1.1

Recharge It!

Normal Distributions

LEARNING GOALS

In this lesson, you will:

- Differentiate between discrete data and continuous data.
- Draw distributions for continuous data.
- Recognize the difference between normal distributions and non-normal distributions.
- Recognize and interpret properties of a normal curve and a normal distribution.
- Describe the effect of changing the mean and standard deviation on a normal curve.

KEY TERMS

- discrete data
- continuous data
- sample
- population
- normal curve
- normal distribution
- mean (μ)
- standard deviation (σ)

Imagine carrying around a cell phone that weighed 80 pounds, provided 30 minutes of talk time on a 100% charged battery, needed 10 hours to fully recharge the battery, and worked in only one assigned local calling area! That's a snapshot of a cell phone in the 1950s.

Cell phones have come a long way since then. Today's cell phone users send and receive texts, emails, photos and videos, they surf the web, play games, use GPS, listen to music, and much more—all on a device that fits in the palm of your hand.

1

PROBLEM 1 Low Battery

Recall that a discrete graph is a graph of isolated points and a continuous graph is a graph of points that are connected by a line or smooth curve on the graph. Data can also be discrete or continuous.

Discrete data are data whose possible values are countable and often finite. The scores of baseball games are examples of discrete data, because a team's score must be a positive whole number or zero.

Continuous data are data which can take any numerical value within a range. Heights of students, times required to complete a test, and distances between cities are examples of continuous data.

Suppose that two cell phone companies, E-Phone and Unlimited, claim that the cell phones of two of their comparable models have a mean battery life of 10 hours.

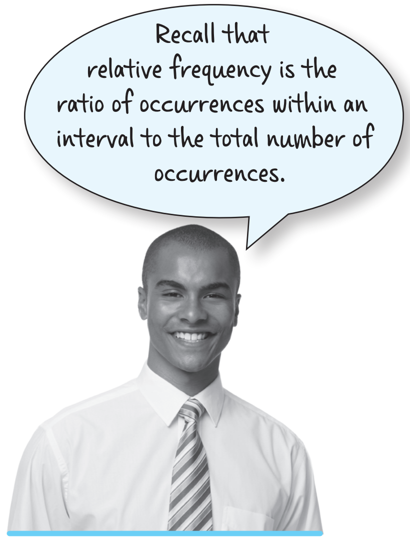
1. Are the durations of the cell phone batteries examples of discrete data or continuous data? Explain your reasoning.

2. If the mean battery life is 10 hours, does that indicate that all of E-Phone's phones and all of Unlimited's phones have a 10-hour battery life? Explain your reasoning.

One way to display continuous data is by using a relative frequency table. The relative frequency tables shown display the battery lives of a *sample* of 100 E-Phone cell phones and 100 Unlimited cell phones.

A **sample** is a subset of data selected from a *population*. A **population** represents all the possible data that are of interest in a study or survey.

The battery lives are divided into intervals. Each interval includes the first value but does not include the second value. For example, the interval 8.0–8.5 includes phones with battery lives greater than or equal to 8 hours and less than 8.5 hours.



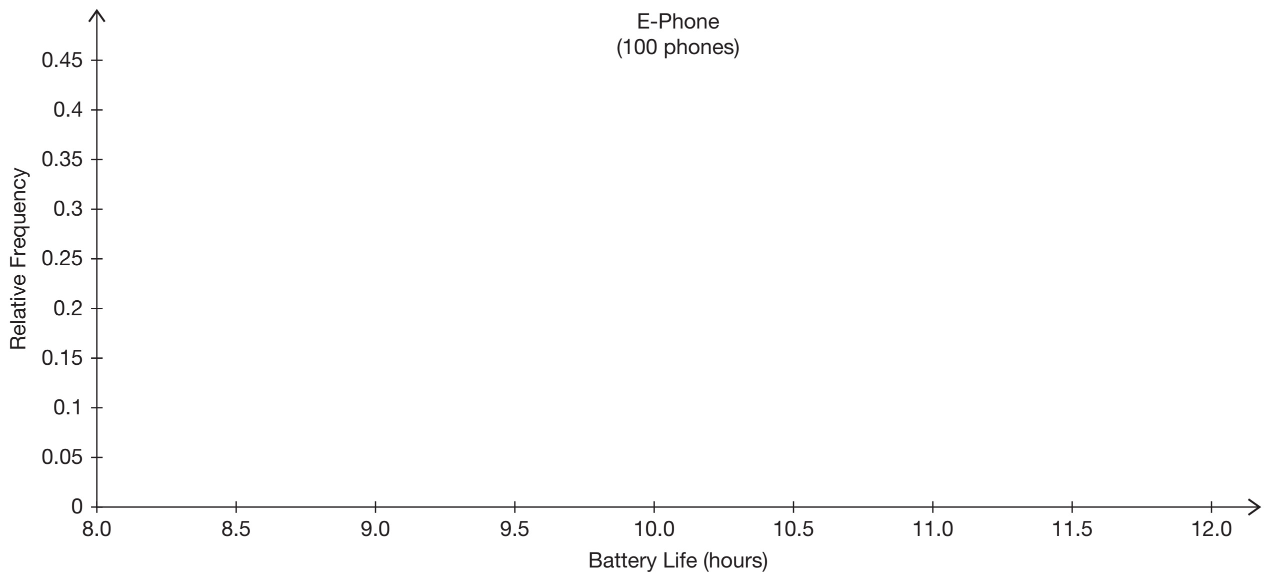
3. Complete the tables by calculating the relative frequency of phones in each interval. Explain how you determined the relative frequencies.

