of 6.199 trillion dollars.

62. ***Airline Passenger Growth*** The following table estimates the growth in the number of airline passengers (in millions) at some of the fastest-growing airports in the United States between 1992 and 2005.*

Airport	1992	2005
Harrisburg Intl.	.7	1.4
Dayton Intl.	1.1	2.4
Austin Robert Mueller	2.2	4.7
Milwaukee Gen. Mitchell Intl.	2.2	4.4
Sacramento Metropolitan	2.6	5.0
Fort Lauderdale–Hollywood	4.1	8.1
Washington Dulles Intl.	5.3	10.9
Greater Cincinnati Airport	5.8	12.3

a. Determine a linear equation that approximates the data using the points $(.7, 1.4)$ and $(5.3, 10.9)$.

b. In 1992, 4.9 million passengers used Raleigh–Durham International Airport. Using the equation from part a, approximate the number of passengers using this airport in 2005 and compare it with the Federal Aviation Administration's estimate of 10.3 million passengers.

LIFE SCIENCES

63. ***HIV Infection*** The time interval between a person's initial infection with HIV and that person's eventual development of AIDS symptoms is an important issue. The method of infection with HIV affects the time interval before AIDS develops. One study of HIV patients who were infected by intravenous drug use found that 17% of the patients had AIDS after 4 years, and 33% had developed the disease after 7 years. The relationship between the time interval and the percentage of patients with AIDS can be modeled accurately with a linear equation.†

a. Write a linear equation $y = mx + b$ that models this data, using the ordered pairs $(4, 9.17)$ and $(7, 9.33)$.

b. Use your equation from part a to predict the number of years before half of these patients will have AIDS.

64. ***Exercise Heart Rate*** To achieve the maximum benefit for the heart when exercising, your heart rate (in beats per minute) should be in the target heart rate zone. The lower limit of this zone is found by taking 70% of the difference between 220 and your age. The upper limit is found by using 85%.‡

a. Find formulas for the upper and lower limits (u and l) as linear equations involving the age x.

b. What is the target heart rate zone for a 20-year-old?

c. What is the target heart rate zone for a 40-year-old?

d. Two women in an aerobics class stop to take their pulse, and are surprised to find that they have the same pulse. One woman is 36 years older than the other and is working at the upper limit of her target heart rate zone. The younger woman is working at the lower limit of her target heart rate zone. What are the ages of the two women, and what is their pulse?

e. Run for 10 minutes, take your pulse, and see if it is in your target heart rate zone. (After all, this is listed as an exercise!)

65. ***Ponies Trotting*** A 1991 study found that the peak vertical force on a trotting Shetland pony increased linearly with the pony's speed, and that when the force reached a critical level, the pony switched from a trot to a gallop.§ For one pony, the critical force was 1.16 times its body weight. It

*Federal Aviation Administration.

†Alcabes, P., A. Munoz, D. Vlahov, and G. Friedland, "Incubation Period of Human Immunodeficiency Virus," *Epidemiologic Review,* Vol. 15, No. 2, The Johns Hopkins University School of Hygiene and Public Health, 1993.

‡Hockey, Robert V., *Physical Fitness: The Pathway to Healthful Living,* Times Mirror/Mosby College Publishing, 1989, pp. 85–87.

§*Science*, July 19, 1991, pp. 306–308.

ISBN: 0-536-10718-1

experienced a force of .75 times its body weight at a speed of 2 meters per second, and a force of .93 times its body weight at 3 meters per second. At what speed did the pony switch from a trot to a gallop?

66. ***Life Span*** Some scientists believe there is a limit to how long humans can live.* One supporting argument is that during the last century, life expectancy from age 65 has increased more slowly than life expectancy from birth, so eventually these two will be equal, at which point, according to these scientists, life expectancy should increase no further. In 1900, life expectancy at birth was 46 yr, and life expectancy at age 65 was 76. In 2000, these figures had risen to 76.9 and 82.9, respectively. In both cases, the increase in life expectancy has been linear. Using these assumptions and the data given, find the maximum life expectancy for humans.

67. ***Deer Ticks*** Deer ticks cause concern because they can carry Lyme disease. One study found a relationship between the density of acorns produced in the fall and the density of deer tick larvae the following spring.† The relationship can be approximated by the linear equation

$$y = 34x + 230,$$

where x is the number of acorns per square meter (m^2) in the fall, and y is the number of deer tick larvae per 400 m^2 the following spring. According to this formula, approximately how many acorns per square meter would result in 1000 deer tick larvae per 400 m^2?

SOCIAL SCIENCES

68. ***Immigration*** In 1974, there were 86,821 people from other countries who immigrated to the state of California. In 2000, the number of immigrants was 217,753.‡

a. If the change in foreign immigration to California is considered to be linear, write an equation expressing the number of immigrants, y, in terms of the number of years after 1974, x.

b. Use your result in part a to predict the foreign immigration to California in the year 2010.

69. ***Cohabitation*** The number of unmarried couples in the United States who are living together has been rising at a roughly linear rate in recent years. The number of cohabiting adults was 1.1 million in 1977 and 5.5 million in 2000.§

a. Write an equation expressing the number of cohabiting adults (in millions), y, in terms of the number of years after 1977, x.

b. Use your result in part a to predict the number of cohabiting adults in the year 2010.

70. ***Older College Students*** The percentage of college students who are age 35 and older has been increasing at roughly a linear rate. In 1970 the percentage was 9.6%. In 2001, an estimated 19.2% were 35 or older.‖

a. Find an equation giving the percentage of college students age 35 and older in terms of time t, where t represents the number of years since 1970.

b. If this linear trend continues, what percentage of college students will be 35 and over in 2010?

c. If this linear trend continues, in what year will the percentage of college students 35 and over reach 31%?

71. ***Consumer Price Index*** The Consumer Price Index (CPI) is a measure of the change in the cost of goods over time. If 1977 is used as the base year of comparison (CPI = 100 in 1977), then the CPI of 252.3 in 2000 would indicate that an item that cost \$1.00 in 1977 would cost \$2.52 in 2000. The CPI has been increasing at an approximately linear rate for the past 30 years.#

a. Use this information to determine a linear function for this data, letting x be the years since 1977.

b. Based on your function, what was the CPI in 1995? Compare this estimate with the actual CPI of 224.7.

c. How is the annual CPI changing?

PHYSICAL SCIENCES

72. ***Global Warming*** In 1990, the Intergovernmental Panel on Climate Change predicted that the average temperature on Earth would rise .3°C per decade in the absence of international controls on greenhouse emissions.** Let t measure the time in years since 1970, when the average global temperature was 15°C.

a. Find a linear equation giving the average global temperature in degrees Celsius in terms of t, the number of years since 1970.

b. Scientists have estimated that the sea level will rise by 65 cm if the average global temperature rises to 19°C.

**Science*, Nov. 15, 1991, pp. 936–938 and *The World Almanac and Book of Facts 2003,* p. 75.

†*Science*, Vol. 281, No. 5375, July 17, 1998, pp. 350–351.

‡*Legal Immigration to California in Federal Fiscal Year 1996,* State of California Demographic Research Unit, June 1999 and *The World Almanac and Book of Facts 2003,* p. 405.

§*The New York Times,* Feb. 15, 2000, p. F8 and U.S. Census Bureau Report, March 13, 2003.

‖National Center for Education Statistics.

#U.S. Census Bureau, Income 2000, http://www.census.gov/hhes/income/income00/cpiurs.html.

***Science News,* June 23, 1990, p. 391.

ISBN: 0-536-10718-1

According to your answer to part a, when would this occur?

73. ***Galactic Distance*** The table lists the distances (in megaparsecs where 1 megaparsec $\approx 3.1 \cdot 10^{19}$ km) and velocities (in kilometers per second) of four galaxies moving rapidly away from Earth.*

Galaxy	Distance	Velocity
Virga	15	1600
Ursa Minor	200	15,000
Corona Borealis	290	24,000
Bootes	520	40,000

a. Plot the data points letting x represent distance and y represent velocity. Do the points lie in an approximately linear pattern?

b. Write a linear equation $y = mx$ to model this data, using the ordered pair $(520, 40{,}000)$.

c. The galaxy Hydra has a velocity of 60,000 km per sec. Use your equation to determine how far away it is from Earth.

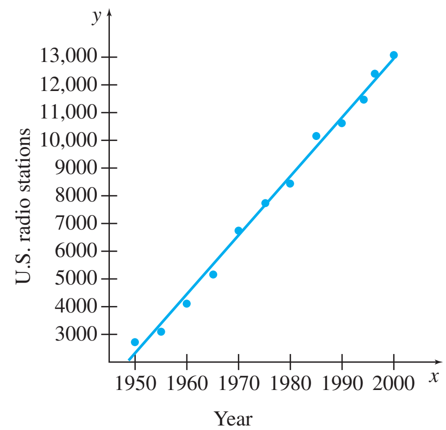
d. The value of m in the equation is called the *Hubble constant*. The Hubble constant can be used to estimate the age of the universe A (in years) using the formula

$$A = \frac{9.5 \times 10^{11}}{m}.$$

Approximate A using your value of m.

GENERAL INTEREST

74. ***Radio Stations*** The graph shows the number of U.S. radio stations on the air along with the graph of a linear equation that models the data.†



a. Use the two ordered pairs $(1950, 2773)$ and $(2000, 13{,}150)$ to find the approximate slope of the line shown. Interpret your answer.

b. Use the same two ordered pairs to write an equation of the line that models the data.

c. Estimate the year when it is expected that the number of stations will first exceed 15,000.

75. ***Tuition*** The table lists the average annual cost (in dollars) of tuition and fees at private four-year colleges for selected years.‡ See Example 15.

Year	Tuition and Fees
1990	9340
1992	10,449
1994	11,719
1996	12,994
1998	14,508
2000	16,332
2002	18,273

a. Sketch a graph of the data. Do the data appear to lie roughly along a straight line?

b. Let $x = 90$ correspond to 1990. Use the points $(90, 9340)$ and $(100, 16{,}332)$ to determine a linear equation that models the data. What does the slope of the graph of the equation indicate?

c. Discuss the accuracy of using this equation to estimate the cost of private college in 2020.

*Acker, A. and C. Jaschek, *Astronomical Methods and Calculations,* John Wiley & Sons, 1986; Karttunen, H. (editor), *Fundamental Astronomy,* Springer-Verlag, 1994.
†National Association of Broadcasters.
‡The College Board.

ISBN: 0-536-10718-1

1.2 LINEAR FUNCTIONS AND APPLICATIONS

THINK ABOUT IT *How many units must be sold for a firm to break even?*

Later in this section, this question will be answered using a linear function.

As we saw in the previous section, many situations involve two variables related by a linear equation. For such a relationship, when we express the variable y in terms of x, we say that y is a **linear function** of x. This means that for any allowed value of x (the **independent variable**), we can use the equation to find the corresponding value of y (the **dependent variable**). Examples of linear functions include $y = 2x + 3$, $y = -5$, and $2x - 3y = 7$, which can be written as $y = (2/3)x - (7/3)$. Equations in the form $x = k$, where k is a constant, are not linear functions. All other linear equations define linear functions.

