



13

Normal Distributions

CASE STUDY Bar graphs and histograms are definitely old technology. Using bars to display data goes back to William Playfair (1759–1823), an English economist who was an early pioneer of data graphics. Histograms require that we choose classes, and their appearance can change with different choices. Surely modern software offers a better way to picture distributions?

Software can replace the separate bars of a histogram with a smooth curve that represents the overall shape of a distribution. Look at Figure 13.1. The data are the number of text messages reported as being sent on a particular day in 2013 by a random sample of 447 high school seniors. The curve is generated from statistical software as a way of replacing the histogram. The software doesn't start from the histogram—it starts with the actual observations and cleverly draws a curve to describe their distribution.

In Figure 13.1, the software has caught the overall shape and shows the ripples in the long right tail more effectively than does the histogram. It struggles a bit with the peak: it has extended the curve beyond zero in an attempt to smooth out the sharp peak. In Figure 13.2, we apply the same software to a set of data with a more regularly shaped distribution. These are the body temperatures from a sample of 130 healthy adults. The software draws a curve that shows a distinctive symmetric, single-peaked bell shape.

For the irregular distribution in Figure 13.1, we can't do better. In the case of the very symmetric sampling data in Figure 13.2, however, there is another way to get a smooth curve. The distribution can be described by a specific kind of smooth curve called a *Normal curve*. Figure 13.3 shows the Normal curve for these data. The curve looks a lot like the one in Figure 13.2, but a close look shows that it is smoother. The Normal curve is much easier to work with and does not require clever software.

In this chapter, we will learn that Normal curves have special properties that help us use them and think about them. By the end of this chapter, you will be able to use these properties to answer questions about the underlying distributions represented in Figures 13.2 and 13.3 that cannot easily be determined from the histograms.



Leonardo Patrizi/Getty Images

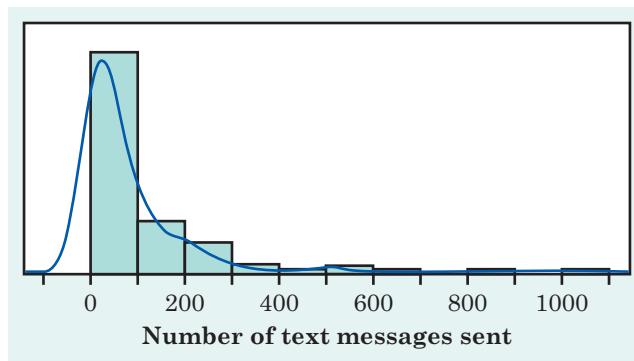


Figure 13.1 A histogram and a computer-drawn curve. Both picture the distribution of number of text messages reported as being sent on a particular day in 2013 by a random sample of 447 high school seniors. This distribution is skewed to the right. (This figure was created using the JMP software package.)

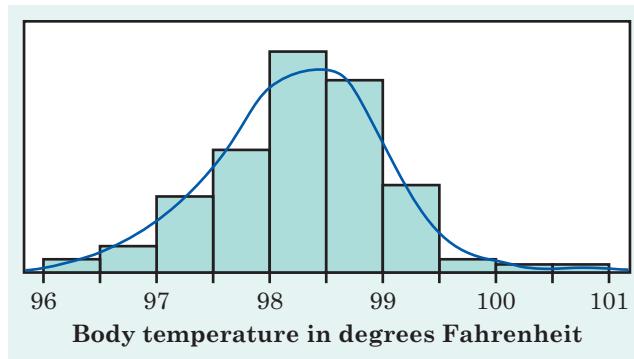


Figure 13.2 A histogram and a computer-drawn curve. Both picture the distribution of body temperatures from a sample of 130 healthy adults. This distribution is quite symmetric. (This figure was created using the JMP software package.)

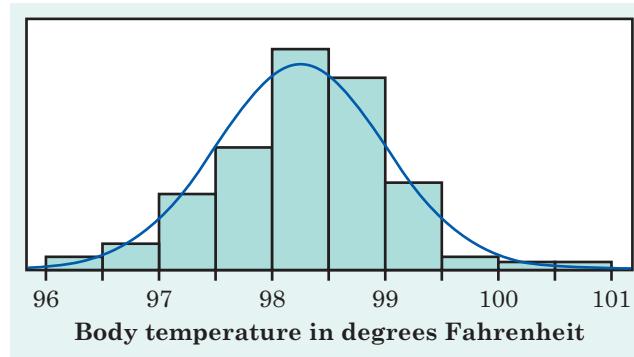


Figure 13.3 A perfectly symmetric Normal curve used to describe the distribution of body temperatures.

We now have a kit of graphical and numerical tools for describing distributions. What is more, we have a clear strategy for exploring data on a single quantitative variable:

1. Always plot your data: make a graph, usually a histogram or a stemplot.
2. Look for the overall pattern (shape, center, variability) and for striking deviations such as outliers.
3. Choose either the five-number summary or the mean and standard deviation to briefly describe center and variability in numbers.

Here is one more step to add to this strategy:

4. Sometimes, the overall pattern of a large number of observations is so regular that we can describe it by a smooth curve.

Density curves

Figures 13.1 and 13.2 show curves used in place of histograms to picture the overall shape of a distribution of data. You can think of drawing a curve through the tops of the bars in a histogram, smoothing out the irregular ups and downs of the bars. There are two important distinctions between histograms and these curves. First, most histograms show the *counts* of observations in each class by the heights of their bars and, therefore, by the areas of the bars. We set up curves to show the *proportion* of observations in any region by areas under the curve. To do that, we choose the scale so that the total area under the curve is exactly 1. We then have a **density curve**. Second, a histogram is a plot of data obtained from a sample. We use this histogram to understand the actual distribution of the population from which the sample was selected. The density curve is intended to reflect the idealized shape of the population distribution.

EXAMPLE 1 Using a density curve

Figure 13.4 copies Figure 13.3, showing the histogram and the Normal density curve that describe this data set of 130 body temperatures. What proportion of the temperatures are greater than or equal to 99 degrees Fahrenheit? From the actual 130 observations, we can count that exactly 19 are greater than or equal to 99°F. So the proportion is 19/130, or 0.146. Because 99 is one of the break points between the classes in the histogram, the area of the shaded bars in Figure 13.4(a) makes up 0.146 of the total area of all the bars.