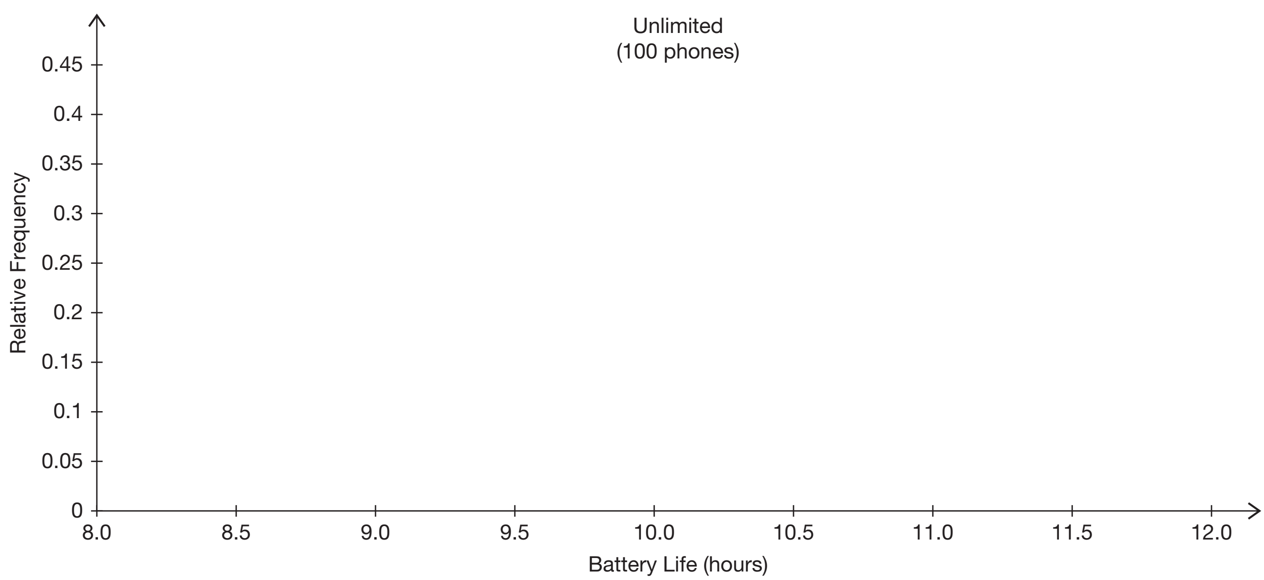
E-Phone		
Battery Life (hours)	**Number of Phones**	**Relative Frequency**
8.0–8.5	1	
8.5–9.0	2	
9.0–9.5	17	
9.5–10.0	30	
10.0–10.5	32	
10.5–11.0	15	
11.0–11.5	3	
11.5–12.0	0	

Unlimited		
Battery Life (hours)	**Number of Phones**	**Relative Frequency**
8.0–8.5	0	
8.5–9.0	1	
9.0–9.5	14	
9.5–10.0	37	
10.0–10.5	36	
10.5–11.0	11	
11.0–11.5	0	
11.5–12.0	1	

For continuous data, a relative frequency histogram displays continuous intervals on the horizontal axis and relative frequency on the vertical axis.