$f(x)$ Notation

Letters such as f, g, or h are often used to name functions. For example, f might be used to name the function

$$y = 5 - 3x.$$

To show that this function is named f, it is common to replace y with $f(x)$ (read "f of x") to get

$$f(x) = 5 - 3x.$$

By choosing 2 as a value of x, $f(x)$ becomes $5 - 3 \cdot 2 = 5 - 6 = -1$, written

$$f(2) = -1.$$

The corresponding ordered pair is $(2, -1)$. In a similar manner,

$$f(-4) = 5 - 3(-4) = 17, \qquad f(0) = 5, \qquad f(-6) = 23,$$

and so on.

EXAMPLE 1 *Function Notation*

Let $g(x) = -4x + 5$. Find $g(3)$, $g(0)$, and $g(-2)$.

Solution To find $g(3)$, substitute 3 for x.

$$g(3) = -4(3) + 5 = -12 + 5 = -7$$

Similarly,

$$g(0) = -4(0) + 5 = 0 + 5 = 5,$$

and

$$g(-2) = -4(-2) + 5 = 8 + 5 = 13.$$

We summarize the discussion below.

ISBN: 0-536-10718-1

LINEAR FUNCTION

A relationship f defined by

$$y = f(x) = mx + b,$$

for real numbers m and b, is a **linear function.**

Supply and Demand Linear functions are often good choices for **supply and demand curves.** Typically, as the price of an item increases, the demand for the item decreases, while the supply increases. On the other hand, when demand for an item increases, so does its price, causing the supply of the item to decrease.

For example, during the 1970s the price of gasoline increased rapidly. As the price continued to escalate, most buyers became more and more prudent in their use of gasoline in order to restrict their demand to an affordable amount. Consequently, the overall demand for gasoline decreased and the supply increased, to a point where there was an oversupply of gasoline. This caused prices to fall until supply and demand were approximately balanced. Many other factors were involved in the situation, but the relationship between price, supply, and demand was nonetheless typical. Some commodities, however, such as medical care, college education, and certain luxury items, may be exceptions to these typical relationships.

Although economists consider price to be the independent variable, they have the unfortunate habit of plotting price, usually denoted by p, on the vertical axis, while everyone else graphs the independent variable on the horizontal axis. This custom was started by the English economist Alfred Marshall (1842–1924). In order to abide by this custom, we will write p, the price, as a function of q, the quantity produced, and plot p on the vertical axis. But remember, it is really *price* that determines how much consumers demand and producers supply, not the other way around.

Supply and demand functions are not necessarily linear, the simplest kind of function. Yet most functions are approximately linear if a small enough piece of the graph is taken, allowing applied mathematicians to often use linear functions for simplicity. That approach will be taken in this chapter.

EXAMPLE 2 *Supply and Demand*

Suppose that Greg Tobin, an economist, has studied the supply and demand for vinyl siding and has determined that the price (in dollars) per square yard (yd^2), p, and the quantity demanded monthly (in thousands of square yards), q, are related by the linear function

$$p = D(q) = 60 - \frac{3}{4}q, \qquad \text{Demand}$$

while the price p and the supply q are related by

$$p = S(q) = \frac{3}{4}q. \qquad \text{Supply}$$

(a) Find the demand at a price of \$45 and at a price of \$18.

Solution Start with the demand function

$$p = 60 - \frac{3}{4}q,$$

ISBN: 0-536-10718-1

FOR REVIEW

In the second-to-last step of the solution in Example 2(a), q was multiplied by $-3/4$, so both sides of the equation had to be divided by $-3/4$. This was done by multiplying both sides by $-4/3$. The way to divide by a fraction is to multiply by its reciprocal. In other words, to divide by the fraction a/b, multiply by b/a.

and replace p with 45.

$$45 = 60 - \frac{3}{4}q$$

$$-15 = -\frac{3}{4}q \qquad \text{Subtract 60 from both sides.}$$

$$20 = q \qquad \text{Multiply both sides by } -\tfrac{4}{3}.$$

Thus, at a price of \$45, the demand is 20,000 yd^2 per month.

Similarly, replace p with 18 to find the demand when the price is \$18. Verify that this leads to $q = 56$. When the price is lowered from \$45 to \$18, the demand increases from 20,000 yd^2 to 56,000 yd^2.

(b) Find the supply at a price of \$60 and at a price of \$12.

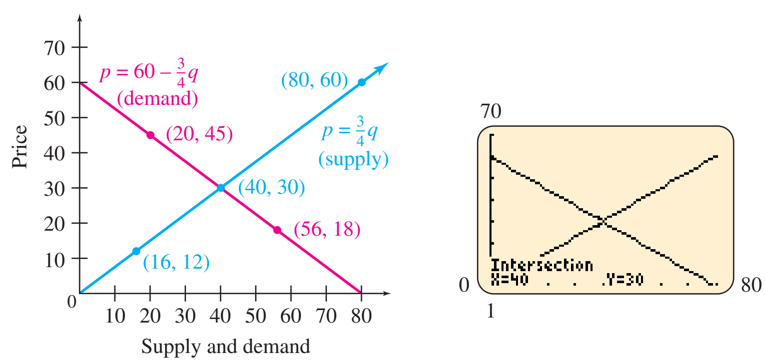
Solution Substitute 60 for p in the supply equation,

$$p = \frac{3}{4}q,$$

to find that $q = 80$, so the supply is 80,000 yd^2. Similarly, replacing p with 12 in the supply equation gives a supply of 16,000 yd^2. If the price decreases from \$60 to \$12, the supply also decreases, from 80,000 yd^2 to 16,000 yd^2.

(c) Graph both functions on the same axes.

Solution The results of part (a) are written as the ordered pairs $(20, 45)$ and $(56, 18)$. The line through those points is the graph of $p = 60 - (3/4)q$, shown in red in Figure 11(a). We used the ordered pairs $(80, 60)$ and $(16, 12)$ from the work in part (b) to get the supply graph shown in blue in Figure 11(a).



(a) **(b)**

FIGURE 11

A calculator-generated graph of the lines representing the supply and demand in Example 2 is shown in Figure 11(b). The equation of each line, using x and y instead of q and p, was entered along with an appropriate window to get this graph. A special menu choice gives the coordinates of the intersection point, shown at the bottom of the graph.

ISBN: 0-536-10718-1

NOTE Not all supply/demand problems will have the same scale on both axes. It helps to consider the intercepts of both the supply and demand graphs to decide what scale to use. For example, in Figure 11, the y-intercept of the demand function is 60, so the scale should allow values from 0 to at least 60 on the vertical axis. The x-intercept of the demand function is 80, so values on the x-axis must go from 0 to 80. Letting each tick mark represent 10 gives a reasonable number of marks on each axis. ■

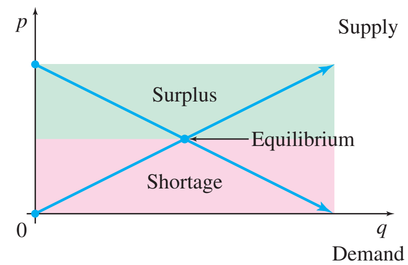


FIGURE 12

As shown in the graphs of Figure 11, both the supply and the demand graphs pass through the point $(40, 30)$. If the price of a square yard of siding is more than \$30, the supply will exceed the demand. At a price less than \$30, the demand will exceed the supply. Only at a price of \$30 will demand and supply be equal. For this reason, \$30 is called the *equilibrium price.* When the price is \$30, demand and supply both equal 40,000 yd^2, the *equilibrium quantity.* In general, the **equilibrium price** of a commodity is the price found at the point where the supply and demand graphs for that commodity intersect. The **equilibrium quantity** is the demand and the supply at that same point. Figure 12 illustrates a general supply and demand situation.

EXAMPLE 3 *Equilibrium Quantity*

Use algebra to find the equilibrium quantity for the vinyl siding in Example 2.

Solution The equilibrium quantity is found when the prices from both supply and demand are equal. Set the two expressions for p equal to each other and solve.

$$
\begin{aligned}
60 - \frac{3}{4}q &= \frac{3}{4}q \\
240 - 3q &= 3q && \text{Multiply both sides by 4.} \\
240 &= 6q && \text{Add } 3q \text{ to both sides.} \\
40 &= q
\end{aligned}
$$

The equilibrium quantity is 40,000 yd^2, the same answer found earlier.

You may prefer to find the equilibrium quantity by solving the equation with your calculator. Or, if your calculator has a TABLE feature, you can use it to find the value of q that makes the two expressions equal.

Another important issue is how, in practice, the equations of the supply and demand functions can be found. This issue is important for many problems involving linear functions in this section and the next. Data need to be collected, and if they lie perfectly along a line, then the equation can easily be found with any two points. What usually happens, however, is that the data are scattered, and there is no line that goes through all the points. In this case we must find a line that approximates the linear trend of the data as closely as possible (assuming the points lie approximately along a line) as in Example 15 in the previous section. This is usually done by the *method of least squares,* also referred to as *linear regression.* We will discuss this method in Section 1.3.

ISBN: 0-536-10718-1

Cost Analysis The cost of manufacturing an item commonly consists of two parts. The first is a **fixed cost** for designing the product, setting up a factory, training workers, and so on. Within broad limits, the fixed cost is constant for a particular product and does not change as more items are made. The second part is a *cost per item* for labor, materials, packing, shipping, and so on. The total value of this second cost *does* depend on the number of items made.

EXAMPLE 4 *Cost Analysis*

Suppose that the cost of producing video games can be approximated by

$$C(x) = 12x + 100,$$

where $C(x)$ is the cost in dollars to produce x games. The cost to produce 0 games is

$$C(0) = 12(0) + 100 = 100,$$

or \$100. This sum, \$100, is the fixed cost.

Once the company has invested the fixed cost into the video game project, what will the additional cost per game be? As an example, let's compare the costs of making 5 games and 6 games.

$$C(5) = 12(5) + 100 = 160 \qquad \text{and} \qquad C(6) = 12(6) + 100 = 172,$$

or \$160 and \$172, respectively.

So the 6th game itself costs $\$172 - \$160 = \$12$ to produce. In the same way, the 81st game costs $C(81) - C(80) = \$1072 - \$1060 = \$12$ to produce. In fact, the $(n + 1)$st game costs

$$\begin{aligned} C(n + 1) - C(n) &= [12(n + 1) + 100] - (12n + 100) \\ &= 12, \end{aligned}$$

or \$12, to produce. The number 12 is also the slope of the graph of the cost function $C(x) = 12x + 100$; the slope gives us the cost to produce an additional item.

In economics, **marginal cost** is the rate of change of cost $C(x)$ at a level of production x and is equal to the slope of the cost function at x. It approximates the cost of producing one additional item. In fact, some books define the marginal cost to be the cost of producing one additional item. With *linear functions,* these two definitions are equivalent, and the marginal cost, which is equal to the slope of the cost function, is *constant.* For instance, in the video game example, the marginal cost of each game is \$12. For other types of functions, these two definitions are only approximately equal. Marginal cost is important to management in making decisions in areas such as cost control, pricing, and production planning.

The work in Example 4 can be generalized. Suppose the total cost to make x items is given by the linear cost function $C(x) = mx + b$. The fixed cost is found by letting $x = 0$:

$$C(0) = m \cdot 0 + b = b;$$

ISBN: 0-536-10718-1

thus, the fixed cost is b dollars. The additional cost of the $(n + 1)$st item, the marginal cost, is m, the slope of the line $C(x) = mx + b$.

