14-4

The Normal Distribution

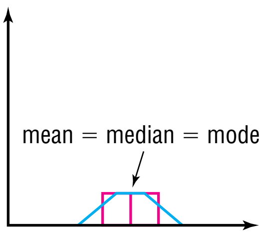
OBJECTIVES
- Use the normal distribution curve.

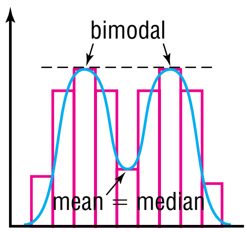
Real World Application

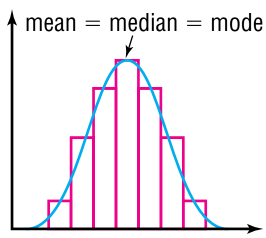
TESTING The class of 1996 was the first class to take the adjusted Scholastic Assessment Test. The test was adjusted so that the median of the scores for the verbal section and the math section would be 500. For each section, the lowest score is 200 and the highest is 800. Suppose the verbal and math scores follow the *normal distribution.* What percent of the students taking the test would have a math score between 375 and 625? *This problem will be solved in Example 4.*

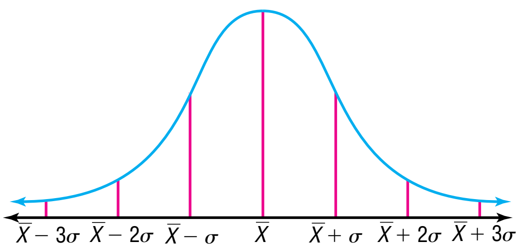
A frequency polygon displays a limited number of data and may not represent an entire population. To display the frequency of an entire population, a smooth curve is used rather than a polygon.

If the curve is symmetric, then information about the measures of central tendency can be gathered from the graph. Study the graphs below.









A **normal distribution** is a frequency distribution that often occurs when there is a large number of values in a set of data. The graph of this distribution is a symmetric, bell-shaped curve, shown at the left. This is known as a **normal curve.** The shape of the curve indicates that the frequencies in a normal distribution are concentrated around the center portion of the distribution. A small portion of the population occurs at the extreme values.

In a normal distribution, small deviations are much more frequent than large ones. Negative deviations and positive deviations occur with the same frequency. The points on the horizontal axis represent values that are a certain number of standard deviations from the mean $\overline{X}$. In the curve shown above, each interval represents one standard deviation. So, the section from $\overline{X}$ to $\overline{X} + \sigma$ represents those values between the mean and one standard deviation greater than the mean, the section from $\overline{X} + \sigma$ to $\overline{X} + 2\sigma$ represents the interval one standard deviation greater than the mean to two standard deviations greater than the mean, and so on. The total area under the normal curve and above the horizontal axis represents the total probability of the distribution, which is 1.

Example 1 Real World Application

MEDICINE **The average healing time of a certain type of incision is 240 hours with a standard deviation of 20 hours. Sketch a normal curve that represents the frequency of healing times.**

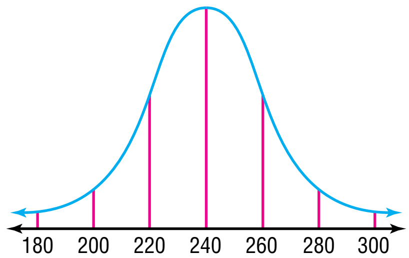
First, find the values defined by the standard deviation in a normal distribution.

$\overline{X} - 1\sigma = 240 - 1(20)$ or 220 $\qquad$ $\overline{X} + 1\sigma = 240 + 1(20)$ or 260

$\overline{X} - 2\sigma = 240 - 2(20)$ or 200 $\qquad$ $\overline{X} + 2\sigma = 240 + 2(20)$ or 280

$\overline{X} - 3\sigma = 240 - 3(20)$ or 180 $\qquad$ $\overline{X} + 3\sigma = 240 + 3(20)$ or 300

Sketch the general shape of a normal curve. Then, replace the horizontal scale with the values you have calculated.