Now concentrate on the density curve drawn through the histogram. The total area under this curve is 1, and the shaded area in Figure 13.4(b)

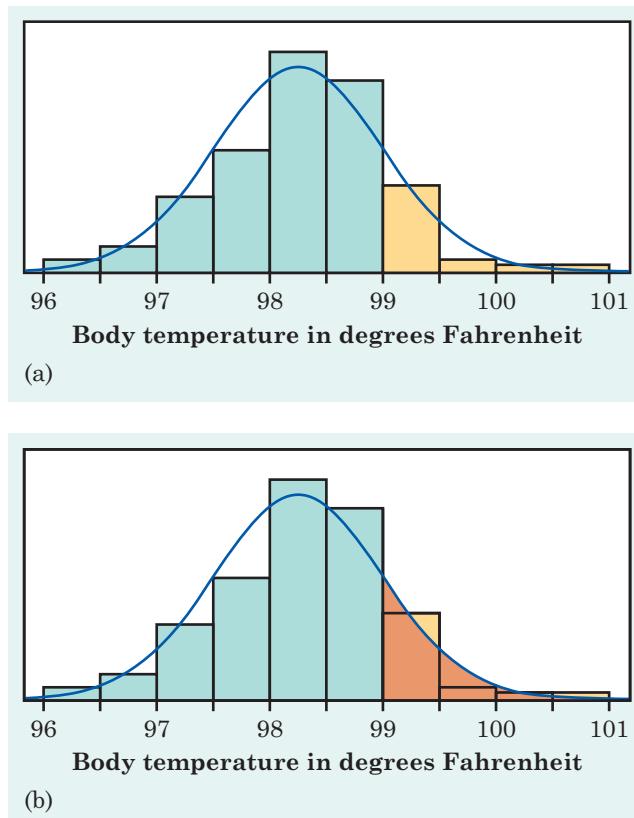


Figure 13.4 A histogram and a Normal density curve, Example 1. (a) The area of the shaded bars in the histogram represents observations greater than 99°F. These make up 19 of the 130 observations. (b) The shaded area under the Normal curve represents the proportion of observations greater than 99°F. This area is 0.1587.

represents the proportion of observations that are greater than or equal to 99°F. This area is 0.1587. You can see that the density curve is a quite good approximation—0.1587 is close to 0.146.

The area under the density curve in Example 1 is not exactly equal to the true proportion because the curve is an idealized picture of the distribution. For example, the curve is exactly symmetric, but the actual data are only approximately symmetric. Because density curves are smoothed-out idealized pictures of the overall shapes of distributions, they are most useful for describing large numbers of observations.

The center and variability of a density curve

Density curves help us better understand our measures of center and variability. The median and quartiles are easy. Areas under a density curve represent proportions of the total number of observations. The median is

the point with half the observations on either side. So, *the median of a density curve is the equal-areas point*, the point with half the area under the curve to its left and the remaining half of the area to its right. The quartiles divide the area under the curve into quarters. One-fourth of the area under the curve is to the left of the first quartile, and three-fourths of the area is to the left of the third quartile. You can roughly locate the median and quartiles of any density curve by eye by dividing the area under the curve into four equal parts.

Because density curves are idealized patterns, a symmetric density curve is exactly symmetric. The median of a symmetric density curve is, therefore, at its center. Figure 13.5(a) shows the median of a symmetric curve. We can roughly locate the equal-areas point on a skewed curve like that in Figure 13.5(b) by eye.

What about the mean? The mean of a set of observations is their arithmetic average. If we think of the observations as weights stacked on a seesaw, the mean is the point at which the seesaw would balance. This fact is also true of density curves. *The mean is the point at which the curve would balance if made of solid material.* Figure 13.6 illustrates this fact about the mean. A symmetric curve balances at its center because

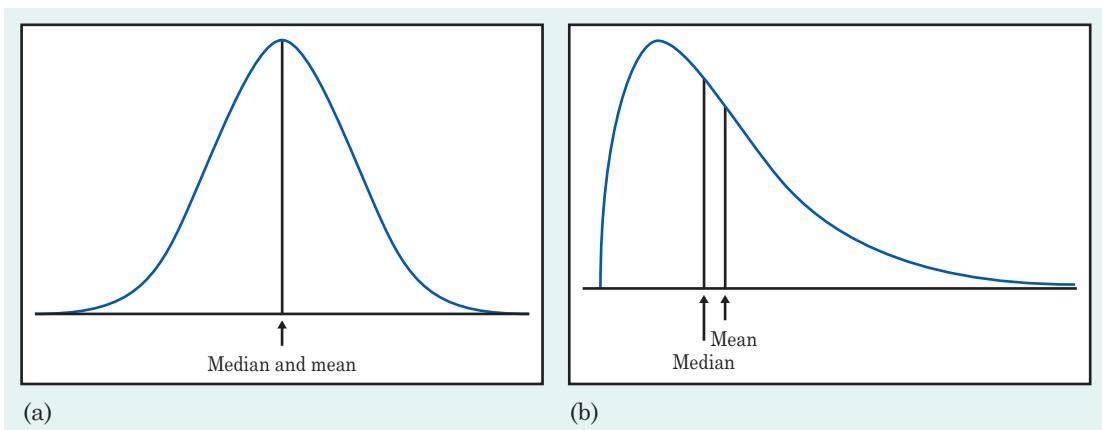


Figure 13.5 The median and mean for two density curves: (a) a symmetric Normal curve; (b) a curve that is skewed to the right.

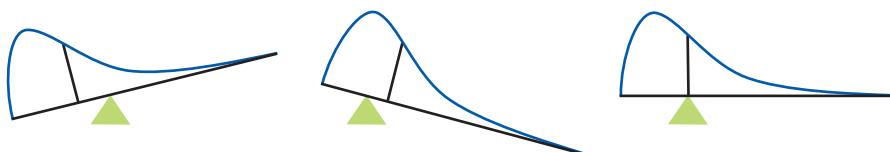


Figure 13.6 The mean of a density curve is the point at which it would balance.

the two sides are identical. *The mean and median of a symmetric density curve are equal*, as in Figure 13.5(a). We know that the mean of a skewed distribution is pulled toward the long tail. Figure 13.5(b) shows how the mean of a skewed density curve is pulled toward the long tail more than is the median.