COST FUNCTION

In a cost function of the form $C(x) = mx + b$, the m represents the marginal cost per item and b the fixed cost. Conversely, if the fixed cost of producing an item is b and the marginal cost is m, then the **cost function** $C(x)$ for producing x items is $C(x) = mx + b$.

EXAMPLE 5 *Cost Function*

The marginal cost to make x tablets of a prescription medication is \$10 per batch, while the cost to produce 100 batches is \$1500. Find the cost function $C(x)$, given that it is linear.

Solution Since the cost function is linear, it can be expressed in the form $C(x) = mx + b$. The marginal cost is \$10 per batch, which gives the value for m, leading to $C(x) = 10x + b$. To find b, use the fact that the cost of producing 100 batches of tablets is \$1500, or $C(100) = 1500$. Substituting $C(x) = 1500$ and $x = 100$ into $C(x) = 10x + b$ gives

$$1500 = 10 \cdot 100 + b$$
$$1500 = 1000 + b$$
$$500 = b. \qquad \text{Subtract 1000 from both sides.}$$

The cost function is given by $C(x) = 10x + 500$, where the fixed cost is \$500.

Break-Even Analysis

The **revenue** $R(x)$ from selling x units of an item is the product of the price per unit p and the number of units sold (demand) x, so that

$$R(x) = px.$$

The corresponding **profit** $P(x)$ is the difference between revenue $R(x)$ and cost $C(x)$. That is,

$$P(x) = R(x) - C(x).$$

A company can make a profit only if the revenue received from its customers exceeds the cost of producing and selling its goods and services. The number of units at which revenue just equals cost is the **break-even quantity;** the corresponding ordered pair gives the **break-even point.**

EXAMPLE 6 *Break-Even Analysis*

A firm producing poultry feed finds that the total cost $C(x)$ in dollars of producing and selling x units is given by

$$C(x) = 20x + 100.$$

ISBN: 0-536-10718-1

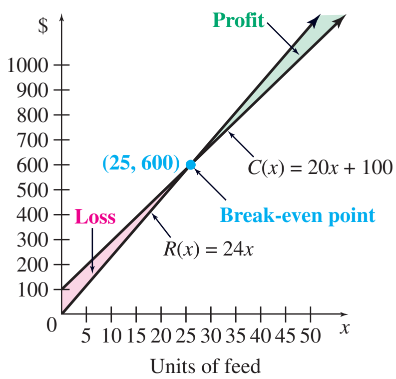


FIGURE 13

Management plans to charge $24 per unit for the feed.

(a) How many units must be sold for the firm to break even?

Solution The firm will break even (no profit and no loss) as long as revenue just equals cost, or $R(x) = C(x)$. From the given information, since $R(x) = px$ and $p = \$24$,

$$R(x) = 24x.$$

Substituting for $R(x)$ and $C(x)$ in the equation $R(x) = C(x)$ gives

$$24x = 20x + 100,$$

from which $x = 25$. The firm breaks even by selling 25 units, which is the break-even quantity. The graphs of $C(x) = 20x + 100$ and $R(x) = 24x$ are shown in Figure 13. The break-even point (where $x = 25$) is shown on the graph. If the company sells more than 25 units (if $x > 25$), it makes a profit. If it sells less than 25 units, it loses money.

(b) What is the profit if 100 units of feed are sold?

Solution Use the formula for profit $P(x)$.

$$\begin{aligned} P(x) &= R(x) - C(x) \\ &= 24x - (20x + 100) \\ &= 4x - 100 \end{aligned}$$

Then $P(100) = 4(100) - 100 = 300$. The firm will make a profit of $300 from the sale of 100 units of feed.

(c) How many units must be sold to produce a profit of $900?

Solution Let $P(x) = 900$ in the equation $P(x) = 4x - 100$ and solve for x.

$$\begin{aligned} 900 &= 4x - 100 \\ 1000 &= 4x \\ x &= 250 \end{aligned}$$

Sales of 250 units will produce $900 profit.

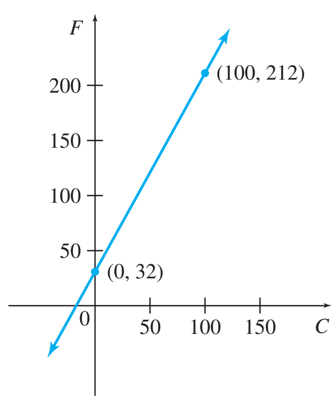


FIGURE 14

Temperature One of the most common linear relationships found in everyday situations deals with temperature. Recall that water freezes at 32° Fahrenheit and 0° Celsius, while it boils at 212° Fahrenheit and 100° Celsius.* The ordered pairs $(0, 32)$ and $(100, 212)$ are graphed in Figure 14 on axes showing Fahrenheit (F) as a function of Celsius (C). The line joining them is the graph of the function.

*Gabriel Fahrenheit (1686–1736), a German physicist, invented his scale with 0° representing the temperature of an equal mixture of ice and ammonium chloride (a type of salt), and 96° as the temperature of the human body. (It is often said, erroneously, that Fahrenheit set 100° as the temperature of the human body. Fahrenheit's own words are quoted in *A History of the Thermometer and Its Use in Meteorology* by W. E. Knowles, Middleton: The Johns Hopkins Press, 1966, p. 75.) The Swedish astronomer Anders Celsius (1701–1744) set 0° and 100° as the freezing and boiling points of water.

ISBN: 0-536-10718-1

EXAMPLE 7 *Temperature*

Derive an equation relating F and C.

Solution To derive the required linear equation, first find the slope using the given ordered pairs, $(0, 32)$ and $(100, 212)$.

$$m = \frac{212 - 32}{100 - 0} = \frac{9}{5}$$

The F-intercept of the graph is 32, so by the slope-intercept form, the equation of the line is

$$F = \frac{9}{5}C + 32.$$

With simple algebra this equation can be rewritten to give C in terms of F:

$$C = \frac{5}{9}(F - 32).$$

1.2 EXERCISES

In Exercises 1–4, decide whether the statement is true or false.

1. To find the x-intercept of the graph of a linear function, we solve $y = f(x) = 0$, and to find the y-intercept, we evaluate $f(0)$.

2. The graph of $f(x) = -3$ is a vertical line.

3. The slope of the graph of a linear function cannot be undefined.

4. The graph of $f(x) = ax$ is a straight line that passes through the origin.

5. Describe what fixed costs and marginal costs mean to a company.

6. In a few sentences, explain why the price of a commodity not already at its equilibrium price should move in that direction.

7. Explain why a linear function may not be adequate for describing the supply and demand functions.

Write a linear cost function for each situation. Identify all variables used.

8. A chain saw rental firm charges \$12 plus \$1 per hour.

9. A trailer-hauling service charges \$45 plus \$2 per mile.

10. A parking garage charges 50 cents plus 35 cents per half-hour.

11. For a one-day rental, a car rental firm charges \$44 plus 28 cents per mile.

Assume that each situation can be expressed as a linear cost function. Find the cost function in each case.

12. Fixed cost: \$100; 50 items cost \$1600 to produce.

13. Fixed cost: \$400; 10 items cost \$650 to produce.

14. Marginal cost: \$90; 150 items cost \$16,000 to produce.

15. Marginal cost: \$120; 700 items cost \$96,500 to produce.

16. How is the average rate of change related to the graph of a function?

17. In your own words, describe the break-even quantity, how to find it, and what it indicates.

ISBN: 0-536-10718-1

Applications

BUSINESS AND ECONOMICS

18. ***Supply and Demand*** Suppose that the demand and price for a certain model of electric can opener are related by

$$p = D(q) = 16 - \frac{5}{4}q,$$

where p is the price (in dollars) and q is the demand (in hundreds). Find the price at each level of demand.

a. 0 can openers **b.** 400 can openers

c. 800 can openers

Find the demand for the electric can opener at each price.

d. \$8 **e.** \$10 **f.** \$12

g. Graph $p = 16 - \frac{5}{4}q$.

Suppose the price and supply of the electric can opener are related by

$$p = S(q) = \frac{3}{4}q,$$

where p is the price (in dollars) and q is the supply (in hundreds) of can openers. Find the supply at each price.

h. \$0 **i.** \$10 **j.** \$20

k. Graph $p = \frac{3}{4}q$ on the same axes used for part g.

l. Find the equilibrium quantity and the equilibrium price.

19. ***Supply and Demand*** Let the supply and demand functions for strawberry-flavored licorice be given by

$$p = S(q) = \frac{3}{2}q \quad \text{and} \quad p = D(q) = 81 - \frac{3}{4}q,$$

where p is the price in dollars and q is the number of batches.

a. Graph these on the same axes.

b. Find the equilibrium quantity and the equilibrium price.

20. ***Supply and Demand*** Let the supply and demand functions for butter pecan ice cream be given by

$$p = S(q) = \frac{2}{5}q \quad \text{and} \quad p = D(q) = 100 - \frac{2}{5}q,$$

where p is the price in dollars and q is the number of 10-gallon tubs.

a. Graph these on the same axes.

b. Find the equilibrium quantity and the equilibrium price.

21. ***Supply and Demand*** Let the supply and demand functions for sugar be given by

$$p = S(q) = 1.4q - .6 \quad \text{and}$$
$$p = D(q) = -2q + 3.2,$$

where p is the price per pound and q is the quantity in thousands of pounds.

a. Graph these on the same axes.

b. Find the equilibrium quantity and the equilibrium price.

22. ***T-Shirt Cost*** Yoshi Yamamura sells silk-screened T-shirts at community festivals and crafts fairs. Her marginal cost to produce one T-shirt is \$3.50. Her total cost to produce 60 T-shirts is \$300, and she sells them for \$9 each.

a. Find the linear cost function for Yoshi's T-shirt production.

b. How many T-shirts must she produce and sell in order to break even?

c. How many T-shirts must she produce and sell to make a profit of \$500?

23. ***Publishing Costs*** Enrique Gonzales owns a small publishing house specializing in Latin American poetry. His fixed cost to produce a typical poetry volume is \$525, and his total cost to produce 1000 copies of the book is \$2675. His books sell for \$4.95 each.

a. Find the linear cost function for Enrique's book production.

b. How many poetry books must he produce and sell in order to break even?

c. How many books must he produce and sell to make a profit of \$1000?

24. ***Marginal Cost of Coffee*** The manager of a restaurant found that the cost to produce 100 cups of coffee is \$11.02, while the cost to produce 400 cups is \$40.12. Assume the cost $C(x)$ is a linear function of x, the number of cups produced.

a. Find a formula for $C(x)$.

b. What is the fixed cost?

c. Find the total cost of producing 1000 cups.

d. Find the total cost of producing 1001 cups.

e. Find the marginal cost of the 1001st cup.

f. What is the marginal cost of *any* cup and what does this mean to the manager?

25. ***Marginal Cost of a New Plant*** In deciding whether to set up a new manufacturing plant, company analysts have

ISBN: 0-536-10718-1

decided that a linear function is a reasonable estimation for the total cost $C(x)$ in dollars to produce x items. They estimate the cost to produce 10,000 items as \$547,500, and the cost to produce 50,000 items as \$737,500.

a. Find a formula for $C(x)$.

b. Find the fixed cost.

c. Find the total cost to produce 100,000 items.

d. Find the marginal cost of the items to be produced in this plant and what does this mean to the manager?