The tables below give the fractional parts of a normally distributed set of data for selected areas about the mean. The letter t represents the number of standard deviations from the mean (that is, $\overline{X} \pm t\sigma$). When $t = 1$, t represents 1 standard deviation above and below the mean.

P represents the fractional part of the data that lies in the interval $\overline{X} \pm t\sigma$. The percent of the data within these limits is $100P$.

t	P
0.0	0.000
0.1	0.080
0.2	0.159
0.3	0.236
0.4	0.311
0.5	0.383
0.6	0.451
0.7	0.516
0.8	0.576

t	P
0.9	0.632
1.0	0.683
1.1	0.729
1.2	0.770
1.3	0.807
1.4	0.838
1.5	0.866
1.6	0.891
1.65	0.900

t	P
1.7	0.911
1.8	0.929
1.9	0.943
1.96	0.950
2.0	0.955
2.1	0.964
2.2	0.972
2.3	0.979
2.4	0.984

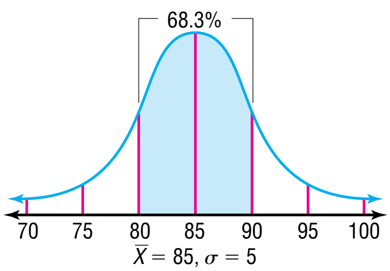
t	P
2.5	0.988
2.58	0.990
2.6	0.991
2.7	0.993
2.8	0.995
2.9	0.996
3.0	0.997
3.5	0.9995
4.0	0.9999

The P value also corresponds to the probability that a randomly selected member of the sample lies within t standard deviation units of the mean. For example, suppose the mean of a set of data is 85 and the standard deviation is 5.

Boundaries:

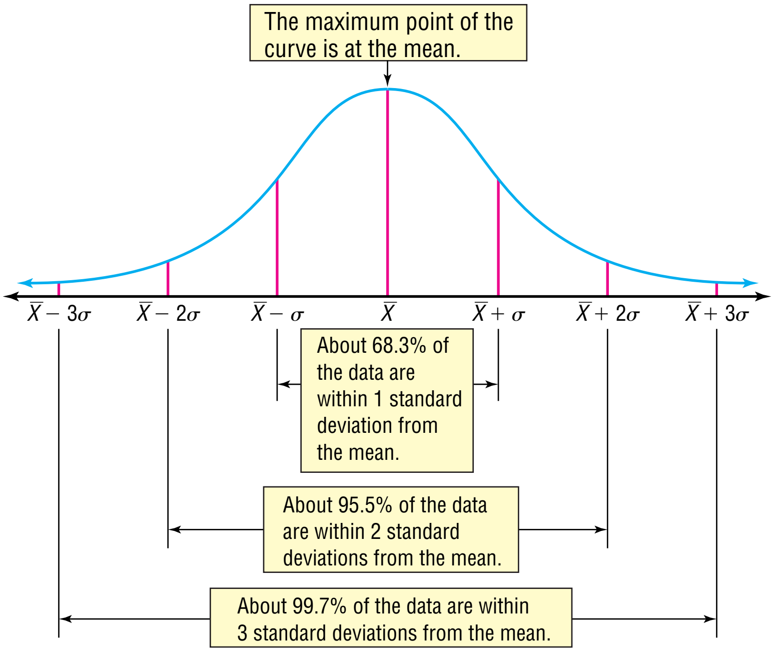
$$\overline{X} - t\sigma \text{ to } \overline{X} + t\sigma$$
$$85 - t(5) \text{ to } 85 + t(5)$$
$$85 - 1(5) \text{ to } 85 + 1(5)$$
$$80 \text{ to } 90$$

68.3% of the values in this set of data lie within one standard deviation of 85; that is, between 80 and 90.



If you randomly select one item from the sample, the probability that the one you pick will be between 80 and 90 is 0.683. If you repeat the process 1000 times, approximately 68.3% (about 683) of those selected will be between 80 and 90.