Median and mean of a density curve

The **median** of a density curve is the equal-areas point, the point that divides the area under the curve in half.

The **mean** of a density curve is the balance point, or center of gravity, at which the curve would balance if made of solid material.

The median and mean are the same for a symmetric density curve. They both lie at the center of the curve. The mean of a skewed curve is pulled away from the median in the direction of the long tail.

Normal distributions

The density curves in Figures 13.3 and 13.4 belong to a particularly important family: the Normal curves. (We capitalize “Normal” to remind you that these curves are special.) Figure 13.7 presents two more Normal density curves. Normal curves are symmetric, single-peaked, and bell-shaped. Their tails fall off quickly, so we do not expect outliers. Because Normal distributions are symmetric, the mean and median lie together at the peak in the center of the curve.

Normal curves also have the special property that we can locate the standard deviation of the distribution by eye on the curve. This isn’t true for most other density curves. Here’s how to do it. Imagine that you are skiing

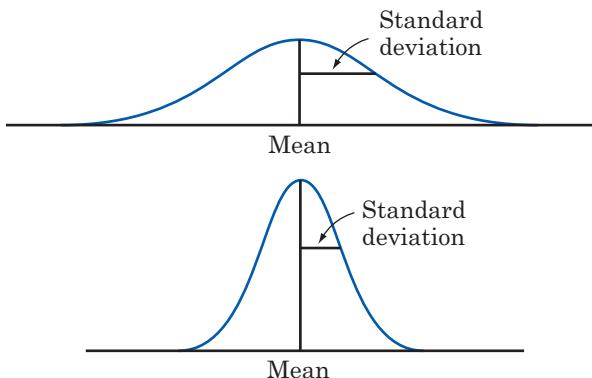
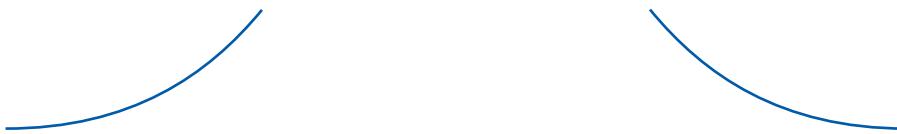


Figure 13.7 Two Normal curves. The standard deviation determines the variability of a Normal curve.

down a mountain that has the shape of a Normal curve. At first, you descend at an ever-steepener angle as you go out from the peak:



Fortunately, before you find yourself going straight down, the slope begins to grow flatter rather than steeper as you go out and down:



The points at which this change of curvature takes place are located one standard deviation on either side of the mean. The standard deviations are marked on the two curves in Figure 13.7. You can feel the change as you run a pencil along a Normal curve, and so find the standard deviation.

Normal curves have the special property that giving the mean and the standard deviation completely specifies the curve. The mean determines the center of the curve, and the standard deviation determines its shape. Changing the mean of a Normal distribution does *not* change its shape, only its location on the axis. Changing the standard deviation *does* change the shape of a Normal curve, as Figure 13.7 illustrates. The distribution with the smaller standard deviation is less variable and more sharply peaked. Here is a summary of basic facts about Normal curves.

Normal density curves

The **Normal curves** are symmetric, bell-shaped curves that have these properties:

- A specific Normal curve is completely described by giving its mean and its standard deviation.
- The mean determines the center of the distribution. It is located at the center of symmetry of the curve.
- The standard deviation determines the shape of the curve. It is the distance from the mean to the change-of-curvature points on either side.

Why are the Normal distributions important in statistics? First, Normal distributions are good descriptions for some distributions of *real data*. Normal curves were first applied to data by the great mathematician



The bell curve? Does the distribution of human intelligence follow the “bell curve” of a Normal distribution? Scores on IQ tests do roughly follow a Normal distribution. That is because a test score is calculated from a person’s answers in a way that is designed to produce a Normal distribution. To conclude that intelligence follows a bell curve, we must agree that the test scores directly measure intelligence. Many psychologists don’t think there is one human characteristic that we can call “intelligence” and can measure by a single test score.

Carl Friedrich Gauss (1777–1855), who used them to describe the small errors made by astronomers and surveyors in repeated careful measurements of the same quantity. You will sometimes see Normal distributions labeled “Gaussian” in honor of Gauss. For much of the nineteenth century, Normal curves were called “error curves” because they were first used to describe the distribution of measurement errors. As it became clear that the distributions of some biological and psychological variables were at least roughly Normal, the “error curve” terminology was dropped. The curves were first called “Normal” by Francis Galton in 1889. Galton, a cousin of Charles Darwin, pioneered the statistical study of biological inheritance.

Normal curves also describe the distribution of statistics such as sample proportions (when the sample size is large and the value of the proportion is moderate) and sample means when we take many samples from the same population. We will explore distributions of statistics in more detail in later chapters. The margins of error for the results of sample surveys are usually calculated from Normal curves. However, even though many sets of data follow a Normal distribution, many do not. Most income distributions, for example, are skewed to the right and so are not Normal. Non-Normal data, like non-normal people, not only are common but are sometimes more interesting than their normal counterparts.

The 68–95–99.7 rule

There are many Normal curves, each described by its mean and standard deviation. All Normal curves share many properties. In particular, the standard deviation is the natural unit of measurement for Normal distributions. This fact is reflected in the following rule.

The 68–95–99.7 rule

In any Normal distribution, approximately

- **68%** of the observations fall within one standard deviation of the mean.
- **95%** of the observations fall within two standard deviations of the mean.
- **99.7%** of the observations fall within three standard deviations of the mean.