26. ***Bread Sales*** Bread Boutiques, which sell freshly baked bread with no preservatives, are located in many malls around the United States and are growing rapidly. The Saint Louis Bread Company (now called Panera Bread) claims a sales growth of 5000% in its first five years.*

a. Suppose sales were \$100,000 in 1991. What would they be in 1996 at that growth rate?

b. Let 1991 correspond to $x = 1$. Write two ordered pairs representing sales in 1991 and 1996.

c. Assuming sales increased linearly, write a linear sales function for this company.

d. If sales continue to increase at the same rate, when will they reach one billion dollars?

e. The actual sales were expected to be \$1 billion in 2003. Discuss the assumption that the growth rate has been linear.

27. ***Break-Even Analysis*** Producing x units of tacos costs $C(x) = 5x + 20$; revenue is $R(x) = 15x$, where $C(x)$ and $R(x)$ are in dollars.

a. What is the break-even quantity?

b. What is the profit from 100 units?

c. How many units will produce a profit of \$500?

28. ***Break-Even Analysis*** To produce x units of a religious medal costs $C(x) = 12x + 39$. The revenue is $R(x) = 25x$. Both $C(x)$ and $R(x)$ are in dollars.

a. Find the break-even quantity.

b. Find the profit from 250 units.

c. Find the number of units that must be produced for a profit of \$130.

Break-Even Analysis *You are the manager of a firm. You are considering the manufacture of a new product, so you ask the accounting department for cost estimates and the sales department for sales estimates. After you receive the data, you must decide whether to go ahead with production of the new product. Analyze the data in Exercises 29–32 (find a break-even quantity) and then decide what you would do in each case. Also write the profit function.*

29. $C(x) = 85x + 900$; $R(x) = 105x$; no more than 38 units can be sold.

30. $C(x) = 105x + 6000$; $R(x) = 250x$; no more than 400 units can be sold.

31. $C(x) = 70x + 500$; $R(x) = 60x$ (*Hint*: What does a negative break-even quantity mean?)

32. $C(x) = 1000x + 5000$; $R(x) = 900x$

PHYSICAL SCIENCES

33. ***Temperature*** Use the formula for conversion between Fahrenheit and Celsius derived in Example 7 to convert each temperature.

a. 58°F to Celsius

b. −20°F to Celsius

c. 50°C to Fahrenheit

34. ***Body Temperature*** You may have heard that the average temperature of the human body is 98.6°. Recent experiments show that the actual figure is closer to 98.2°.† The figure of 98.6 comes from experiments done by Carl Wunderlich in 1868. But Wunderlich measured the temperatures in degrees Celsius and rounded the average to the nearest degree, giving 37°C as the average temperature.‡

a. What is the Fahrenheit equivalent of 37°C?

b. Given that Wunderlich rounded to the nearest degree Celsius, his experiments tell us that the actual average human body temperature is somewhere between 36.5°C and 37.5°C. Find what this range corresponds to in degrees Fahrenheit.

35. ***Temperature*** Find the temperature at which the Celsius and Fahrenheit temperatures are numerically equal.

**The New York Times,* Nov. 18, 1995, pp. 19 and 21.
†*Science News,* Sept. 26, 1992, p. 195.
‡*Science News,* Nov. 7, 1992, p. 399.

ISBN: 0-536-10718-1

1.3 THE LEAST SQUARES LINE

THINK ABOUT IT *How has the accidental death rate in the United States changed over time?*

In this section, we show how to answer such questions using the method of least squares. We use past data to find trends and to make tentative predictions about the future. The only assumption we make is that the data are related linearly—that is, if we plot pairs of data, the resulting points will lie close to some line. This method cannot give exact answers. The best we can expect is that, if we are careful, we will get a reasonable approximation.

Year	Death Rate
1910	84.4
1920	71.2
1930	80.5
1940	73.4
1950	60.3
1960	52.1
1970	56.2
1980	46.5
1990	36.9
2000	34.0

The table lists the number of accidental deaths per 100,000 population in the United States through the past century.* If you were a manager at an insurance company, these data could be very important. You might need to make some predictions about how much you will pay out next year in accidental death benefits, and even a very tentative prediction based on past trends is better than no prediction at all.

The first step is to draw a scatterplot, as we have done in Figure 15. Notice that the points lie approximately along a line, which means that a linear function may give a good approximation of the data. If we select two points and find the line that passes through them, as we did in Section 1.1, we will get a different line for each pair of points, and in some cases the lines will be very different. We want to draw one line that is simultaneously close to all the points on the graph, but many such lines are possible, depending upon how we define the phrase "simultaneously close to all the points." How do we decide on the best possible line? Before going on, you might want to try drawing the line you think is best on Figure 15.

The line used most often in applications is that in which the sum of the squares of the vertical distances from the data points to the line is as small as possible. Such a line is called the **least squares line.** The least squares line for the data in Figure 15 is drawn in Figure 16. How does the line compare with the one you drew on Figure 15? It may not be exactly the same, but should appear similar.

In Figure 16, the vertical distances from the points to the line are indicated by d_1, d_2, and so on, up through d_{10} (read "d-sub-one, d-sub-two, d-sub-three," and so on). For n points, corresponding to the n pairs of data, the least squares line is found by minimizing the sum $(d_1)^2 + (d_2)^2 + (d_3)^2 + \cdots + (d_n)^2$.

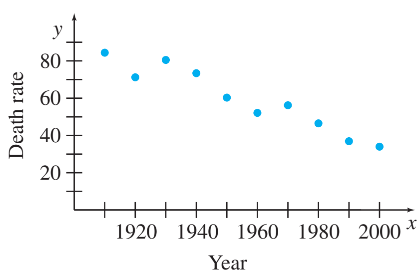


FIGURE 15

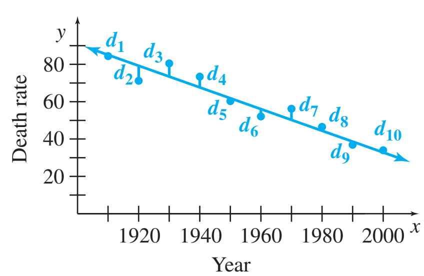


FIGURE 16

*U.S. Department of Health and Human Services, National Center for Health Statistics.

ISBN: 0-536-10718-1

For the points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, if the equation of the desired line is $Y = mx + b$, then

$$d_1 = |Y_1 - y_1| = |mx_1 + b - y_1|,$$
$$d_2 = |Y_2 - y_2| = |mx_2 + b - y_2|,$$

and so on. We use Y in the equation of the line instead of y to distinguish the predicted values (Y) from the y-values of the given data points. The sum to be minimized becomes

$$(mx_1 + b - y_1)^2 + (mx_2 + b - y_2)^2 + \cdots + (mx_n + b - y_n)^2,$$

where $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ are known and m and b are to be found.

The method of minimizing this sum requires advanced techniques and is not given here. The result gives equations that can be solved for the slope m and the y-intercept b of the least squares line.* In these equations, the symbol Σ, the Greek letter sigma, indicates "the sum of"; this notation is known as **summation notation.** For example, we write the sum $x_1 + x_2 + \cdots + x_n$, where n is the number of data points, as

$$x_1 + x_2 + \cdots + x_n = \Sigma x.$$

Similarly, Σxy means $x_1y_1 + x_2y_2 + \cdots + x_ny_n$, and so on.

CAUTION Note that Σx^2 means $x_1^2 + x_2^2 + \cdots + x_n^2$, which is *not* the same as squaring Σx. ■

LEAST SQUARES LINE

The **least squares line** $Y = mx + b$ that gives the best fit to the data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ has slope m and y-intercept b that satisfy the equations

$$nb + (\Sigma x)m = \Sigma y$$
$$(\Sigma x)b + (\Sigma x^2)m = \Sigma xy.$$

Method 1: Calculating by Hand

To find the least squares line for the given data, we first find the required sums. To reduce the size of the numbers, let x represent the years since 1900, so that, for example, $x = 10$ for the year 1910. Let y represent the death rate.

x	y	xy	x^2	y^2
10	84.4	844	100	7123.36
20	71.2	1424	400	5069.44
30	80.5	2415	900	6480.25
40	73.4	2936	1600	5387.56
50	60.3	3015	2500	3636.09
60	52.1	3126	3600	2714.41
70	56.2	3934	4900	3158.44
80	46.5	3720	6400	2162.25
90	36.9	3321	8100	1361.61
100	34.0	3400	10,000	1156.00
$\Sigma x = 550$	$\Sigma y = 595.5$	$\Sigma xy = 28,135$	$\Sigma x^2 = 38,500$	$\Sigma y^2 = 38,249.41$

*Equations for m and b are derived in Exercise 4.

ISBN: 0-536-10718-1

(The column headed y^2 will be used later.) Now we can calculate m and b by solving a system of equations. Our method is to solve the first equation for b in terms of m, and substitute this into the second equation. We then solve the second equation for m. Once we have m, we put this back into the equation for b. Here, $n = 10$ (the number of data points).

$$nb + (\sum x)m = \sum y \qquad \text{First least squares equation.}$$
$$10b + 550m = 595.5 \qquad \text{Substitute from the table.}$$
$$10b = 595.5 - 550m \qquad \text{Subtract } 550m \text{ from both sides.}$$
$$b = (595.5 - 550m)/10 \qquad \text{Divide both sides by } 10.$$
$$(\sum x)b + (\sum x^2)m = \sum xy \qquad \text{Second least squares equation.}$$
$$550(595.5 - 550m)/10 + 38{,}500m = 28{,}135 \qquad \text{Substitute.}$$
$$32{,}752.5 - 30{,}250m + 38{,}500m = 28{,}135 \qquad \text{Multiply.}$$
$$8250m = -4617.5 \qquad \text{Combine terms.}$$
$$m \approx -.5596970 \approx -.560$$

The significance of m is that the death rate per 100,000 population is tending to drop (because of the negative) at a rate of .560 per year.

Now substitute the value of m into the equation for b.

$$b = \frac{595.5 - 550(-.5596970)}{10} \approx 90.333335 \approx 90.3$$

Substitute m and b into the least squares line equation, $Y = mx + b$; the least squares line that best fits the nine data points has equation $Y = -.560x + 90.3$. This gives a mathematical description of the relationship between the year and the number of accidental deaths per 100,000 population. The equation can be used to predict y from a given value of x, as we will show in Example 1. As we mentioned before, however, caution must be exercised when using the least squares equation to predict data points that are far from the range of points on which the equation was modeled.

CAUTION In computing m and b, we rounded the final answer to three digits because the original data were known only to three digits. It is important, however, *not* to round any of the intermediate results (such as $\sum x^2$) because round-off error may have a detrimental effect on the accuracy of the answer. Similarly, it is important not to use a rounded-off value of m when computing b. ■

Method 2: Graphing Calculator

The calculations for finding the least squares line are often tedious, even with the aid of a calculator. Fortunately, many calculators can calculate the least squares line with just a few keystrokes. For purposes of illustration, we will show how the least squares line in the previous example is found with a TI-83/84 Plus graphing calculator.