Thus, normal distributions have the following properties.



Example 2 Real World Application

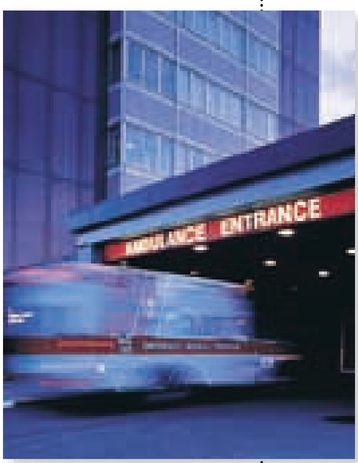
MEDICINE **Refer to Example 1. Suppose a hospital has treated 2000 patients in the past five years having this type of incision. Estimate how many patients healed in each of the following intervals.**

a. 220–260 hours

The interval 220–260 hours represents $\overline{X} \pm 1\sigma$, which represents a probability of 68.3%.

$$68.3\%(2000) = 1366$$

Approximately 1366 patients took between 220 and 260 hours to heal.



b. 200–280 hours

The interval 200–280 hours represents $\overline{X} \pm 2\sigma$, which represents a probability of 95.5%.

$$95.5\%(2000) = 1910$$

Approximately 1910 patients took between 200 and 280 hours to heal.

c. 180–300 hours

The interval 180–300 hours represents $\overline{X} \pm 3\sigma$, which represents a probability of 99.7%.

$$99.7\%(2000) = 1994$$

Approximately 1994 patients took between 180 and 300 hours to heal.

If you know the mean and the standard deviation, you can find a range of values for a given probability.

Example 3 **Find the upper and lower limits of an interval about the mean within which 45% of the values of a set of normally distributed data can be found if $\overline{X} = 110$ and $\sigma = 15$.**

Use the table on page 919 to find the value of t that most closely approximates $P = 0.45$. For $t = 0.6$, $P = 0.451$. Choose $t = 0.6$. Now find the limits.

$\overline{X} \pm t\sigma = 110 \pm 0.6(15)$ *$\overline{X} = 110$, $t = 0.6$, $\sigma = 15$*

$= 101$ and 119

The interval in which 45% of the data lies is 101–119.

If you know the mean and standard deviation, you can also find the percent of the data that lies within a given range of values.

Example 4 Real World Application

TESTING **Refer to the application at the beginning of the lesson.**

a. Determine the standard deviation.

b. What percent of the students taking the test would have a math score between 375 and 625?

c. What is the probability that a senior chosen at random has a math score between 550 and 650?

a. For a normal distribution, both the median and the mean are 500. All of the scores are between 200 and 800. Therefore, all of the scores must be within 300 points from the mean. In a normal distribution, 0.9999 of the data is within 4 standard deviations of the mean. If the scores are to be a normal distribution, the standard deviation should be $\frac{300}{4}$ or 75.

b. Write each of the limits in terms of the mean.

$$375 = 500 - 125 \text{ and } 625 = 500 + 125$$

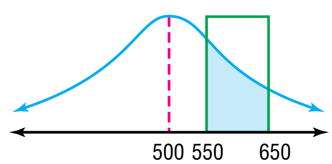
Therefore, $\overline{X} \pm t\sigma = 500 \pm 125$ and $t\sigma = 125$. Solve for t.

$$t\sigma = 125$$
$$t(75) = 125 \quad \textit{σ = 75}$$
$$t \approx 1.7$$

If $t = 1.7$, then $P = 0.911$. *Use the table on page 919.*

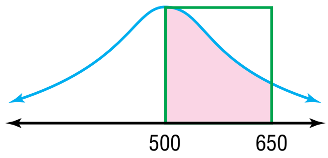
About 91.1% of the students taking the test would have a math score between 375 and 625.

c. The graph shows that 550–650 does not define an interval that can be represented by $\overline{X} \pm t\sigma$. However, the interval can be defined as the difference between the intervals 500–650 and 500–550.