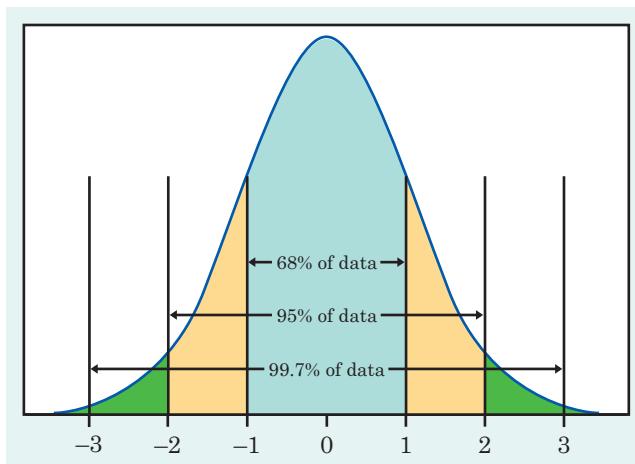


Figure 13.8 The 68–95–99.7 rule for Normal distributions. Note the x axis, or the horizontal axis, displays the number of standard deviations from the mean.

Figure 13.8 illustrates the 68–95–99.7 rule. By remembering these three numbers, you can think about Normal distributions without constantly making detailed calculations. Remember also, though, that no set of data is exactly described by a Normal curve. The 68–95–99.7 rule will be only approximately true for SAT scores or the lengths of crickets because these distributions are approximately Normal.

EXAMPLE 2 Heights of young women

The distribution of heights of women aged 18 to 24 is approximately Normal with mean 65 inches and standard deviation 2.5 inches. To use the 68–95–99.7 rule, always start by drawing a picture of the Normal curve. Figure 13.9 shows what the rule says about women's heights.

Half of the observations in any Normal distribution lie above the mean, so half of all young women are taller than 65 inches.

The central 68% of any Normal distribution lies within one standard deviation of the mean. Half of this central 68%, or 34%, lies above the mean. So 34% of young women are between 65 inches and 67.5 inches tall. Adding the 50% who are shorter than 65 inches, we see that 84% of young women have heights less than 67.5 inches. That leaves 16% who are taller than 67.5 inches.

The central 95% of any Normal distribution lies within two standard deviations of the mean. Two standard deviations is 5 inches here, so



Cultura Creative/Alamy

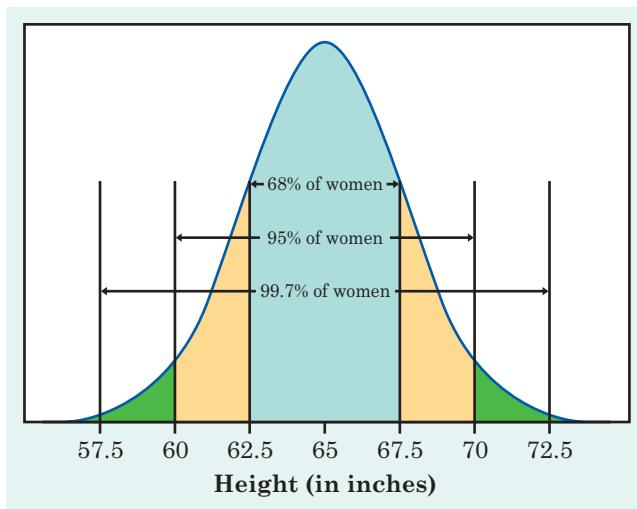


Figure 13.9 The 68–95–99.7 rule for heights of young women, Example 2. This Normal distribution has mean 65 inches and standard deviation 2.5 inches.

the middle 95% of young women's heights are between 60 inches (that's $65 - 5$) and 70 inches (that's $65 + 5$).

The other 5% of young women have heights outside the range from 60 to 70 inches. Because the Normal distributions are symmetric, half of these women are on the short side. The shortest 2.5% of young women are less than 60 inches (5 feet) tall.

Almost all (99.7%) of the observations in any Normal distribution lie within three standard deviations of the mean. Almost all young women are between 57.5 and 72.5 inches tall.

NOW IT'S YOUR TURN

13.1 Heights of young men. The distribution of heights of young men is approximately Normal with mean 70 inches and standard deviation 2.5 inches. Between which heights do the middle 95% of men fall?

Standard scores

Madison scored 600 on the SAT Mathematics college entrance exam. How good a score is this? That depends on where a score of 600 lies in the distribution of all scores. The SAT exams are scaled so that scores should roughly follow the Normal distribution with mean 500 and standard deviation 100. Madison's 600 is one standard deviation above the mean. The 68–95–99.7 rule now tells us just where she stands (Figure 13.10). Half of all scores are below 500, and another 34% are between 500 and 600. So, Madison did better than 84% of the students who took the SAT. Her score report not only will

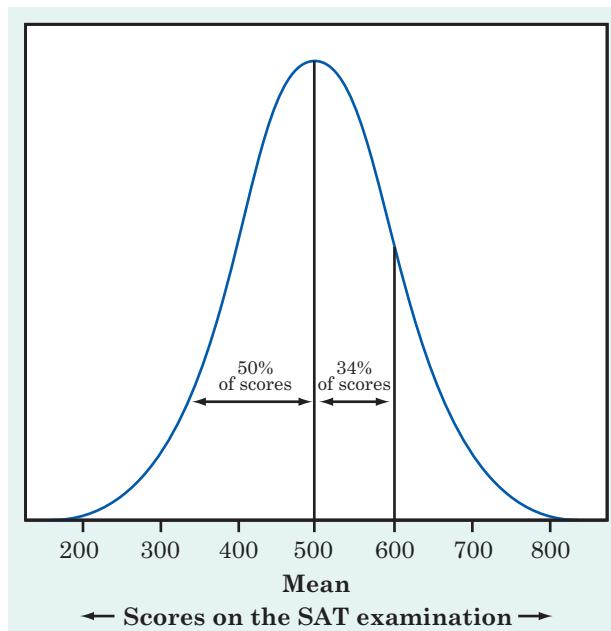


Figure 13.10 The 68–95–99.7 rule shows that 84% of any Normal distribution lies to the left of the point one standard deviation above the mean. Here, this fact is applied to SAT scores.

say she scored 600 but will add that this is at the “84th percentile.” That’s statistics speak for “You did better than 84% of those who took the test.”