We begin by entering the data into the calculator. We will be using the first two lists, called L_1 and L_2. Choosing the STAT menu, then choosing the fourth entry ClrList, we enter L_1, L_2, to indicate the lists to be cleared. Now we press STAT again and choose the first entry EDIT, which brings up the blank lists. As before, we will only use the last two digits of the year, putting the numbers in L_1. We put the death rate in L_2, giving the two screens shown in Figure 17.

ISBN: 0-536-10718-1

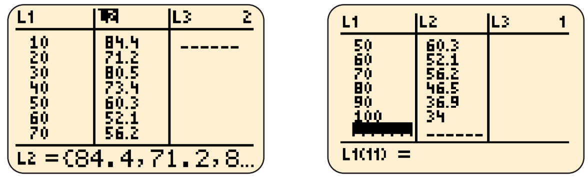


FIGURE 17

Press STAT again and choose CALC instead of EDIT. Then choose item 4 LinReg $(ax + b)$ to get the values of a (the slope) and b (the y-intercept) for the least squares line, as shown in Figure 18. With a and b rounded to three decimal places, the least squares line is $Y = -.560x + 90.3$. A graph of the data points and the line is shown in Figure 19.

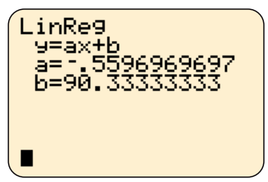


FIGURE 18

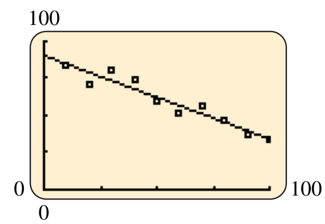


FIGURE 19

For more details on finding the least squares line with a graphing calculator, see *The Graphing Calculator Manual* that is available with this book.

Method 3: Spreadsheet

Many computer spreadsheet programs can also find the least squares line. Figure 20 shows the scatterplot and least squares line for the accidental death rate data using an Excel spreadsheet. The scatterplot was found using the XY(Scatter) command under Chart Wizard, and the line was found using the Add Trendline command under the Chart menu. For details, see *The Spreadsheet Manual* that is available with this book.

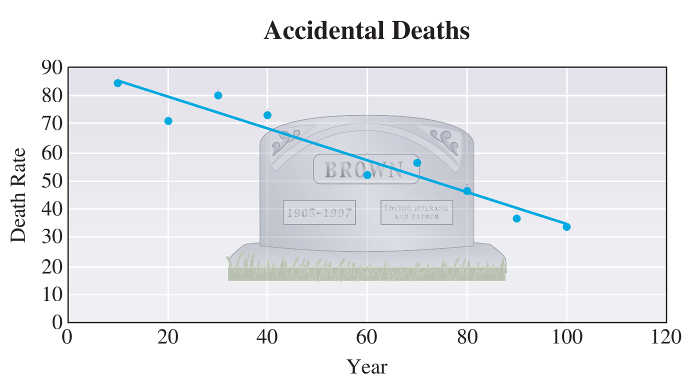


FIGURE 20

ISBN: 0-536-10718-1

EXAMPLE 1 *Least Squares Line*
What do we predict the accidental death rate in 2001 to be?

Solution Use the least squares line equation given above with $x = 101$.

$$\begin{aligned} Y &= -.560x + 90.3 \\ &= -.560(101) + 90.3 \\ &= 33.74 \end{aligned}$$

The death rate in 2001 is predicted to be about 33.7 per 100,000 population. In this case, we have the actual value for 2001. It happens to be 34.3, which is close to the predicted value.

EXAMPLE 2 *Least Squares Line*
In what year is the death rate predicted to drop below 26 per 100,000 population?

Solution Let $Y = 26$ in the equation above and solve for x.

$$\begin{aligned} 26 &= -.560x + 90.3 \\ -64.3 &= -.560x && \text{Subtract 90.3 from both sides.} \\ x &= 114.8 && \text{Divide both sides by } -.560. \end{aligned}$$

This means that after 114 years, the rate will not have quite reached 26 per 100,000, so we must wait 115 years for this to happen. This corresponds to the year 2015 (115 years after 1900), when our equation predicts the death rate to be $-.560(115) + 90.3 = 25.9$ per 100,000 population.

Correlation

Once an equation is found for the least squares line, we need to have some way of judging just how good the equation is for predictive purposes. If the points from the data fit the line quite closely, then we have more reason to expect future data pairs to do so. But if the points are widely scattered about even the best-fitting line, then predictions are not likely to be accurate.

In order to have a quantitative basis for confidence in our predictions, we need a measure of the "goodness of fit" of the original data to the prediction line. One such measure is called the **coefficient of correlation,** denoted r.

COEFFICIENT OF CORRELATION

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \cdot \sqrt{n(\sum y^2) - (\sum y)^2}}$$

Although the expression for r looks daunting, remember that each of the summations, $\sum x$, $\sum y$, $\sum xy$, and so on, are just the totals from a table like the one we prepared for the data on accidental deaths. Also, with a calculator, the arithmetic is no problem!

The coefficient of correlation r is always equal to or between 1 and -1. Values of exactly 1 or -1 indicate that the data points lie *exactly* on the least squares

ISBN: 0-536-10718-1

line. If $r = 1$, the least squares line has a positive slope; $r = -1$ gives a negative slope. If $r = 0$, there is no linear correlation between the data points (but some *nonlinear* function might provide an excellent fit for the data). A correlation of zero may also indicate that the data fit a horizontal line. To investigate what is happening, it is always helpful to sketch a scatterplot of the data. Some scatterplots that correspond to these values of r are shown in Figure 21.

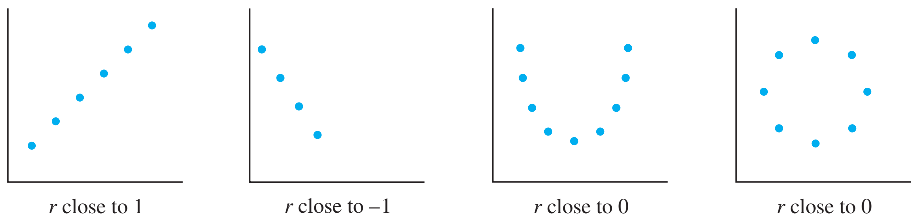


FIGURE 21

A value of r close to 1 or -1 indicates the presence of a linear relationship. The exact value of r necessary to conclude that there is a linear relationship depends upon n, the number of data points, as well as how confident we want to be of our conclusion. For details, consult a text on statistics.*

EXAMPLE 3 *Coefficient of Correlation*

Find r for the data on accidental death rates.

Solution

Method 1: Calculating by Hand

From the table on page 31, $\sum x = 550$, $\sum y = 595.5$, $\sum xy = 28{,}135$, $\sum x^2 = 38{,}500$, and $\sum y^2 = 38{,}249.41$. Also, $n = 10$. Substituting these values into the formula for r gives

$$\begin{aligned} r &= \frac{10(28{,}135) - (550)(595.5)}{\sqrt{10(38{,}500) - (550)^2} \cdot \sqrt{10(38{,}249.41) - (595.5)^2}} \\ &= \frac{-46{,}175}{\sqrt{82{,}500} \cdot \sqrt{27{,}873.85}} \\ &= -.962900585 \approx -.963. \end{aligned}$$

This is a high correlation, which agrees with our observation that the data fit a line quite well.

Method 2: Graphing Calculator

Most calculators that give the least squares line will also give the coefficient of correlation. To do this on the TI-83/84 Plus, press the second function CATALOG and go down the list to the entry DiagnosticOn. Press ENTER at that point, then press STAT, CALC, and choose item 4 to get the display in Figure 22. The result is the same as we got by hand. The command DiagnosticOn need only be entered once, and the coefficient of correlation will always appear in the future.

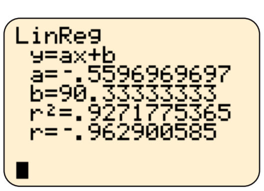


FIGURE 22

*For example, see *Introductory Statistics,* 6th edition, by Neil A. Weiss, Boston, Mass.: Addison-Wesley, 2002.

ISBN: 0-536-10718-1

Method 3: Spreadsheet Many computer spreadsheet programs have a built-in command to find the coefficient of correlation. For example, in Excel, use the command "= CORREL(A1:A10,B1:B10)" to find the correlation of the 10 data points stored in columns A and B. For more details, see *The Spreadsheet Manual* that is available with this text.

EXAMPLE 4 *Airline Passengers*

The following table shows the number of airline passengers in the United States (in millions) from 1993 to 2002.*

Year	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Passengers	488.5	528.8	547.8	581.2	594.7	612.9	636.0	666.2	622.1	611.9

Find the correlation coefficient, as well as the line that best fits the data.

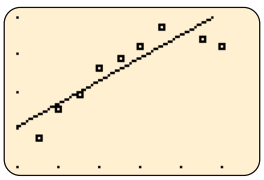


FIGURE 23

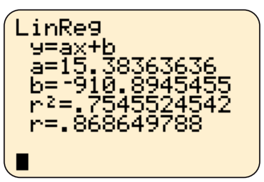


FIGURE 24

Solution A scatterplot of the data, along with the graph of the least squares line, is shown in Figure 23. Notice that the data appear to be linear for 1993 to 2000, but the last two points suggest that the number of passengers is no longer growing linearly. The most likely explanation for this downward spiral is the terrorist attacks on September 11, 2001. Figure 24 shows the result of the LinReg command on the TI-83/84 Plus. Although the correlation coefficient continues to reflect a linear trend, the scatterplot indicates that a significant event has occurred that has apparently stopped the increase, and in fact, has started a decline in the number of airline passengers. If this same analysis had been done in 2000 (that is, if we removed the last two data points), we would get $r = .99$ and $Y = 23.675x - 1702.625$. (Verify this on your calculator.) Therefore, the trend from 1993 to 2000 is very linear and many people would have been confident in predicting that there would have been about 712,225,000 airline passengers in 2002. Notice that the actual number is about 100 million fewer.

This example illustrates the hazards of using current trends to predict the future.

The final example is another illustration of why a plot of the data points is so important.

EXAMPLE 5 *Silicone Implants*

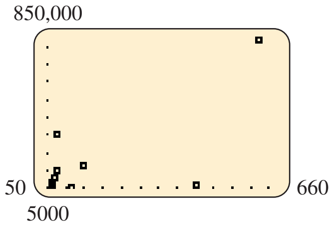
Silicones have long been used for fabricating medical devices on the presumption that they are biocompatible materials. This presumption is not entirely correct. Silicone prostheses, when implanted within the soft tissues of the breast, may evoke an inflammatory reaction. In response to silicone exposure, inflammatory mediator production was observed in experimental studies. After in vitro culture for 24 hours, the levels of four inflammatory mediators from ten patients with silicone breast implants were as shown in the following table.†

*Annual Traffic and Company, Air Transport Association, June 4, 2003, http://www.airlines.org/public/industry/display1.asp?nid=1032.

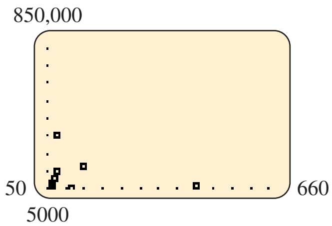
†Mena et al., "Inflammatory Intermediates Produced by Tissues Encasing Silicone Breast Prostheses," *Journal of Investigative Surgery,* Vol. 8, 1995, p. 33. Copyright © 1995. Reproduced by permission of Taylor & Francis, Inc., http://www.routledge-ny.com.