(continued on the next page)

First, find the probability that the score is between the mean 500 and the upper limit 650.

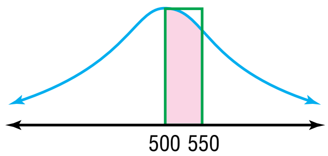


$$\bar{X} + t\sigma = 650$$
$$500 + t(75) = 650$$
$$t = 2$$

The value of P that corresponds to $t = 2$ is 0.955.

$P = 0.955$ describes the probability that a student's score falls $\pm 2(75)$ points about the mean, or between 350 and 650, but we are only considering half that interval. So, the probability that a student's score is between 500 and 650 is $\frac{1}{2}(0.955)$ or about 0.478.

Next, find the probability that a score is between the mean and the lower limit 550.

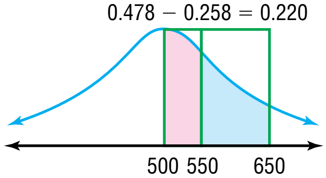


$$\bar{X} + t\sigma = 550$$
$$500 + t(75) = 550$$
$$t \approx 0.7$$

For $t = 0.7$, $P = 0.516$.

Likewise, we will only consider half of this probability or 0.258.

Now find the probability that a student's score falls in the interval 550–650.



$$P(550\text{–}650) = P(500\text{–}650) - P(500\text{–}550)$$
$$P \approx 0.478 - 0.258$$
$$P \approx 0.220 \text{ or } 22\%$$

The probability that a student's score is between 550 and 650 is about 22%.

Students who take the SAT or ACT tests will receive a score as well as a **percentile.** The percentile indicates how the student's score compares with other students taking the test.

Percentile

The nth percentile of a set of data is the value in the set such that n percent of the data is less than or equal to that value.

Therefore if a student scores in the 65th percentile, this means that 65% of the students taking the test scored the same or less than that student.

Check for Understanding

Communicating Mathematics

Read and study the lesson to answer each question.

1. **Compare** the median, mean, and mode of a set of normally distributed data.
2. **Write** an expression for the interval that is within 1.5 standard deviations from the mean.
3. **Sketch** a normal curve with a mean of 75 and a standard deviation of 10 and a normal curve with a mean of 75 and a standard deviation of 5. Which curve displays less variability?

4. **Counterexample** Draw a curve that represents data which is not normally distributed.
5. **Name** the percentile that describes the median.

Guided Practice

6. The mean of a set of normally distributed data is 550 and the standard deviation is 35.
   a. Sketch a curve that represents the frequency distribution.
   b. What percent of the data is between 515 and 585?
   c. Name the interval about the mean in which about 99.7% of the data are located.
   d. If there are 200 values in the set of data, how many would be between 480 and 620?
7. A set of 500 values is normally distributed with a mean of 24 and a standard deviation of 2.
   a. What percent of the data is in the interval 22-26?
   b. What percent of the data is in the interval 20.5-27.5?
   c. Find the interval about the mean that includes 50% of the data.
   d. Find the interval about the mean that includes 95% of the data.
8. **Education** In her first semester of college, Salali earned a grade of 82 in chemistry and a grade of 90 in speech.
   a. The mean of the chemistry grades was 73, and the standard deviation was 3. Draw a normal distribution for the chemistry grades.
   b. The mean of the speech grades was 80, and the standard deviation was 5. Draw a normal distribution for the speech grades.
   c. Which of Salali's grades is relatively better based on standard deviation from the mean? Explain.

EXERCISES

Practice

9. The mean of a set of normally distributed data is 12 and the standard deviation is 1.5.
   a. Sketch a curve that represents the frequency distribution.
   b. Name the interval about the mean in which about 68.3% of the data are located.
   c. What percent of the data is between 7.5 and 16.5?
   d. What percent of the data is between 9 and 15?
10. Suppose 200 values in a set of data are normally distributed.
   a. How many values are within one standard deviation of the mean?
   b. How many values are within two standard deviations of the mean?
   c. How many values fall in the interval between the mean and one standard deviation above the mean?
11. A set of data is normally distributed with a mean of 82 and a standard deviation of 4.
   a. Find the interval about the mean that includes 45% of the data.
   b. Find the interval about the mean that includes 80% of the data.
   c. What percent of the data is between 76 and 88?
   d. What percent of the data is between 80.5 and 83.5?