Because the standard deviation is the natural unit of measurement for Normal distributions, we restated Madison’s score of 600 as “one standard deviation above the mean.” Observations expressed in standard deviations above or below the mean of a distribution are called *standard scores*. Standard scores are also sometimes referred to as *z-scores*.

Standard scores

The **standard score** for any observation is

$$\text{standard score} = \frac{\text{observation} - \text{mean}}{\text{standard deviation}}$$

A standard score of 1 says that the observation in question lies one standard deviation above the mean. An observation with standard score –2 is two standard deviations below the mean. Standard scores can be used to compare values in different distributions. Of course, you should not use standard scores unless you are willing to use the standard deviation to

describe the variability of the distributions. That requires that the distributions be at least roughly symmetric.

EXAMPLE 3 ACT versus SAT scores

Madison scored 600 on the SAT Mathematics exam. Her friend Gabriel took the American College Testing (ACT) test and scored 21 on the math part. ACT scores are Normally distributed with mean 18 and standard deviation 6. Assuming that both tests measure the same kind of ability, who has the higher score?

Madison's standard score is

$$\frac{600 - 500}{100} = \frac{100}{100} = 1.0$$

Compare this with Gabriel's standard score, which is

$$\frac{21 - 18}{6} = \frac{3}{6} = 0.5$$

Because Madison's score is 1 standard deviation above the mean and Gabriel's is only 0.5 standard deviation above the mean, Madison's performance is better.

NOW IT'S YOUR TURN

13.2 Heights of young men. The distribution of heights of young men is approximately Normal with mean 70 inches and standard deviation 2.5 inches. What is the standard score of a height of 72 inches (6 feet)?

Percentiles of Normal distributions*

For Normal distributions, but not for other distributions, standard scores translate directly into *percentiles*.

Percentiles

The **cth percentile** of a distribution is a value such that *c* percent of the observations lie below it and the rest lie above.

The median of any distribution is the 50th percentile, and the quartiles are the 25th and 75th percentiles. In any Normal distribution, the point one standard deviation above the mean (standard score 1) is the 84th percentile. Figure 13.10 shows why. Every standard score for a

*This material is not needed to read the rest of the book.

Normal distribution translates into a specific percentile, which is the same no matter what the mean and standard deviation of the original Normal distribution are. Table B at the back of this book gives the percentiles corresponding to various standard scores. This table enables us to do calculations in greater detail than does the 68–95–99.7 rule.

EXAMPLE 4 Percentiles for college entrance exams

Madison's score of 600 on the SAT translates into a standard score of 1.0. We saw that the 68–95–99.7 rule says that this is the 84th percentile. Table B is a bit more precise: it says that standard score 1 is the 84.13 percentile of a Normal distribution. Gabriel's 21 on the ACT is a standard score of 0.5. Table B says that this is the 69.15 percentile. Gabriel did well, but not as well as Madison. The percentile is easier to understand than either the raw score or the standard score. That's why reports of exams such as the SAT usually give both the score and the percentile.



Jose Luis Pelaez, Inc./CORBIS

EXAMPLE 5 Finding the observation that matches a percentile

How high must a student score on the SAT to fall in the top 10% of all scores? That requires a score at or above the 90th percentile. Look in the body of Table B for the percentiles closest to 90. You see that standard score 1.2 is the 88.49 percentile and standard score 1.3 is the 90.32 percentile. The percentile in the table closest to 90 is 90.32, so we conclude that a standard score of 1.3 is approximately the 90th percentile of *any Normal distribution*.

To go from the standard score back to the scale of SAT scores, “undo” the standard score calculation as follows:

$$\begin{aligned}\text{observation} &= \text{mean} + \text{standard score} \times \text{standard deviation} \\ &= 500 + (1.3)(100) = 630\end{aligned}$$

A score of 630 or higher will be in the top 10%. (More exactly, these scores are in the top 9.68% because 630 is exactly the 90.32 percentile.)

EXAMPLE 6 Finding the percentile below a particular value

Returning again to the distribution of SAT scores, if Ting obtains a score of 430 on this exam, what percentage of individuals score the same as or lower than Ting? To answer this question, we must first convert Ting's score to a standard score. This standard score is

$$\frac{430 - 500}{100} = \frac{-70}{100} = -0.7$$

From Table B, we can see that the standard score of -0.7 corresponds to a percentile of 24.20. Therefore, 24.20% of those individuals who took this exam score at or below Ting's score of 430.

EXAMPLE 7 Finding the percentile above a particular value

If Jordan scores 725 on the SAT Mathematics college entrance exam, what percentage of individuals would score at least as high or higher than Jordan? To answer this question, we must again begin by converting Jordan's score to a standard score. This standard score is

$$\frac{725 - 500}{100} = \frac{225}{100} = 2.25$$

To use Table B, we must first round 2.25 to 2.3. From Table B, we can see that the standard score of 2.3 corresponds to a percentile of 98.93. We might be tempted to assume our final answer will be 98.93%. If we were attempting to determine the percentage of individuals who scored 725 or lower, our final answer would be 98.93%. We want to know, however, the percentage who score 725 or higher. To find this percentage, we need to remember that the total area below the Normal curve is 1. Expressed as a percentage, the total area is 100%. To find the percentage of values that fall at or above 725, we subtract the percentage of values that fall below 725 from 100%. This gives us $100\% - 98.93\% = 1.07\%$. Our final answer is that 1.07% of individuals who take the SAT Mathematics college entrance exam score at or above Jordan's score of 725.

NOW IT'S YOUR TURN

13.3 SAT scores. How high must a student score on the SAT Mathematics test to fall in the top 25% of all scores? What percentage of students obtain scores at or below 475? What percentage of students obtain scores at or above 580?

STATISTICS IN SUMMARY

Chapter Specifics

- Stemplots, histograms, and boxplots all describe the distributions of quantitative variables.
- **Density curves** also describe distributions. A density curve is a curve with area exactly 1 underneath it whose shape describes the overall pattern of a distribution.