4. Create a relative frequency histogram to represent the battery lives of the 100 cell phones in each sample.

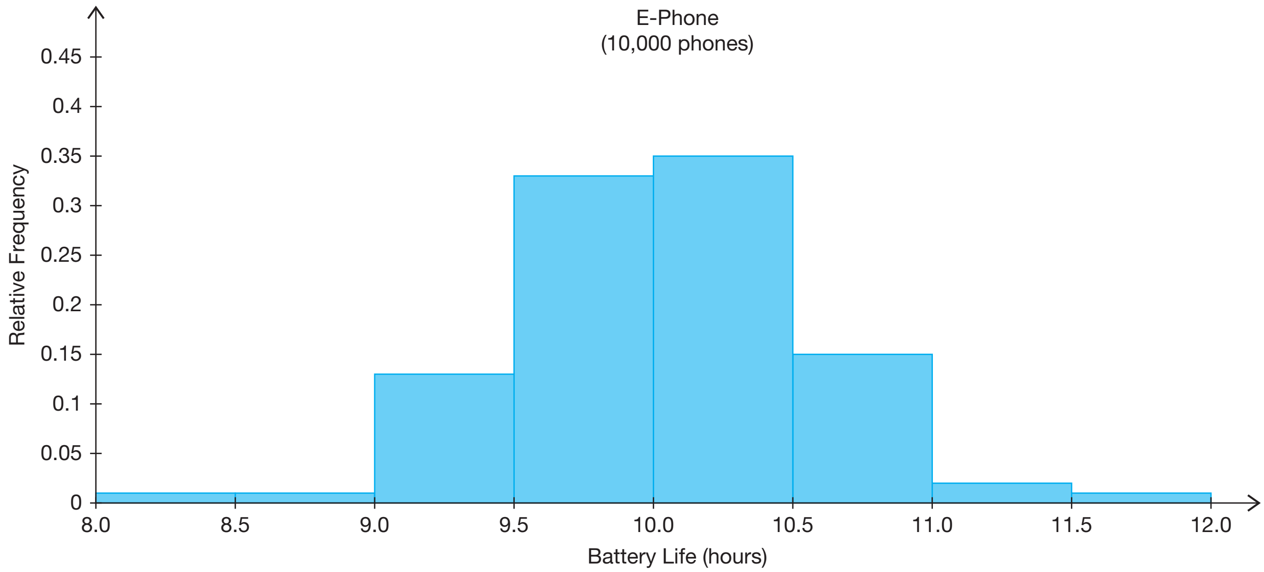


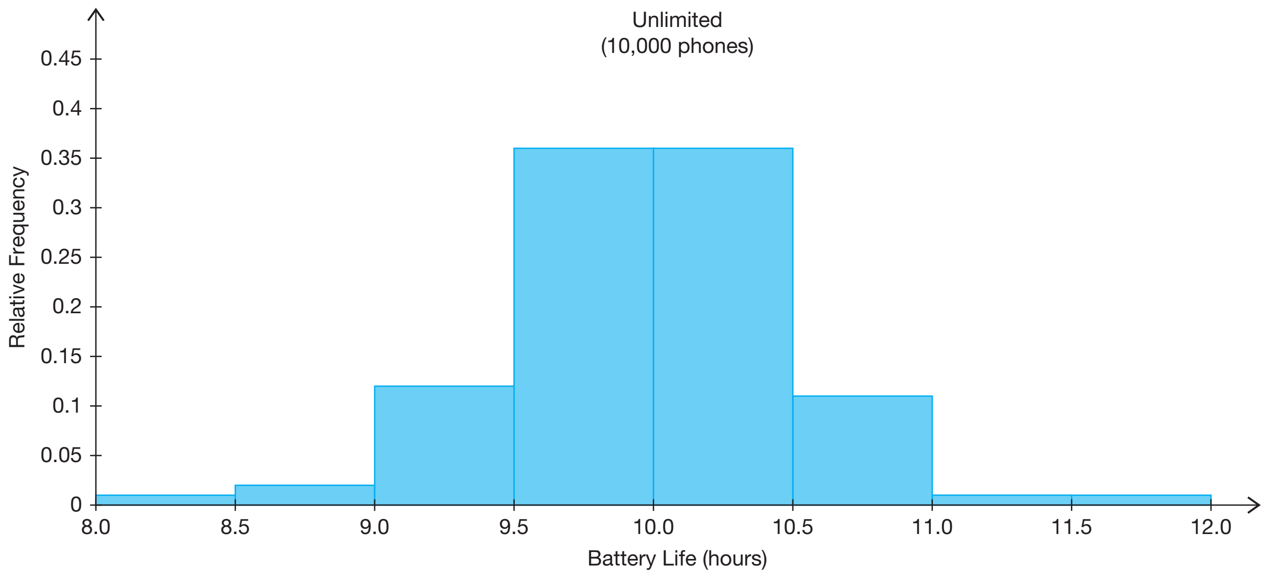


Recall that the shape of a data distribution can reveal information about the data. Data can be widely spread out or packed closer together. Data distributions can also be skewed or symmetric.

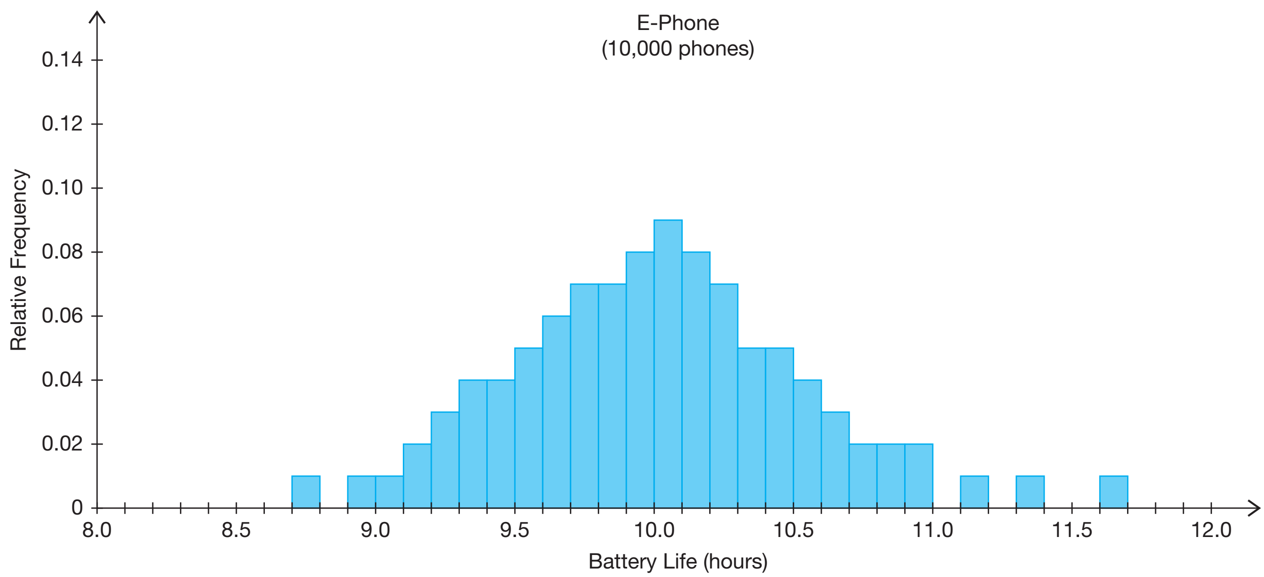
5. Describe the shape and spread of the histograms. What might these characteristics reveal about the data for each company?