ISBN: 0-536-10718-1

Patient	IL-2	TNF-α	IL-6	PGE_2
1	48	ND	231	68
2	219	78	308,287	2710
3	109	65	33,291	1804
4	2179	149	124,550	8053
5	219	451	17,075	7371
6	54	64	22,955	3418
7	6	79	95,102	9768
8	10	115	5649	441
9	42	618	840,585	9585
10	196	69	58,924	4536



(a)



(b)

FIGURE 25

A graphing calculator was used to plot the last nine points in the table for TNF-$\alpha(x)$ and IL-6(y). (There is no data point for patient 1, since there is no TNF-α level.) The graph is shown in Figure 25(a). The correlation for these two mediators is .6936. Notice that one data point is way off by itself. As shown in Example 4, sometimes, by removing such a point from the graph, we can achieve a higher correlation.* Figure 25(b) shows the scatterplot with the remaining eight points. However, the correlation is now $-.2493$![†] With the ninth point, the points were closer to a line with a positive slope. Without the point, there is little linear correlation and it has become negative.

1.3 EXERCISES

1. Suppose a positive linear correlation is found between two quantities. Does this mean that one of the quantities increasing causes the other to increase? If not, what does it mean?

2. Given a set of points, the least squares line formed by letting x be the independent variable will not necessarily be the same as the least squares line formed by letting y be the independent variable. Give an example to show why this is true.

The following problem is reprinted from the November 1989 Actuarial Examination on Applied Statistical Methods.[‡]

3. You are given

X	6.8	7.0	7.1	7.2	7.4
Y	.8	1.2	.9	.9	1.5

Determine r^2, the coefficient of determination for the regression of Y on X. (*Note:* The coefficient of determination is defined as the square of the coefficient of correlation.)

a. .3 **b.** .4 **c.** .5 **d.** .6 **e.** .7

*Before discarding a point, we should investigate the reason it is an outlier.

[†]The observation that removing one point changes the correlation from positive to negative was made by Patrick Fleury, *Chance News* (an electronic newsletter), Vol. 4, No. 16, Dec. 1995.

[‡]"November 1989 Course 120 Examination Applied Statistical Methods" of the *Education and Examination Committee of The Society of Actuaries.* Reprinted by permission of The Society of Actuaries.

ISBN: 0-536-10718-1

4. Follow the steps outlined in this section to solve the least squares line equations

$$nb + (\Sigma x)m = \Sigma y$$
$$(\Sigma x)b + (\Sigma x^2)m = \Sigma xy$$

for m and b to get

$$m = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$
$$b = \frac{\Sigma y - m(\Sigma x)}{n}.$$

Applications

BUSINESS AND ECONOMICS

5. *Recreation Spending* The U.S. Department of Commerce, Bureau of Economic Analysis, has reported the total U.S. expenditures on recreational goods (hobbies, music, sports, spectator admissions, etc.). From 1994 to 2001, expenditures have grown at an approximately linear rate. The results of the report, in which x represents the years since 1900 and y represents the total expenditures (in billions of dollars), provide the following summations.*

$$n = 8 \qquad \Sigma x^2 = 76{,}092$$
$$\Sigma x = 780 \qquad \Sigma xy = 374{,}850.8$$
$$\Sigma y = 3830.7 \qquad \Sigma y^2 = 1{,}878{,}286.33$$

a. Find an equation for the least squares line.

b. Predict the recreational expenditures in 2005.

c. If this growth continues linearly, when will recreational expenditures reach 750 billion dollars?

d. Find and interpret the coefficient of correlation.

6. *Decrease in Banks* The number of banks in the United States has dropped about 30% since 1992. The following data are from a survey in which x represents the years since 1900 and y corresponded to the number of banks, in thousands, in the United States.†

$$n = 10 \qquad \Sigma x^2 = 93{,}205$$
$$\Sigma x = 965 \qquad \Sigma xy = 9165.1$$
$$\Sigma y = 95.3 \qquad \Sigma y^2 = 920.47$$

a. Find an equation of the least squares line.

b. If the trend continues, how many banks will there be in 2004?

c. Find and interpret the coefficient of correlation.

7. *Air Fares* In January 2000, American Airlines ran an ad in *The New York Times* advertising one-way air fares from New York to various cities.‡ Fourteen of the cities are listed below, with the distances from New York to the cities added.

a. Plot the data. Do the data points lie in a linear pattern?

b. Find the correlation coefficient. Combining this with your answer to part a, does the cost of a ticket tend to go up with the distance flown?

c. Find the equation of the least squares line, and use it to find the approximate marginal cost per mile to fly.

d. For similar data in an October 1993 ad, the equation of the least squares line was $Y = 91.9 + .0313x$. Use this information and your answer to part b to compare the cost of flying American Airlines for these two time periods.

City	Distance (x) (miles)	Price (y) (dollars)
Boston	206	109
Chicago	802	124
Denver	1771	154
Kansas City	1198	144
Little Rock	1238	144
Los Angeles	2786	179
Minneapolis	1207	144
Nashville	892	144
Phoenix	2411	179
Portland	2885	179
Reno	2705	179
St. Louis	948	144
San Diego	2762	179
Seattle	2815	179

**The World Almanac and Book of Facts 2003,* p. 109.
†FDIC, Historical Statistics on Banking, http://www2.fdic.gov/hsob/hsobRpt.asp.
‡*The New York Times,* Jan. 7, 2000, p. A9.

ISBN: 0-536-10718-1

Year (x)	95	96	97	98	99	00	01	02
Debt (y)	5832	6487	6900	7188	7564	8123	8367	8562

8. ***Consumer Debt*** Credit card debt has risen steadily over the years. The table above gives the average U.S. credit card debt (in dollars) per household. Years are represented as the number of years since 1900. (The table above includes all credit cards and U.S. households with at least one credit card.)[*]

a. Plot the data. Does the graph show a linear pattern?

b. Find the equation of the least squares line and graph it on the same axes. Does the line appear to be a good fit?

c. Find and interpret the coefficient of correlation.

d. If this linear trend continues, when will household debt reach $10,000?

9. ***Used Car Sales*** As cars are becoming more expensive, used car sales have increased at a faster rate since 1984 than new car sales.[†] Sales in millions from 1984 to 1996 are given in the table below.

Year	Sales	Year	Sales
84	12.3	91	12.3
85	13.2	92	12.8
86	13.6	93	13.9
87	13.2	94	15.0
88	14.5	95	14.7
89	14.5	96	14.6
90	13.8		

a. Find the equation of the least squares line and the coefficient of correlation.

b. Find the equation of the least squares line using only the data for every other year starting with 1985, 1987, and so on. Find the coefficient of correlation.

c. Compare your answers for parts a and b. What do you find? Why do you think this happens?

10. ***Medical School Admissions*** According to the American Association of Medical Colleges, the number of applications to medical schools in the United States began to decrease since 1996 as indicated in the following table.[‡] Years are represented as the number of years since 1900 and applications are given in thousands.

Year (x)	94	95	96	97	98	99	00
Applications (y)	45.4	46.6	47.0	43.0	41.0	38.5	37.1

a. Plot the data. Do the data points lie in a linear pattern?

b. Determine the least squares line for this data and graph it on the same coordinate axes. Does the line fit the data reasonably well?

c. Find the coefficient of correlation. Does it agree with your estimate of the fit in part b?

d. Explain why the coefficient of correlation is close to 1, even though some of the data points do not appear to be linear.

11. ***Bird Eggs*** The average length and width of various bird eggs are given in the following table.[§]

Bird Name	Width (cm)	Length (cm)
Canada goose	5.8	8.6
Robin	1.5	1.9
Turtledove	2.3	3.1
Hummingbird	1.0	1.0
Raven	3.3	5.0

a. Plot the points, putting the length on the y-axis and the width on the x-axis. Do the data appear to be linear?

b. Find the least squares line, and plot it on the same graph as the data.

*MSN Money, *Is the Debt Binge Over?* http://money.msn.com/articles/smartbuy/basics/9526.asp.
†*The New York Times,* March 3, 1996.
‡Gabriel, B., "Medical School Applicant Pool Still Runs Deep," *New Room Reporter*, American Association of Medical Colleges, Vol. 6, No. 4, May 2003.
§www.nctm.org/wlme/wlme6/five.htm.

ISBN: 0-536-10718-1

c. Suppose there are birds with eggs even smaller than those of hummingbirds. Would the equation found in part b continue to make sense for all positive widths, no matter how small? Explain.

d. Find the coefficient of correlation.

12. ***Crickets Chirping*** Biologists have observed a linear relationship between the temperature and the frequency with which a cricket chirps. The following data were measured for the striped ground cricket.*

Temperature °F (x)	Chirps per Second (y)
88.6	20.0
71.6	16.0
93.3	19.8
84.3	18.4
80.6	17.1
75.2	15.5
69.7	14.7
82.0	17.1
69.4	15.4
83.3	16.2
79.6	15.0
82.6	17.2
80.6	16.0
83.5	17.0
76.3	14.4

a. Find the equation for the least squares line for the data.

b. Use the results of part a to determine how many chirps per second you would expect to hear from the striped ground cricket if the temperature were 73°F.

c. Use the results of part a to determine what the temperature is when the striped ground crickets are chirping at a rate of 18 times per sec.

d. Find the coefficient of correlation.

SOCIAL SCIENCES

13. ***Educational Expenditures*** A 2000 report issued by the U.S. Department of Education listed the expenditure per pupil and the average mathematics proficiency in grade 8 for 39 states and the District of Columbia. Letting x equal the expenditure per pupil (ranging from \$4378 in Utah to \$10,107 in Washington, D.C.) and y equal the average mathematics proficiency score (ranging from 234 in Washington, D.C., to 288 in Minnesota), the data can be summarized as follows:†

$$\sum x = 266{,}947 \qquad \sum x^2 = 1{,}865{,}325{,}667$$
$$\sum y = 10{,}922 \qquad \sum y^2 = 2{,}986{,}764$$
$$\sum xy = 72{,}987{,}414 \qquad n = 40$$

Compute the coefficient of correlation for the given data. Does there appear to be a trend in the amount of money spent per pupil and the proficiency of eighth graders in mathematics?

14. ***Poverty Levels*** The following table lists how poverty level income cutoffs (in dollars) for a family of four have changed over time.‡

Year	Income
1970	3968
1975	5500
1980	8414
1985	10,989
1990	13,359
1995	15,569
2000	17,603

Let x be the year, with $x = 0$ corresponding to 1970, and y be the income in thousands of dollars. (*Note:* $\sum x = 105$, $\sum x^2 = 2275$, $\sum y = 75.402$, $\sum y^2 = 968.270792$, $\sum xy = 1460.97$.)

a. Plot the data. Do the data appear to lie along a straight line?

b. Calculate the coefficient of correlation. Does your result agree with your answer to part a?

c. Find the equation of the least squares line.

d. Use your answer from part c to predict the poverty level in the year 2015.

15. ***SAT Scores*** At Hofstra University, all students take the math SAT before entrance, and most students take a mathematics placement test before registration. Recently, one professor collected the data on the next page for 19 students in his Finite Mathematics class:

a. Find an equation for the least squares line. Let x be the math SAT and y be the placement test score.