- An area under the curve gives the proportion of the observations that fall in an interval of values.
- You can roughly locate the median (equal-areas point) and the mean (balance point) by eye on a density curve.
- Stemplots, histograms, and boxplots are created from samples. Density curves are intended to display the idealized shape of the distribution of the population from which the samples are taken.
- **Normal curves** are a special kind of density curve that describes the overall pattern of some sets of data. Normal curves are symmetric and bell-shaped. A specific Normal curve is completely described by its mean and standard deviation. You can locate the mean (center point) and the standard deviation (distance from the mean to the change-of-curvature points) on a Normal curve. All Normal distributions obey the **68–95–99.7 rule**.
- **Standard scores** express observations in standard deviation units about the mean, which has standard score 0. A given standard score corresponds to the same **percentile** in any Normal distribution. Table B gives percentiles of Normal distributions.



Chapters 10, 11, and 12 provided us with a strategy for exploring data on a single quantitative variable.

- Make a graph, usually a histogram or stemplot.
- Look for the overall pattern (shape, center, variability) and striking deviations from the pattern.
- Choose the five-number summary or the mean and standard deviation to briefly describe the center and variability in numbers.

In this chapter, we added another step: sometimes the overall pattern of a large number of observations is so regular that we can describe it by a smooth density curve, such as the Normal curve. This step also allows us to identify “a large number of observations” as a population and use density curves to describe the distribution of a population. We did precisely this when we used the Normal distribution to describe the distribution of the heights of all young women or the scores of all students on the SAT exam.

Using a density curve to describe the distribution of a population is a convenient summary, allowing us to determine percentiles of the distribution without having to see a list of all the values in the population. It also suggests the nature of the conclusions we might draw about a single quantitative variable. Use statistics that describe the

distribution of the sample to draw conclusions about parameters that describe the distribution of a population. We will explore this in future chapters.

CASE STUDY The Normal curve that best approximates the distribution of body temperatures in Figures 13.2 and 13.3 has mean 98.25°F and standard deviation 0.73°F. Use what you have learned in this chapter to answer the following questions.

1. According to the 68–95–99.7 rule, 68% of body temperatures fall between what two values? Between what two values do 95% of body temperatures fall?
2. There was a time when 98.6°F was considered the average body temperature. Given what you know about the distribution of body temperatures given in Figures 13.2 and 13.3, what percentage of individuals would you expect to have body temperatures greater than 98.6°F? What percentage would you expect to have body temperature less than 98.6°F?



LaunchPad Online Resources

macmillan learning

- *LearningCurve* has good questions to check your understanding of the concepts.

CHECK THE BASICS

For Exercise 13.1, see page 302; for Exercise 13.2, see page 304; for Exercise 13.3, see page 306.

13.4 Density curves. One characteristic of a density curve is that there is a specific total area under the curve. What is this area equal to?

- (a) Exactly 1.
- (b) Approximately 1.
- (c) It depends on what is being measured.
- (d) It depends on whether the distribution is Normal.

13.5 The mean and median. Which of the following is an incorrect statement?

(a) If a density curve is skewed to the right, the mean will be larger than the median.

(b) In a symmetric density curve, the mean is equal to the median.

(c) The median is the balance point in a density curve.

(d) The mean of a skewed distribution is pulled toward the long tail.

13.6 Pulse rates. Suppose that resting pulse rates for healthy adults are found to follow a Normal distribution, with a mean of 69 beats per minute and a standard deviation of 9.5 beats per minutes. If Bonnie has a pulse rate of 78.5 beats per minute, this means that

- (a) approximately 32% of adults have pulse rates higher than Bonnie's.
- (b) approximately 16% of adults have pulse rates higher than Bonnie's.
- (c) Bonnie's pulse rate is two standard deviations above the mean.
- (d) Bonnie's pulse rate, when converted to a standard score, would be 1.5.

13.7 More on pulse rates. Let's again assume that the resting pulse rates for healthy adults follow a Normal distribution with a mean of 69 beats per minute and a standard deviation of 9.5 beats per minute. When converted to a standard score, Adam's pulse rate becomes -0.5 . How should this standard score be interpreted?

- (a) Adam should see a doctor because his pulse rate is unusually low.
- (b) Adam's pulse rate is above the mean.

- (c) Exactly 5% of healthy adults have pulse rates less than Adam's pulse rate.
- (d) Adam's pulse rate is about one-half of a standard deviation below the mean.

13.8 Reading test scores. A standardized reading test is given to fifth-grade students. Scores on this test are Normally distributed, with a mean of 32 points and a standard deviation of 8 points. When Corey gets his test results, he is told that his score is at the 95th percentile. What does this mean?

- (a) Corey's score is higher than 95% of the students who took this test.
- (b) Corey's score is 48 points.
- (c) Corey's score is exactly two standard deviations above the mean.
- (d) All of the above.

CHAPTER 13 EXERCISES

13.9 Density curves

- (a) Sketch a density curve that is symmetric but has a shape different from that of the Normal curves.
- (b) Sketch a density curve that is strongly skewed to the right.

13.10 Mean and median. Figure 13.11 shows density curves of several shapes. Briefly describe the overall shape of each distribution. Two or more points are marked on each curve. The mean and the median are among these points. For each curve, which point is the median and which is the mean?

13.11 Random numbers. If you ask a computer to generate "random numbers" between 0 and 1, you will get

observations from a *uniform distribution*. Figure 13.12 shows the density curve for a uniform distribution. This curve takes the constant value 1 between 0 and 1 and is zero outside that range. Use this density curve to answer these questions.

- (a) Why is the total area under the curve equal to 1?
- (b) The curve is symmetric. What is the value of the mean and median?
- (c) What percentage of the observations lie between 0 and 0.1?
- (d) What percentage of the observations lie between 0.6 and 0.9?