12. The mean of a set of normally distributed data is 402, and the standard deviation is 36.

a. Find the interval about the mean that includes 25% of the data.

b. What percent of the data is between 387 and 417?

c. What percent of the data is between 362 and 442?

d. Find the interval about the mean that includes 45% of the data.

13. A set of data is normally distributed with a mean of 140 and a standard deviation of 20.

a. What percent of the data is between 100 and 150?

b. What percent of the data is between 150 and 180?

c. Find the value that defines the 75th percentile.

14. The mean of a set of normally distributed data is 6, and the standard deviation is 0.35.

a. What percent of the data is between 6.5 and 7?

b. What percent of the data is between 5.5 and 6.2?

c. Find the limit above which 90% of the data lies.

Applications and Problem Solving

15. **Probability** Tossing six coins is a binomial experiment.

a. Find each probability.

P(no tails) P(one tail) P(two tails) P(three tails)
P(four tails) P(five tails) P(six tails)

b. Assume that the experiment was repeated 64 times. Make a bar graph showing how many times you would expect each outcome to occur.

c. Use the bar graph to determine the mean number of tails.

d. Find the standard deviation of the number of tails.

e. Compare the bar graph to a normal distribution.

16. **Critical Thinking** Consider the percentile scores on a standardized test.

a. Describe the 92nd percentile in terms of standard deviation.

b. Name the percentile of a student whose score is 0.8 standard deviation above the mean.

17. **Health** The lengths of babies born in City Hospital in the last year are normally distributed. The mean length is 20.4 inches, and the standard deviation is 0.8 inch. Trey was 22.3 inches long at birth.

a. What percent of the babies born at City Hospital were longer than Trey?

b. What percent of the babies born at City Hospital were shorter than Trey?

18. **Business** The length of time a brand of CD players can be used before needing service is normally distributed. The mean length of time is 61 months, and the standard deviation is 5 months. The manufacturer plans to issue a guarantee that it will replace any CD player that breaks within a certain length of time. If the manufacturer does not want to replace any more than 2% of the CD players, how many months should they limit the guarantee?

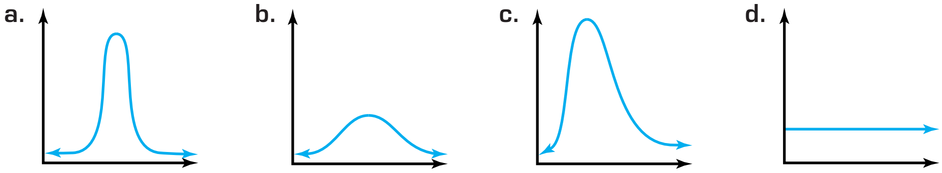
19. **Education** A college professor plans to grade a test on a curve. The mean score on the test is 65, and the standard deviation is 7. The professor wants 15% A's, 20% B's, 30% C's, 20% D's and 15% F's. Assume the grades are normally distributed.

a. What is the lowest score for an A?

b. Find the lowest passing score.

c. What is the interval for the B's?

20. Critical Thinking Describe the frequency distribution represented by each graph.



21. Industry A machine is used to fill cans of cola. The amount of cola dispensed into each can varies slightly. Suppose the amount of cola dispensed into the cans is normally distributed.

a. If at least 95% of the cans must have between 350 and 360 milliliters of cola, find the greatest standard deviation that can be allowed.

b. What percent of the cans will have between 353 and 357 milliliters of cola?