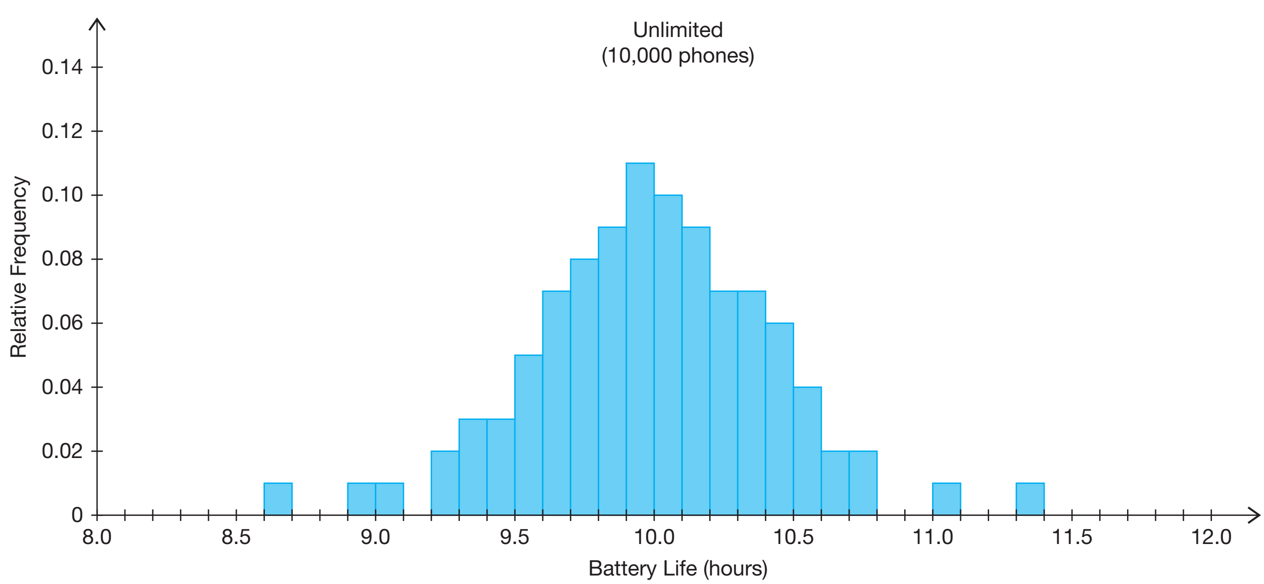
6. The relative frequency histograms shown represent samples of 10,000 phones from each of the two companies. Compare the histograms created from a sample of 10,000 cell phones to the histograms created from a sample of 100 cell phones. How does increasing the sample size change the appearance of the data distributions?





7. The histograms shown represent the same samples of 10,000 phones, but now the data have been divided into intervals of 0.1 hour instead of 0.5 hour. Compare these histograms with the histograms from the previous question. How does decreasing the interval size change the appearance of the data distributions?

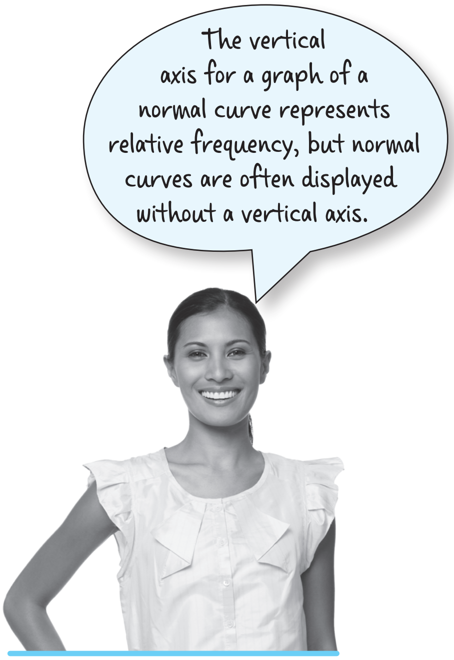




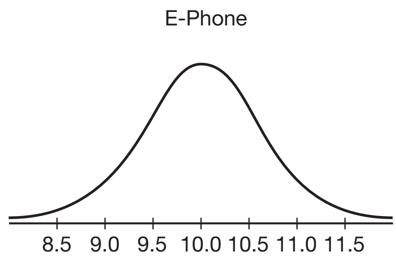
8. Explain why the scale of the *y*-axis changed when the interval size increased.

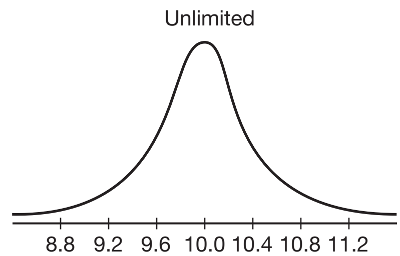
As the sample size continues to increase and the interval size continues to decrease, the shape of each relative frequency histogram will likely start to resemble a *normal curve*. A **normal curve** is a bell-shaped curve that is symmetric about the mean of the data.

A normal curve models a theoretical data set that is said to have a **normal distribution**.



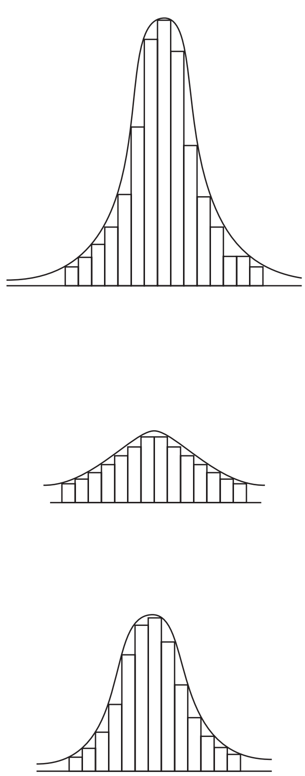
The normal curves for the E-Phone and Unlimited cell phone battery lives are shown. In order to display normal curves for each data set, different intervals were used on the horizontal axis in each graph.



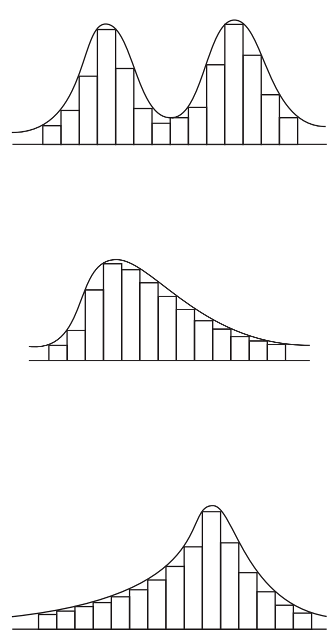


Although normal curves can be narrow or wide, all normal curves are symmetrical about the mean of the data.