*Pierce, George W., Data from *The Songs of Insects,* Cambridge, Mass., Harvard University Press, Copyright © 1948 by the President and Fellows of Harvard College.

†National Center for Educational Statistics.

‡U.S. Census Bureau, *Current Population Reports*.

ISBN: 0-536-10718-1

Math SAT	Placement Test	Math SAT	Placement Test	Math SAT	Placement Test
540	20	580	8	440	10
510	16	680	15	520	11
490	10	560	8	620	11
560	8	560	13	680	8
470	12	500	14	550	8
600	11	470	10	620	7
540	10				

b. Use your answer from part a to predict the mathematics placement test score for a student with a math SAT score of 420.

c. Use your answer from part a to predict the mathematics placement test score for a student with a math SAT score of 620.

d. Calculate the coefficient of correlation.

e. Based on your answer to part d, what can you conclude about the relationship between a student's math SAT and mathematics placement test score?

PHYSICAL SCIENCES

16. ***Air Conditioning*** While shopping for an air conditioner, Adam Bryer consulted the following table giving a machine's BTUs and the square footage (ft^2) that it would cool.

ft^2 (x)	BTUs (y)
150	5000
175	5500
215	6000
250	6500
280	7000
310	7500
350	8000
370	8500
420	9000
450	9500

a. Find the equation for the least squares line for the data.

b. To check the fit of the data to the line, use the results from part a to find the BTUs required to cool a room of 150 ft^2, 280 ft^2, and 420 ft^2. How well does the actual data agree with the predicted values?

c. Suppose Adam's room measures 230 ft^2. Use the results from part a to decide how many BTUs it requires. If air conditioners are available only with the BTU choices in the table, which would Adam choose?

d. Why do you think the table gives ft^2 instead of ft^3, which would give the volume of the room?

17. ***Length of a Pendulum*** Grandfather clocks use pendulums to keep accurate time. The relationship between the length of a pendulum L and the time T for one complete oscillation can be determined from the data in the table.*

L (ft)	T (sec)
1.0	1.11
1.5	1.36
2.0	1.57
2.5	1.76
3.0	1.92
3.5	2.08
4.0	2.22

a. Plot the data from the table with L as the horizontal axis and T as the vertical axis.

b. Find the least squares line equation and graph it simultaneously, if possible, with the data points. Does it seem to fit the data?

c. Find the coefficient of correlation and interpret it. Does it confirm your answer to part b?†

GENERAL INTEREST

18. ***Athletic Records*** The table on the next page shows the men's and women's world records (in seconds) in the 800-m run.‡

a. Find the equation for the least squares line for the men's record (y) in terms of the year (x). Use 5 for 1905, 15 for 1915, and so on.

*Data provided by Gary Rockswold, Mankato State University, Minnesota.

†The actual relationship is $L = .81T^2$, which is not a linear equation. This illustrates that even if the relationship is not linear, a line can give a good approximation.

‡Whipp, Brian J. and Susan A. Ward, "Will Women Soon Outrun Men?" *Nature,* Vol. 355, Jan. 2, 1992, p. 25. The data are from Peter Matthews, *Track and Field Athletics: The Records,* Guinness, 1986, pp. 11, 44, and from Robert W. Schutz and Yuanlong Liu in *Statistics in Sport,* edited by Jay Bennett, Arnold, 1998, p. 189.

ISBN: 0-536-10718-1

Year	Men's Record	Women's Record
1905	113.4	—
1915	111.9	—
1925	111.9	144.0
1935	109.7	135.6
1945	106.6	132.0
1955	105.7	125.0
1965	104.3	118.0
1975	103.7	117.48
1985	101.73	113.28
1995	101.11	113.28

b. Find the equation for the least squares line for the women's record.

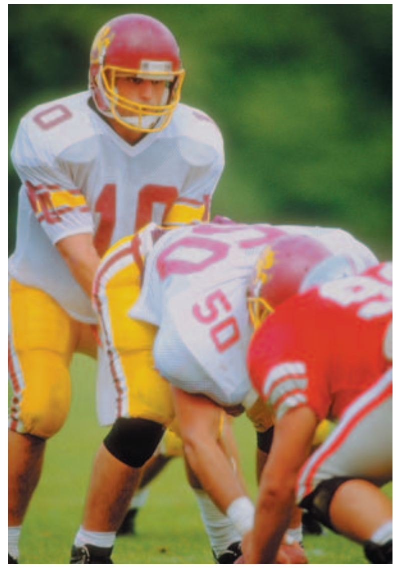
c. Suppose the men's and women's records continue to improve as predicted by the equations found in parts a and b. In what year will the women's record catch up with the men's record? Do you believe that will happen? Why or why not?

d. Calculate the coefficient of correlation for both the men's and the women's record. What do these numbers tell you?

19. ***Football*** The following data give the expected points for a football team with first down and 10 yards to go from various points on the field.* (*Note:* $\sum x = 500$, $\sum x^2 = 33{,}250$, $\sum y = 20.668$, $\sum y^2 = 91.927042$, $\sum xy = 399.16$.)

a. Calculate the coefficient of correlation. Does there appear to be a linear correlation?

Yards from Goal (x)	Expected Points (y)
5	6.041
15	4.572
25	3.681
35	3.167
45	2.392
55	1.538
65	.923
75	.236
85	−.637
95	−1.245



b. Find the equation of the least squares line.

c. Use your answer from part a to predict the expected points when a team is at the 50-yd line.

20. ***Baseball*** Some baseball fans are concerned about the recent increase in time to complete the game. The following table shows the average time (in hours and minutes) to complete baseball games in recent years.†

Year	Average Completion Time	Year	Average Completion Time
1981	2:33	1989	2:46
1982	2:34	1990	2:48
1983	2:36	1991	2:49
1984	2:35	1992	2:49
1985	2:40	1993	2:48
1986	2:44	1994	2:54
1987	2:48	1995	2:57
1988	2:45		

Let x be the number of years since 1980, and let y be the number of minutes beyond 2 hours. (*Note:* $\sum x = 120$, $\sum x^2 = 1240$, $\sum y = 666$, $\sum xy = 5765$.)

a. Find the equation of the least squares line.

b. If the trend in the data continues, in what year will the average completion time be 3 hours and 15 minutes?

*Carter, Virgil and Robert E. Machol, *Operations Research,* Vol. 19, 1971, pp. 541–545.
†*The New York Times,* May 30, 1995, p. B9.

ISBN: 0-536-10718-1

21. ***Running*** If you think a marathon is a long race, consider the Hardrock 100, a 101.7-mile running race held in southwestern Colorado. The following table lists the times that the 2000 winner, Kirk Apt, arrived at various mileage points along the way.*

a. What was Apt's average speed?

b. Graph the data, plotting time on the x-axis and distance on the y-axis. You will need to convert the time from hours and minutes into hours. Do the data appear to lie approximately on a straight line?

c. Find the equation for the least squares line, fitting distance as a linear function of time.

d. Calculate the coefficient of correlation. Does it indicate a good fit of the least squares line to the data?

e. Based on your answer to part d, what is a good value for Apt's average speed? Compare this with your answer to part a. Which answer do you think is better? Explain your reasoning.

Miles	Time (hr:min)
0	0
9.6	2:14
16.5	4:08
21.6	6:10
31.6	7:10
42.4	10:51
49.8	12:42
58.0	14:20
65.2	16:30
68.4	18:02
73.7	19:25
83.1	23:07
89.6	26:09
95.8	28:18
101.7	29:35

CHAPTER SUMMARY

In this chapter we have seen how to find the equation of a line, given a point and the slope or given two points. We have also seen how to express the result as a linear function. Equations of lines have a broad range of applications, as demonstrated in this chapter. They are used through the rest of this book, so fluency in their use is important. The method of least squares shows how mathematical models, such as many of those used throughout this book, are derived.

KEY TERMS

To understand the concepts presented in this chapter, you should know the meaning and use of the following terms. For easy reference, the section in the chapter where a word (or expression) was first used is provided.

mathematical model

1.1 **ordered pair**
Cartesian coordinate system
axes
origin
coordinates
quadrants
graph
intercepts
slope
linear equation
slope-intercept form
proportional
point-slope form
parallel
perpendicular
scatterplot

1.2 **linear function**
independent variable
dependent variable
supply curve
demand curve
equilibrium price
equilibrium quantity
fixed cost
marginal cost
cost function
break-even quantity
break-even point

1.3 **least squares line**
summation notation
coefficient of correlation

*Hardrock Hundred Mile Endurance Run, 2000 Hardrock Results Spreadsheet, http://www.run100s.com/HR/.

ISBN: 0-536-10718-1

CHAPTER 1 REVIEW EXERCISES

1. What is marginal cost? Fixed cost?

2. What six quantities are needed to compute a coefficient of correlation?

Find the slope for each line that has a slope.

3. Through $(-2, 5)$ and $(4, 7)$

4. Through $(4, -1)$ and $(3, -3)$

5. Through the origin and $(11, -2)$

6. Through the origin and $(0, 7)$

7. $2x + 3y = 15$

8. $4x - y = 7$

9. $y + 4 = 9$

10. $3y - 1 = 14$

11. $y = -3x$

12. $x = 5y$

Find an equation in the form $y = mx + b$ (where possible) for each line.

13. Through $(5, -1)$; slope $= 2/3$

14. Through $(8, 0)$; slope $= -1/4$

15. Through $(5, -2)$ and $(1, 3)$

16. Through $(2, -3)$ and $(-3, 4)$

17. Through $(-1, 4)$; undefined slope

18. Through $(-2, 5)$; slope $= 0$

19. Through $(2, -1)$, parallel to $3x - y = 1$

20. Through $(0, 5)$, perpendicular to $8x + 5y = 3$

21. Through $(2, -10)$, perpendicular to a line with undefined slope

22. Through $(3, -5)$, parallel to $y = 4$

23. Through $(-7, 4)$, perpendicular to $y = 8$

Graph each linear equation defined as follows.

24. $y = 4x + 3$

25. $y = 6 - 2x$

26. $3x - 5y = 15$

27. $2x + 7y = 14$

28. $x + 2 = 0$

29. $y = 1$

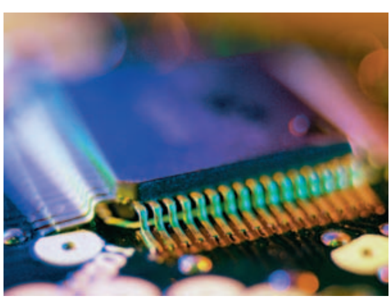
30. $y = 2x$

31. $x + 3y = 0$

Applications

BUSINESS AND ECONOMICS

32. ***Profit*** To manufacture x thousand computer chips requires fixed expenditures of \$352 plus \$42 per thousand chips. Receipts from the sale of x thousand chips amount to \$130 per thousand.



a. Write an expression for expenditures.

b. Write an expression for receipts.

c. For profit to be made, receipts must be greater than expenditures. How many chips must be sold to produce a profit?