Mixed Review

22. Nutrition The numbers of Calories in one serving of twenty different cereals are listed below. *(Lesson 14-3)*

110, 110, 330, 200, 88, 110, 88, 110, 165, 390, 150, 440, 536, 200, 110, 165, 88, 147, 110, 165

a. What is the median of the data?

b. Find the first quartile point and the third quartile point.

c. Find the interquartile range of the data.

d. What is the semi-quartile range of the data?

e. Draw a box-and-whisker plot of the data.

23. Find the mean, median, and mode of {33, 42, 71, 19, 42, 45, 79, 48, 55}. *(Lesson 14-2)*

24. Write an equation for a secant function with a period of $\frac{\pi}{2}$, a phase shift of $-\pi$, and a vertical shift of 3. *(Lesson 6-7)*

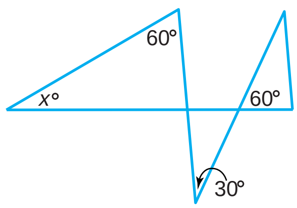
25. Education The numbers of students attending Wilder High School during various years are listed below. *(Lesson 4-8)*

Year	1965	1970	1975	1980	1985	1990	1995	2000
Enrollment	365	458	512	468	426	401	556	629

a. Write an equation that models the enrollment as a function of the number of years since 1965.

b. Use the model to predict the enrollment in 2015.

26. SAT/ACT Practice What is the value of x in the figure at the right?



A 50 **B** 45

C 40 **D** 35

E 30

Extra Practice See p. A54.

GRAPHING CALCULATOR EXPLORATION

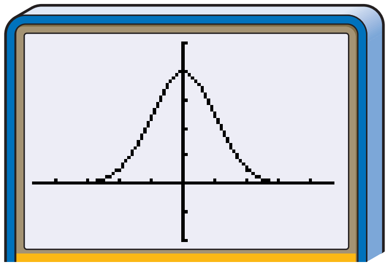
14-4B The Standard Normal Curve

An Extension of Lesson 14-4

WHAT YOU'LL LEARN

- Use the standard normal curve to study properties of normal distributions.

The graph shown at the right is known as the *standard normal curve.* The standard normal curve is the graph of $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. You can use a graphing calculator to investigate properties of this function and its graph. Enter the function for the normal curve in the **Y=** list of a graphing calculator.



[−4.7, 4.7] scl:1 by [−0.2, 0.5] scl:0.1

TRY THESE

1. What can you say about the function and its graph? Be sure to include information about the domain, the range, symmetry, and end behavior.
2. The standard normal curve models a probability distribution. As a result, probabilities for intervals of x-values are equal to areas of regions bounded by the curve, the x-axis, and the vertical lines through the endpoints of the intervals. The calculator can approximate the areas of such regions. To find the area of the region bounded by the curve, the x-axis, and the vertical lines $x = -1$ and $x = 1$, go to the **CALC** menu and select **7:∫f(x) dx**. Move the cursor to the point where $x = -1$. Press ENTER . Then move the cursor to the point where $x = 1$ and press ENTER . The calculator will shade the region described above and display its approximate area. What number does the calculator display for the area of the shaded region?
3. Refer to the diagram on page 920. For normal distributions, about what percent of the data are within one standard deviation from the mean? How is this number related to the area you found in Exercise 2?
4. Enter 2nd **[DRAW]** 1. This causes the calculator to clear the shading and redisplay the graph. Find the area of the region bounded by the curve, the x-axis, and the vertical lines $x = -2$ and $x = 2$.
5. Find the area of the region bounded by the curve, the x-axis, and the vertical lines $x = -3$ and $x = 3$.
6. How do your answers for Exercises 4 and 5 compare to the percents in the diagram on page 920?

WHAT DO YOU THINK?

7. Without using a calculator, estimate the area of the region bounded by the curve, the x-axis, and the vertical lines $x = -4$ and $x = 4$ to four decimal places.
8. Change the graphing window to **Xmin** $= -47$ and **Xmax** $= 47$. Find the area of the region bounded by the curve, the x-axis, and the vertical lines $x = -20$ and $x = 20$. Do you think that your answer is the exact area for the region? Explain.