Normal Distributions

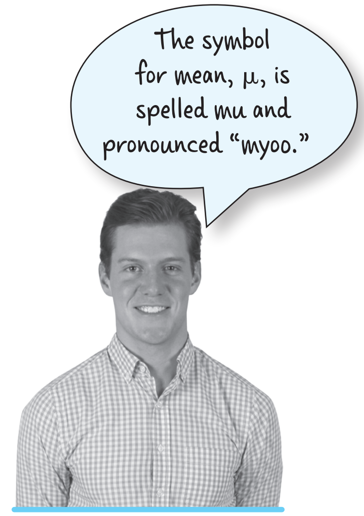


Not Normal Distributions



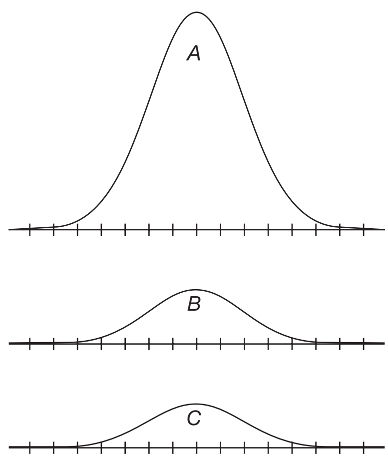
PROBLEM 2 Deviating Slightly

You already know a lot about the mean. With normal curves, the **mean** of a population is represented with the symbol μ. The mean of a sample is represented with the symbol $\bar{x}$. The **standard deviation** of data is a measure of how spread out the data are from the mean. The symbol used for the standard deviation of a population is the sigma symbol (σ). The standard deviation of a sample is represented with the variable *s*. When interpreting the standard deviation of data:



- A lower standard deviation represents data that are more tightly clustered near the mean.
- A higher standard deviation represents data that are more spread out from the mean.

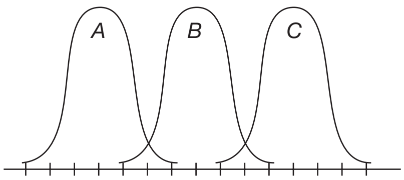
1. Normal curves *A*, *B*, and *C* represent the battery lives of a population of cell phones of comparable models from three different companies.



The normal curves represent distributions with standard deviations of $\sigma = 0.1$, $\sigma = 0.4$, and $\sigma = 0.5$. Match each standard deviation value with one of the normal curves and explain your reasoning.

1

2. Normal curves *A*, *B*, and *C* represent the battery lives of cell phones from three different companies.



a. Compare the mean of each company.

b. Compare the standard deviation of each distribution.

Be prepared to share your solutions and methods.

1

1.2

#I'mOnline

The Empirical Rule for Normal Distributions

LEARNING GOALS

In this lesson, you will:

- Recognize the connection between normal curves, relative frequency histograms, and the Empirical Rule for Normal Distributions.
- Use the Empirical Rule for Normal Distributions to determine the percent of data in a given interval.

KEY TERMS

- standard normal distribution
- Empirical Rule for Normal Distributions

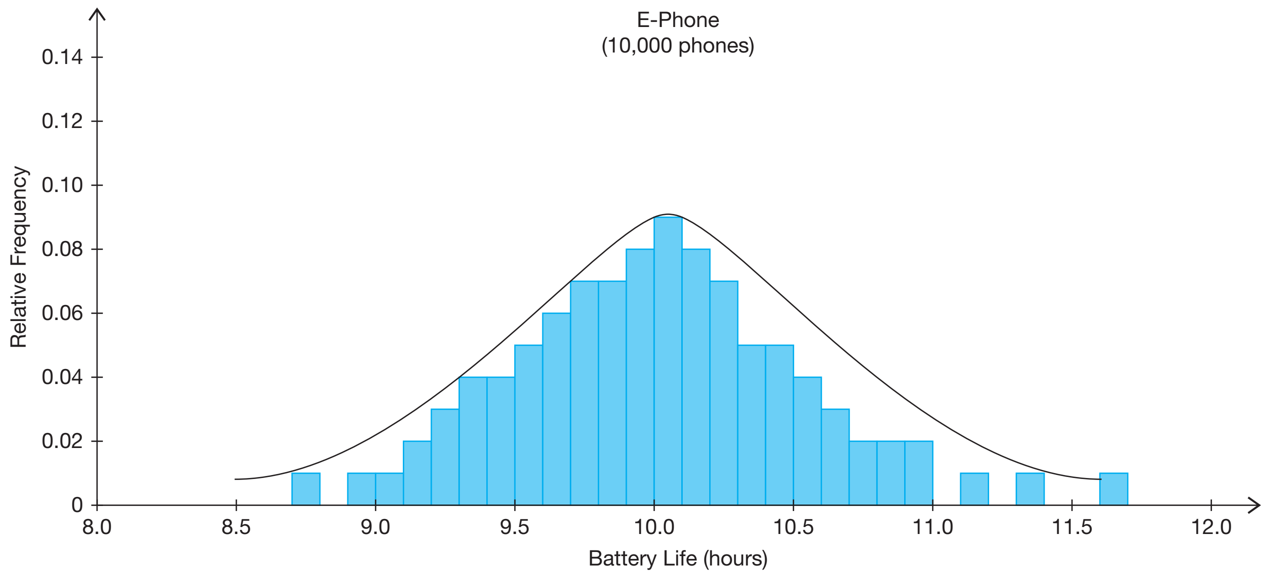
On October 19, 1987, stock markets around the world fell into sharp decline. In the United States, the Dow Jones Industrial Average dropped 508 points—a 22% loss in value. Black Monday, as the day came to be called, represented at the time the largest one-day decline in the stock market ever.