33. ***Supply and Demand*** The supply and demand for crabmeat in a local fish store are related by the equations

$$\text{Supply: } p = S(q) = 6q + 3$$

and

$$\text{Demand: } p = D(q) = 19 - 2q,$$

where p represents the price in dollars per pound and q represents the quantity of crabmeat in pounds per day. Find the supply and demand at each of the following prices.

a. \$10 **b.** \$15 **c.** \$18

d. Graph both the supply and the demand functions on the same axes.

e. Find the equilibrium price.

f. Find the equilibrium quantity.

ISBN: 0-536-10718-1

34. ***Supply*** For a new diet pill, 60 pills will be supplied at a price of \$40, while 100 pills will be supplied at a price of \$60. Write a linear supply function for this product.

35. ***Demand*** The demand for the diet pills in Exercise 34 is 50 pills at a price of \$47.50 and 80 pills at a price of \$32.50. Determine a linear demand function for these pills.

36. ***Supply and Demand*** Find the equilibrium price and quantity for the diet pills in Exercises 34 and 35.

Cost *Find a linear cost function in Exercises 37–40.*

37. Eight units cost \$300; fixed cost is \$60.

38. Fixed cost is \$2000; 36 units cost \$8480.

39. Twelve units cost \$445; 50 units cost \$1585.

40. Thirty units cost \$1500; 120 units cost \$5640.

41. ***Break-Even Analysis*** The cost of producing x cartons of CDs is $C(x)$ dollars, where $C(x) = 200x + 1000$. The CDs sell for \$400 per carton.

   a. Find the break-even quantity.

   b. What revenue will the company receive if it sells just that number of cartons?

42. ***Break-Even Analysis*** The cost function for flavored coffee at an upscale coffeehouse is given in dollars by $C(x) = 3x + 160$, where x is in pounds. The coffee sells for \$7 per pound.

   a. Find the break-even quantity.

   b. What will the revenue be at that point?

43. ***U.S. Imports from China*** The U.S. is China's largest export market. Imports from China have grown from about 19 billion dollars in 1991 to 102 billion dollars in 2001.* This growth has been approximately linear. Use the given data pairs to write a linear equation that describes this growth in imports over the years. Let $x = 91$ represent 1991 and $x = 101$ represent 2001.

44. ***U.S. Exports to China*** U.S. exports to China have grown (although at a slower rate than imports) since 1991. In 1991, about 10 billion dollars of goods were exported to China. By 2001, this amount had grown to 19 billion dollars.* Write a linear equation describing the number of exports each year, with $x = 91$ representing 1991 and $x = 101$ representing 2001.

45. ***Median Income*** The U.S. Census Bureau reported that the median income for all U.S. households in 2000 was \$42,148. In 1993, the median income (in 2000 dollars) was \$36,746.† The median income is approximately linear and is a function of time. Find a formula for the median income, I, as a function of the year x, where x is the number of years since 1900.

46. ***New Car Cost*** The average new car cost for the years from 1975 to 2000 is given in the table where x is the number of years since 1900.‡

Year (x)	75	80	85	90	95	00
Cost (y)	6000	7500	12,000	16,000	20,400	24,900

   a. Find an equation for the least squares line.

   b. Use your equation from part a to predict the average cost of a new car in the year 2005 ($x = 105$).

   c. Find and interpret the coefficient of correlation. Does it indicate that the line is a good fit for the data?

   d. Plot the data. Does the scatterplot suggest the trend might not be linear?

LIFE SCIENCES

47. ***World Health*** In general, people tend to live longer in countries that have a greater supply of food. Listed below is the 1997 daily calorie supply and 2000 life expectancy at birth for 10 randomly selected countries.§

Country	**Calories (x)**	**Life Expectancy (y)**
Afghanistan	1523	43
Belize	2862	74
Cambodia	1974	56
France	3551	79
India	2415	64
Mexico	3137	73
New Zealand	3405	78
Peru	2310	70
Sweden	3160	80
United States	3642	78

   a. Find the coefficient of correlation. Do the data seem to fit a straight line?

*International Trade Administration, Trade and Economy: Data and Analysis, Tables 55 and 56.
†U.S. Census Bureau, Current Population Survey, March 1994, 2000, 2001.
‡*Chicago Tribune,* Feb. 4, 1996, Sec. 5, p. 4 and NADA Industry Analysis Division 2002.
§*The New York Times 2003 Almanac,* pp. 479–481.

ISBN: 0-536-10718-1

b. Draw a scatterplot of the data. Combining this with your results from part a, do the data seem to fit a straight line?

c. Find the equation for the least squares line.

d. Use your answer from part c to predict the life expectancy in the United Kingdom, which has a daily calorie supply of 3237. Compare your answer with the actual value of 78 years.

e. Briefly explain why countries with a higher daily calorie supply might tend to have a longer life expectancy.

f. (For the ambitious!) Find the coefficient of correlation and least squares line using the data for a larger sample of countries, as found in an almanac or other reference. Is the result in general agreement with the previous results?

48. ***Blood Sugar and Cholesterol Levels*** The following data show the connection between blood sugar levels and cholesterol levels for 8 different patients.

Patient	Blood Sugar Level (x)	Cholesterol Level (y)
1	130	170
2	138	160
3	142	173
4	159	181
5	165	201
6	200	192
7	210	240
8	250	290

For the data given in the preceding table, $\sum x = 1394$, $\sum y = 1607$, $\sum xy = 291{,}990$, $\sum x^2 = 255{,}214$, and $\sum y^2 = 336{,}155$.

a. Find the equation of the least squares line, $Y = mx + b$.

b. Predict the cholesterol level for a person whose blood sugar level is 190.

c. Find r.

SOCIAL SCIENCES

49. ***Red Meat Consumption*** The per capita consumption of red meat in the United States decreased from 129.5 lb in 1969 to 117.2 pounds in 1999.* Assume a linear function describes the decrease. Write a linear equation defining the function. Let x represent the number of years since 1900 and y represent the number of pounds of red meat consumed.

50. ***Marital Status*** More people are staying single longer in the United States. In 1990, the number of never-married adults, age 15 and over, was 52.6 million. By 2000, it was 59.9 million.[†] Assume the data increase linearly, and write an equation that defines a linear function for this data. Let x represent the number of years since 1900.

51. ***Governors' Salaries*** In general, the larger a state's population, the more its governor earns. Listed below are the estimated 2001 populations (in millions) and the salary of the governor (in thousands of dollars) for 8 randomly selected states.[‡]

a. Find the coefficient of correlation. Do the data seem to fit a straight line?

b. Draw a scatterplot of the data. Compare this with your answer from part a.

c. Find the equation for the least squares line.

d. Based on your answer to part c, how much does a governor's salary increase, on average, for each additional million in population?

e. Use your answer from part c to predict the governor's salary in your state. Based on your answers from parts a and b, would this prediction be very accurate? Compare with the actual salary, as listed in an almanac or other reference.

f. (For the ambitious!) Find the coefficient of correlation and least squares line using the data for all 50 states, as found in an almanac or other reference. Is the result in general agreement with the previous results?

State	AZ	DE	MD	MA	NY	PA	TN	WY
Population (x)	5.31	.80	5.38	6.38	19.01	12.29	5.74	.49
Governor's Salary (y)	95	114	120	135	179	142	85	95

**The World Almanac and Book of Facts 2003,* p. 100.
[†]U.S. Census Bureau, http://factfinder.census.gov/servlet/DTTable.
[‡]*The World Almanac and Book of Facts 2003,* pp. 364, 368.

ISBN: 0-536-10718-1

EXTENDED APPLICATION: Using Extrapolation to Predict Life Expectancy

One reason for developing a mathematical model is to make predictions. If your model is a least squares line, you can predict the y value corresponding to some new x by substituting this x into an equation of the form $Y = mx + b$. (We use a capital Y to remind us that we're getting a predicted value rather than an actual data value.) Data analysts distinguish two very different kinds of prediction, *interpolation* and *extrapolation*. An interpolation uses a new x inside the x range of your original data. For example, if you have inflation data at five-year intervals from 1950 to 2000, estimating the rate of inflation in 1957 is an interpolation problem. But if you use the same data to estimate what the inflation rate was in 1920, or what it will be in 2020, you are extrapolating.

In general, interpolation is much safer than extrapolation, because data that are approximately linear over a short interval may be nonlinear over a larger interval. One way to detect nonlinearity is to look at *residuals,* which are the differences between the actual data values and the values predicted by the line of best fit. Here is a simple example:

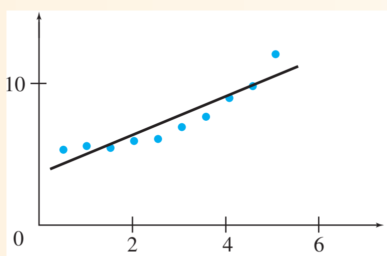


FIGURE 26

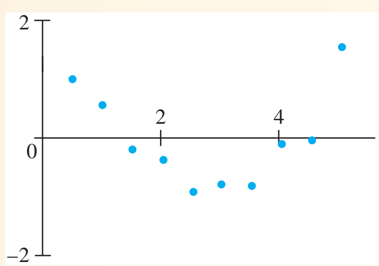


FIGURE 27

The regression equation for the linear fit on the top is $Y = 3.431 + 1.334x$. Since the r value for this regression line is .93, our linear model fits the data very well. But we might notice that the predictions are a bit low at the ends and high in the middle. We can get a better look at this pattern by plotting the residuals. To find them, we put each value of the independent variable into the regression equation, calculate the predicted value Y, and subtract it from the actual y value. The residual plot is below the linear fit graph, with the vertical axis rescaled to exaggerate the pattern. The residuals indicate that our data has a nonlinear, U-shaped component that is not captured by the linear fit. Extrapolating from this data set is probably not a good idea; our linear prediction for the value of y when x is 10 may be much too low.

Exercises

*The following table gives the life expectancy at birth of females born in the United States in various years from 1950 to 2000.**

Year of Birth	Life Expectancy (years)
1950	71.3
1960	73.1
1970	74.7
1980	77.4
1985	78.2
1990	78.8
1995	78.9
2000	79.5

1. Find an equation for the least squares line for this data, using year of birth as the independent variable.
2. Use your regression equation to guess a value for the life expectancy of females born in 1900.
3. Compare your answer with the actual life expectancy for females born in 1900, which was 48.3 years. Are you surprised?
4. Find the life expectancy predicted by your regression equation for each year in the table, and subtract it from the actual value in the second column. This gives you a table of residuals. Plot your residuals as points on a graph.
5. Now look at the residuals as a fresh data set and see if you can sketch the graph of a smooth function that fits the residuals well. How easy do you think it will be to predict the life expectancy at birth of females born in 2010?

**The World Almanac and Book of Facts 2003,* p. 75.

ISBN: 0-536-10718-1

6. What will happen if you try linear regression on the *residuals?* If you're not sure, use your calculator or software to find the regression equation for the residuals. Why does this result make sense?
7. Since most of the females born in 1985 are still alive, how did the Public Health Service come up with a life expectancy of 78.2 years for these women?

Directions for Group Project

Assume that you and your group (3–5 students) are preparing a report for a local health agency that is interested in using linear regression to predict life expectancy. Using the questions above as a guide, write a report that addresses the spirit of each question and any issues related to that question. The report should be mathematically sound, grammatically correct, and professionally crafted. Provide recommendations as to whether the health agency should proceed with the linear equation or whether it should seek other means of making such predictions.

ISBN: 0-536-10718-1