IQ test scores. Figure 13.13 is a stemplot of the IQ test scores of 74

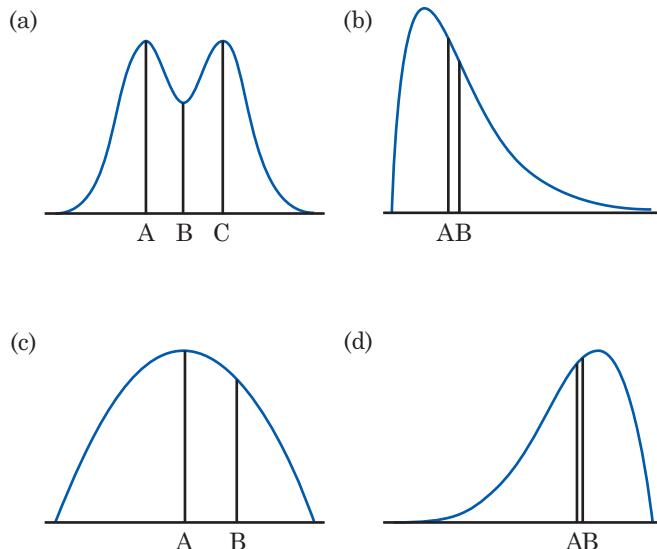


Figure 13.11 Four density curves of different shapes, Exercise 13.10. In each case, the mean and the median are among the marked points.

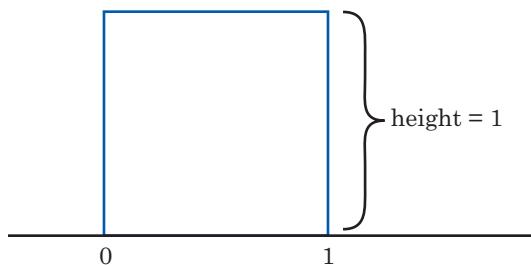


Figure 13.12 The density curve of a uniform distribution, Exercise 13.11. Observations from this distribution are spread “at random” between 0 and 1.

seventh-grade students. This distribution is very close to Normal with mean 111 and standard deviation 11. It includes all the seventh-graders in a rural Midwest school except for four low outliers who were dropped because they may have been ill or otherwise not paying attention to the test. Take the Normal distribution with mean 111 and standard deviation 11 as a description of the IQ test scores of all rural

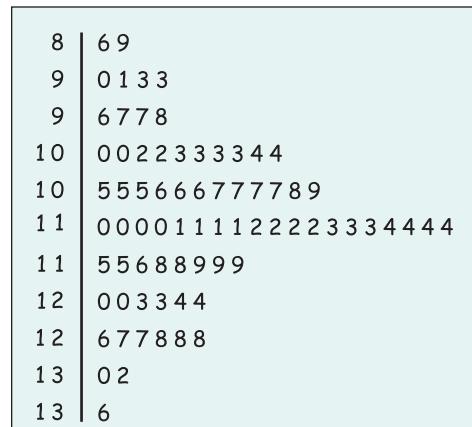


Figure 13.13 Stemplot of the IQ scores of 74 seventh-grade students, Exercises 13.12, 13.13, and 13.14.

Midwest seventh-grade students. Use this distribution and the 68–95–99.7 rule to answer Exercises 13.12 to 13.14.

13.12 Between what values do the IQ scores of 68% of all rural Midwest seventh-graders lie?

13.13 What percentage of IQ scores for rural Midwest seventh-graders are more than 100?

13.14 What percentage of all students have IQ scores below 78? None of the 74 students in our sample school had scores this low. Are you surprised at this? Why?

13.15 Length of pregnancies. The length of human pregnancies from conception to birth varies according to a distribution that is approximately Normal with mean 266 days and standard deviation 16 days. Use the 68–95–99.7 rule to answer the following questions.

- Almost all (99.7%) pregnancies fall in what range of lengths?
- How long are the longest 2.5% of all pregnancies?

13.16 A Normal curve. What are the mean and standard deviation of the Normal curve in Figure 13.14?

13.17 Horse pregnancies. Bigger animals tend to carry their young longer before birth. The length of horse pregnancies from conception to birth varies according to a roughly Normal distribution with mean 336 days and standard deviation of three days. Use the 68–95–99.7 rule to answer the following questions.

- Between what values do the lengths of the middle 95% of all horse pregnancies fall?
- What percentage of horse pregnancies are less than 333 days?

13.18 Three great hitters. Three landmarks of baseball achievement are Ty Cobb's batting average of .420 in 1911, Ted Williams's .406 in 1941, and

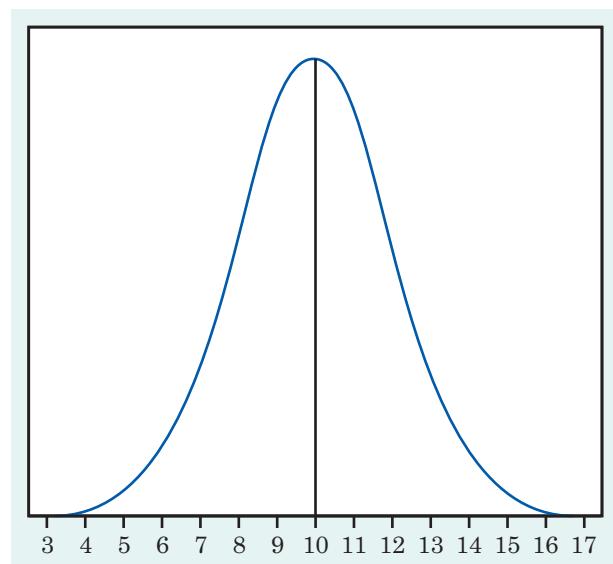


Figure 13.14 What are the mean and standard deviation of this Normal density curve? Exercise 13.16.

George Brett's .390 in 1980. These batting averages cannot be compared directly because the distribution of major league batting averages has changed over the years. The distributions are quite symmetric and (except for outliers such as Cobb, Williams, and Brett) reasonably Normal. While the mean batting average has been held roughly constant by rule changes and the balance between hitting and pitching, the standard deviation has dropped over time. Here are the facts:

Decade	Mean	Std. dev.
1910s	.266	.0371
1940s	.267	.0326
1970s	.261	.0317

Compute the standard scores for the batting averages of Cobb, Williams, and Brett to compare how far each stood above his peers.