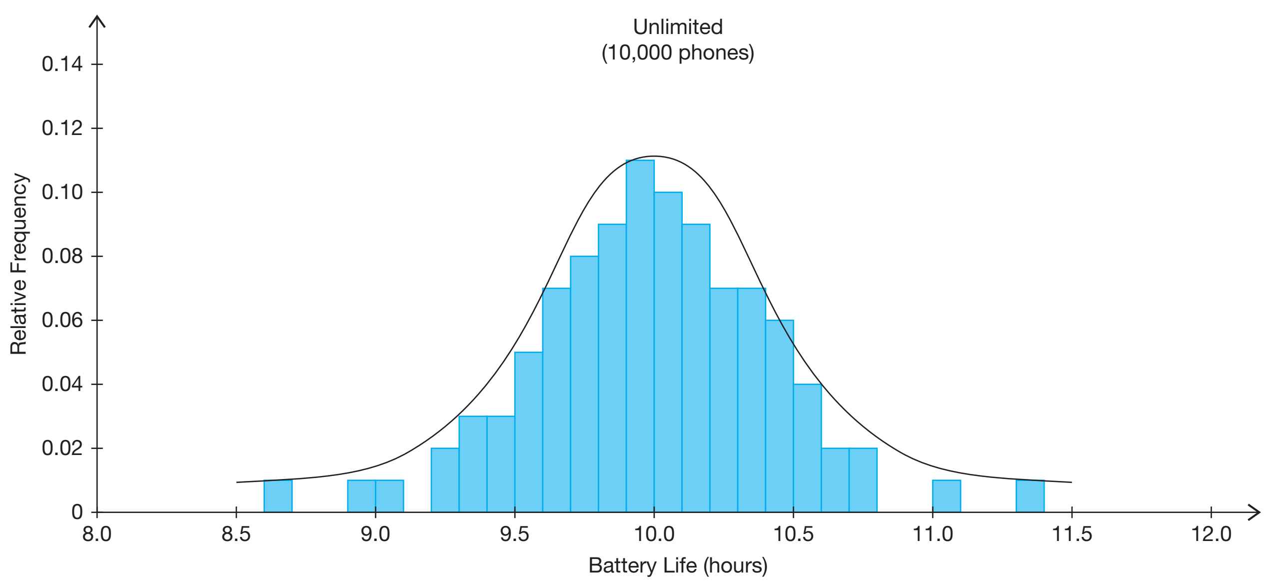
According to some economic models, the crash that occurred on Black Monday represented an event that was 20 standard deviations away from the normal behavior of the market. Mathematically, the odds of a Black Monday–type event occurring were 1 in 10^{50}.

PROBLEM 1 Count 'em Up

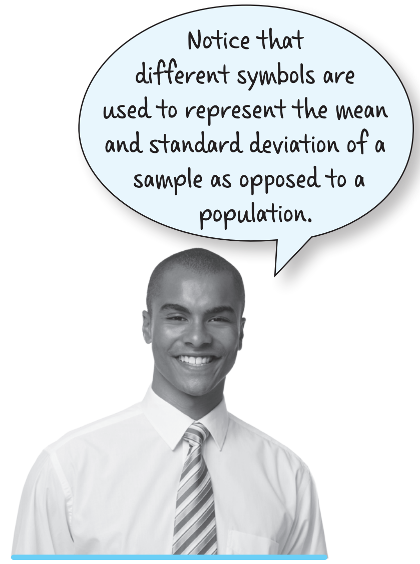
Let's investigate what the standard deviation can tell us about a normal distribution.

The relative frequency histograms for the battery lives of E-Phone and Unlimited cell phones are shown. The normal curves for each data set are mapped on top of the histogram.





Normal curves can be graphed with units of standard deviation on the horizontal axis. The normal curve for the E-Phone sample has a standard deviation of 0.5 hour ($s = 0.5$), and the normal curve for the Unlimited sample has a standard deviation of 0.4 hour ($s = 0.4$). The mean of each sample is $\bar{x} = 10.0$ hours.



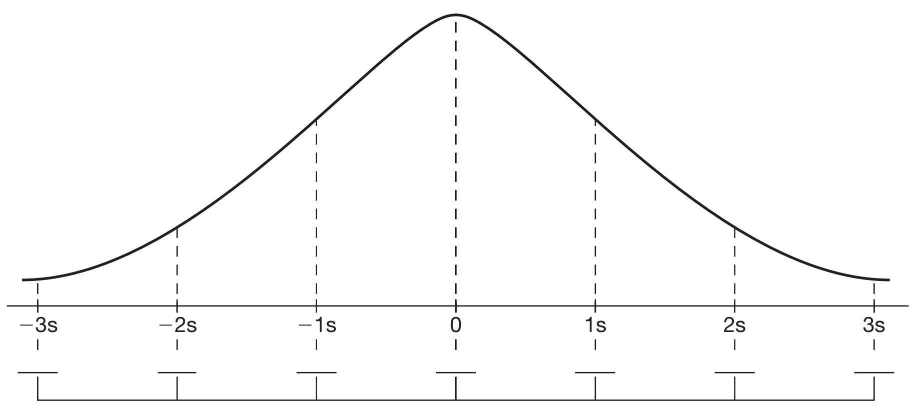
1. Study the graphs shown.

a. For each graph, label each standard deviation unit with its corresponding battery life.

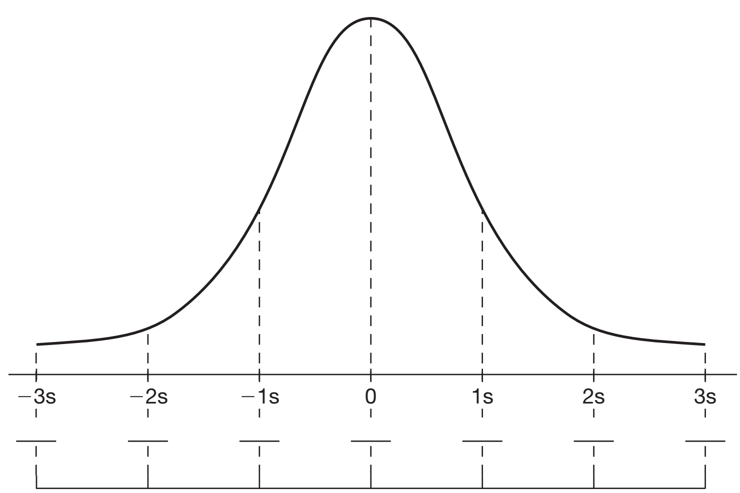
b. What value is represented at $s = 0$ for both graphs?

2. Use the histograms on the previous page to estimate the percent of data within each standard deviation. Write each percent in the appropriate space below the horizontal axis.

E-Phone



Unlimited



1

3. Compare the percents in each standard deviation interval for E-Phone with the percents in each standard deviation interval for Unlimited. What do you notice?

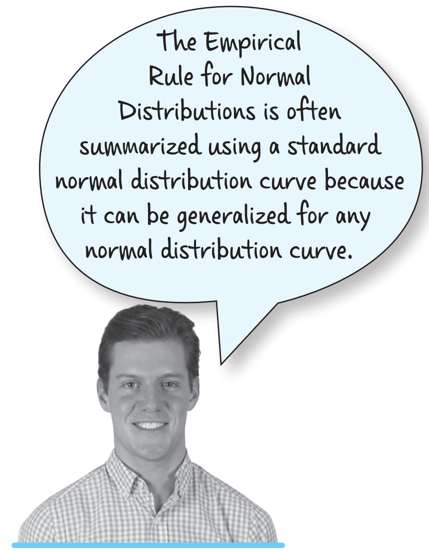
4. Use your results to answer each question. Explain your reasoning.

a. Estimate the percent of data within 1 standard deviation of the mean.

b. Estimate the percent of data within 2 standard deviations of the mean.

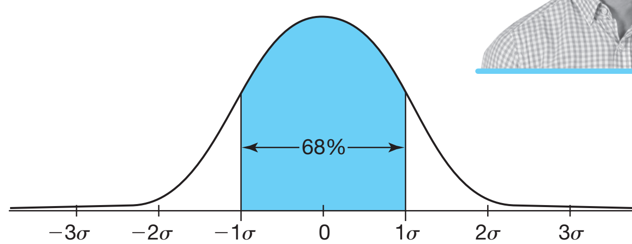
c. Estimate the percent of data within 3 standard deviations of the mean.

The **standard normal distribution** is a normal distribution with a mean value of 0 and a standard deviation of 1σ or $1s$. In a standard normal distribution, 0 represents the mean. Positive integers represent standard deviations greater than the mean. Negative integers represent standard deviations less than the mean.

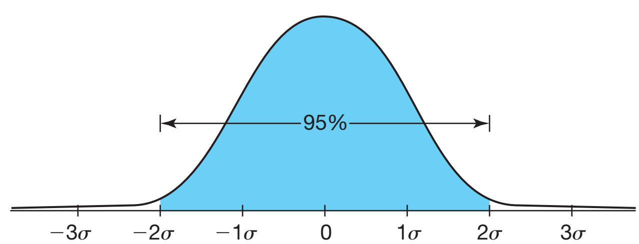


The **Empirical Rule for Normal Distributions** states:

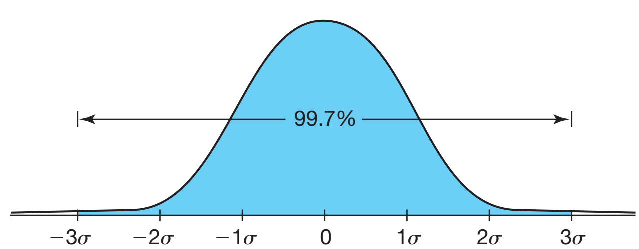
- Approximately 68% of the data in a normal distribution for a population is within 1 standard deviation of the mean.



- Approximately 95% of the data in a normal distribution for a population is within 2 standard deviations of the mean.

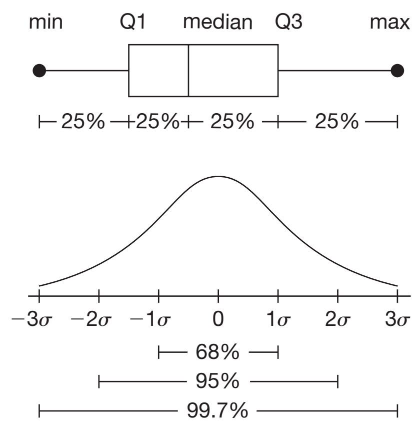


- Approximately 99.7% of the data in a normal distribution for a population is within 3 standard deviations of the mean.



The Empirical Rule applies most accurately to population data rather than sample data. However, the Empirical Rule is often applied to data in large samples.

Recall that a box-and-whisker plot is a graph that organizes, summarizes, and displays data based on quartiles that each contains 25% of the data values.