13.19 Comparing IQ scores. The Wechsler Adult Intelligence Scale (WAIS) is an IQ test. Scores on the WAIS for the 20–34 age group are approximately Normally distributed with mean 110 and standard deviation 15. Scores for the 60–64 age group are approximately Normally distributed with mean 90 and standard deviation 15. Sarah, who is 30, scores 130 on the WAIS. Her mother, who is 60, takes the test and scores 110.

- Express both scores as standard scores that show where each woman stands within her own age group.
- Who scored higher relative to her age group, Sarah or her mother? Who has the higher absolute level of the variable measured by the test?

13.20 Men's heights. The distribution of heights of young men is approximately Normal with mean 70 inches and standard deviation 2.5 inches. Sketch a Normal curve on which this mean and standard deviation are correctly located. (*Hint:* Draw the curve first, locate the points where the curvature changes, and then mark the horizontal axis.)

13.21 More on men's heights. The distribution of heights of young men is approximately Normal with mean 70 inches and standard deviation 2.5 inches. Use the 68–95–99.7 rule to answer the following questions.

- What percentage of men are taller than 77.5 inches?
- Between what heights do the middle 68% of men fall?
- What percentage of men are shorter than 65 inches?

13.22 Heights of men and women. The heights of young women are approximately Normal with mean 65 inches and standard deviation 2.5 inches. The heights of men in the same age group have mean 70 inches and standard deviation 2.5 inches. What percentage of women are taller than a man of average (mean) height?

13.23 Heights of young adults. The mean height of men aged 18–24 is about 70 inches. Women that age have a mean height of about 65 inches. Do you think that the distribution of heights for all Americans aged 18–24 is approximately Normal? Explain your answer.

13.24 Sleep. The distribution of hours of sleep per week night, among college students, is found to be Normally distributed, with a mean of seven hours and a standard deviation of one hour. Use this information and the 68–95–99.7 rule to answer the following questions.

- What percentage of college students sleep more than seven hours? More than eight hours?
- What range contains the middle 95% of hours slept per week night by college students?

13.25 Cholesterol. Low density lipoprotein, or LDL, is the main source of cholesterol buildup and blockage in the arteries. This is why LDL is known as “bad cholesterol.” LDL is measured in milligrams per deciliter of blood, or mg/dL. In a population of adults at risk for cardiovascular problems, the distribution of LDL levels is Normal, with a mean of 123 mg/dL and a standard deviation of 41 mg/dL.

If an individual's LDL is at least one standard deviation or more above the mean, he or she will be monitored carefully by a doctor. What percentage of individuals from this population will have LDL levels one or more standard deviations above the mean?

The following optional exercises require use of Table B of Normal distribution percentiles.

13.26 NCAA rules for athletes. The National Collegiate Athletic Association (NCAA) requires Division II athletes to get a combined score of at least 820 on the Mathematics and Critical Reading sections of the SAT exam in order to compete in their first college year. In 2014, the combined scores of the millions of college-bound seniors taking the SATs were approximately Normal with mean 1010 and standard deviation approximately 218. What percentage of all college-bound seniors had scores less than 820?

13.27 More NCAA rules. For Division I athletes, the NCAA uses a sliding scale, based on both core GPA and the combined Mathematics and Critical Reading SAT score, to determine eligibility to compete in the first year of college. For athletes with a core GPA of 3.0, a score of at least 620 on the combined Mathematics and Critical Reading sections of the SAT exam is required. Use the information in the previous exercise to find the percentage of all SAT scores of college-bound seniors that are less than 620.

13.28 800 on the SAT. It is possible to score higher than 800 on the SAT, but scores above 800 are reported as 800. (That is, a student can get a reported

score of 800 without a perfect paper.) In 2014, the scores of college-bound senior men on the SAT Math test followed a Normal distribution with mean 530 and standard deviation 123. What percentage of scores were above 800 (and so reported as 800)?

13.29 Women's SAT scores. The average performance of women on the SAT, especially the math part, is lower than that of men. The reasons for this gender gap are controversial. In 2014, college-bound senior women's scores on the SAT Math test followed a Normal distribution with mean 499 and standard deviation 114. The mean for men was 530. What percentage of women scored higher than the male mean?

13.30 Are we getting smarter? When the Stanford-Binet IQ test came into use in 1932, it was adjusted so that scores for each age group of children followed roughly the Normal distribution with mean 100 and standard deviation 15. The test is readjusted from time to time to keep the mean at 100. If present-day American children took the 1932 Stanford-Binet test, their mean score would be about 120. The reasons for the increase in IQ over time are not known but probably include better childhood nutrition and more experience in taking tests.

- IQ scores above 130 are often called "very superior." What percentage of children had very superior scores in 1932?
- If present-day children took the 1932 test, what percentage would have very superior scores? (Assume that the standard deviation 15 does not change.)

13.31 Japanese IQ scores. The Wechsler Intelligence Scale for Children is used (in several languages) in the United States and Europe. Scores in each case are approximately Normally distributed with mean 100 and standard deviation of 15. When the test was standardized in Japan, the mean was 111. To what percentile of the American-European distribution of scores does the Japanese mean correspond?

13.32 The stock market. The annual rate of return on stock indexes (which combine many individual stocks) is very roughly Normal. Since 1945, the Standard & Poor's 500 index has had a mean yearly return of 12.5%, with a standard deviation of 17.8%. Take this Normal distribution to be the distribution of yearly returns over a long period.

- (a) In what range do the middle 95% of all yearly returns lie?
- (b) The market is down for the year if the return on the index is less than

zero. In what proportion of years is the market down?

- (c) In what proportion of years does the index gain 25% or more?

13.33 Locating the quartiles. The quartiles of any distribution are the 25th and 75th percentiles. About how many standard deviations from the mean are the quartiles of any Normal distribution?

13.34 Young women's heights. The heights of women aged 18–24 are approximately Normal with mean 65 inches and standard deviation 2.5 inches. How tall are the tallest 10% of women? (Use the closest percentile that appears in Table B.)

13.35 High IQ scores. Scores on the Wechsler Adult Intelligence Scale for the 20 to 34 age group are approximately Normally distributed with mean 110 and standard deviation 15. How high must a person score to be in the top 10% of all scores?



EXPLORING THE WEB

Follow the QR code to access exercises.