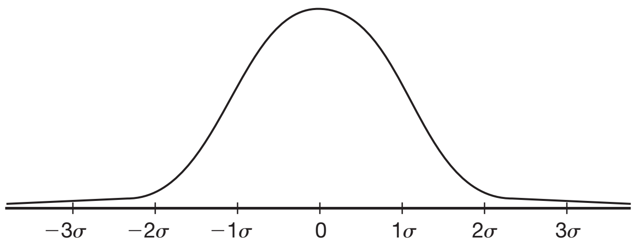
5. What similarities and/or differences do you notice about the box-and-whisker plot and the standard normal distribution?

PROBLEM 2 Interval Training

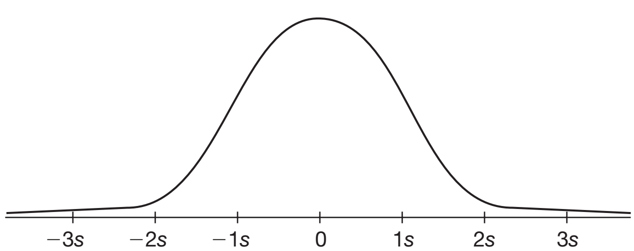
You can use the Empirical Rule for Normal Distributions to estimate the percent of data within specific intervals of a normal distribution.

1. Determine each percent and explain your reasoning. Shade the corresponding region under each normal curve. Then tell whether the distribution represents population data or sample data.

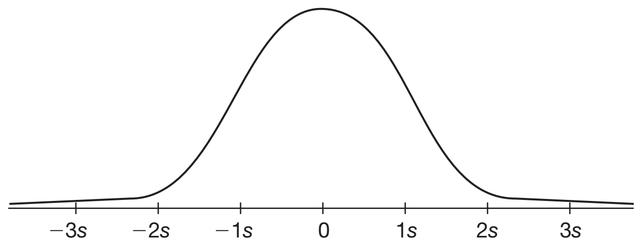
   a. What percent of the data is greater than the mean?



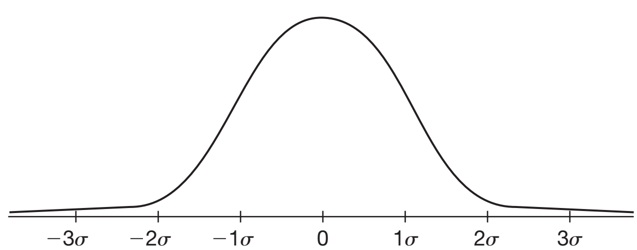
   b. What percent of the data is between the mean and 2 standard deviations below the mean?



c. What percent of the data is between 1 and 2 standard deviations above the mean?

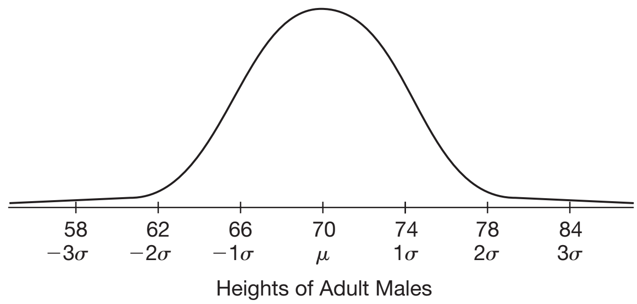


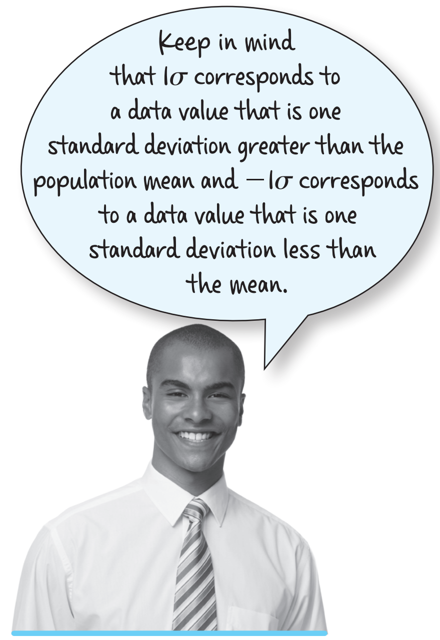
d. What percent of the data is more than 2 standard deviations below the mean?



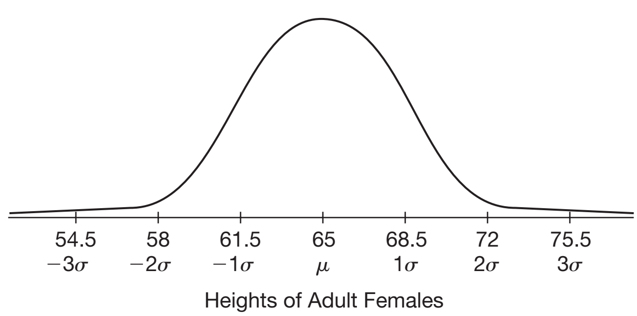
2. Use the normal curve to answer each question and explain your reasoning. Shade the region under each normal curve to represent your answer. Then tell whether the distribution represents population data or sample data.

   a. What percent of adult males have a height between 62 inches and 74 inches?



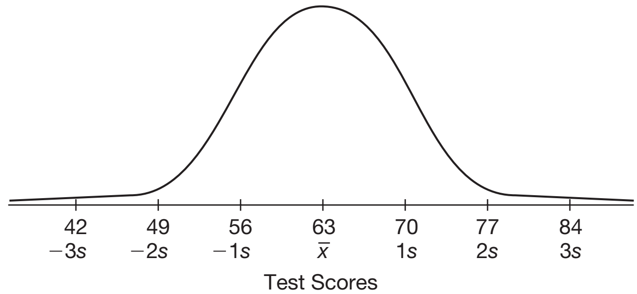


   b. What percent of adult females are taller than 68.5 inches?



1

c. What percent of history test scores are between 63 points and 70 points?



Be prepared to share your solutions and methods.