

Interpreting Data in Normal Distributions



This curve is kind of a big deal. It shows the distribution of a set of test scores, the results of rolling a die a million times, the heights of people on Earth, the battery life of cell phones, and the fuel efficiency of hybrid cars.



1.1 Recharge It!	
Normal Distributions	3
1.2 #I'mOnline	
The Empirical Rule for Normal Distributions	13
1.3 Below the Line, Above the Line, and Between the Lines	
Z-Scores and Percentiles	23
1.4 You Make the Call	
Normal Distributions and Probability	35

Chapter 1 Overview

The first lesson of this chapter leverages student knowledge of relative frequency histograms to introduce normal distributions. Students explore the characteristics of normal distributions. In the second lesson, students build their knowledge of normal distributions using the Empirical Rule for Normal Distributions. Students use the Empirical Rule for Normal Distributions to determine the percent of data between given intervals that are bounded by integer multiples of the standard deviation from the mean. In the third lesson, students use a z-score table and a graphing calculator to determine the percent of data in given intervals that are bounded by non-integer multiples of the standard deviation from the mean. In the last lesson, students use their knowledge of probability and normal distributions to analyze scenarios and make decisions.

Lesson		CCSS	Pacing	Highlights	Models	Worked Examples	Peer Analysis	Talk the Talk	Technology
1.1	Normal Distributions	S.ID.1 S.ID.2 S.ID.4	2	This lesson introduces normal distributions. In the context of a cell phone battery life scenario, questions guide students through building a normal distribution based on the concepts of continuous data and relative frequency histograms. Students explore and identify characteristics of normal distributions.	X				
1.2	The Empirical Rule for Normal Distributions	S.ID.1 S.ID.4	2	In this lesson, students use normal curves and relative frequency histograms of the cell phone battery scenario to build conceptual understanding of the Empirical Rule for Normal Distributions. The standard normal distribution is introduced. Students compare and contrast a box-and-whisker plot with a normal distribution. Using contextual and non-contextual scenarios, questions ask students to use the Empirical Rule for Normal Distributions to estimate the percentage of data in a given interval for normal distributions and for standard normal distributions.	X				

Lesson		CCSS	Pacing	Highlights	Models	Worked Examples	Peer Analysis	Talk the Talk	Technology
1.3	Z-Scores and Percentiles	S.ID.4	2	<p>Using the contexts of hybrid car gas mileage and teen texting, this lesson focuses on determining the percent of data below any given data value, above any given data value, and between any two data values of a normal distribution.</p> <p>Questions ask students to determine the percent of data in a given interval using a z-score table and using a graphing calculator. The lesson culminates with questions that ask students to determine data values that represent given percentiles, using a z-score table and a graphing calculator.</p>	X	X	X		X
1.4	Normal Distributions and Probability	S.MD.6(+) S.MD.7(+)	1	In this lesson, students will use their knowledge of normal distributions and probability to analyze situations and make informed decisions. The context of the scenarios are teen texting, pizza delivery times, and prize-winning tomatoes.	X				

Skills Practice Correlation for Chapter 1

1

Lesson		Problem Set	Objectives
1.1	Normal Distributions		Vocabulary
		1 – 6	Label data as either continuous or discrete
		7 – 12	Sketch a relative frequency histogram for a distribution and determine if it is a normal or non-normal distribution
		13 – 18	Identify the mean and standard deviation for normal distributions
		19 – 24	Draw a curve on the same axes as a given curve with new properties
1.2	The Empirical Rule for Normal Distributions		Vocabulary
		1 – 8	Shade the corresponding region under a standard normal curve
		9 – 16	Estimate the percent of data within specific intervals of a normal distribution, shade the corresponding region, and label the distribution
1.3	Z-Scores and Percentiles		Vocabulary
		1 – 6	Calculate a percent using a z-score table given a normal distribution
		7 – 12	Determine a percent using a graphing calculator given a normal distribution
		13 – 18	Calculate a percentile using a z-score table given a normal distribution
		19 – 24	Determine a percentile using a graphing calculator given a normal distribution
1.4	Normal Distributions and Probability	1 – 6	Calculate probabilities given a normal distribution
		7 – 12	Calculate probabilities given a mean and a standard deviation
		13 – 18	Answer questions given a mean and a standard deviation

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.

ESSENTIAL IDEAS

- A discrete graph is a graph of isolated points.
- A continuous graph is a graph of points that are connected by a line or a smooth curve on the graph.
- Discrete data is data whose possible values are countable and often finite.
- Continuous data are data which can take any numerical value within a range.
- 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
- Relative frequency is a ratio of occurrences within a category to the total number of occurrences.
- A relative frequency histogram displays data values on one axis and the relative frequency on the other axis.
- A normal curve is a bell shaped curve that is symmetric about the mean of the data set.
- A data set that can be modeled using a normal curve has a normal distribution.

KEY TERMS

- discrete data
 - continuous data
 - sample
 - population
 - normal curve
 - normal distribution
 - mean (μ)
 - standard deviation (σ)
- The mean is the number along the number line that is aligned with the peak of the curve. Each tick mark to the left and right of the mean represents one standard deviation from the mean.

COMMON CORE STATE STANDARDS FOR MATHEMATICS

S.ID Interpreting Categorical and Quantitative Data

Summarize, represent, and interpret data on a single count or measurement variable

1. Represent data with plots on the real number line (dot plots, histograms, and box plots).
2. Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.

4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

Overview

Data sets are given and students calculate the relative frequency which is used to create relative frequency histograms within the context of a situation. They explore how the relative frequency histograms begin to resemble a bell-shaped curve when the sample size increases and the interval size of the bars decrease. The terms mean, normal curve, normal distribution and standard deviation are defined. Sets of normal curves are given and students identify the mean and standard deviation related to each curve.

Warm Up

1

Determine whether each situation is best modeled by a discrete graph or a continuous graph.

1. The temperature of an oven when using it to cook food.

The data is best modeled by a continuous graph because the oven temperature can be any numerical value between zero and the highest temperature used to cook the food.

2. The test scores on the math test in Ms. Brown's class.

The data is best modeled by a discrete graph because test scores are commonly reported as whole numbers.

3. The amount of water in a bathtub as the tub fills, is used, and then drains.

The data is best modeled by a continuous graph because the amount of water can be any numerical value between zero and the maximum amount of water contained by the tub.

4. The number of cookies in a box of cookies.

The data is best modeled by a discrete graph because the number of cookies is a whole number.

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.

Problem 1

Battery life data of cell phones from two different companies are given. Students calculate relative frequencies and create relative frequency histograms for each phone company. They compare the spread and shape of each histogram and explore how the relative frequency histograms change when the sample size increases and when the interval size of the histogram bars decrease. Increasing the sample size and decreasing the interval size of the histogram bars transforms the data distribution into a bell-shaped, symmetrical curve.

Grouping

- Ask a student to read the information and definitions. Discuss as a class.
- Have students complete Questions 1 and 2 with a partner. Then have students share their responses as a class.
- Ask a student to read the information and definitions before Question 3. Discuss as a class.
- Have students complete Questions 3 through 5 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Questions 1 and 2

- If the average battery life is 10 hours, can some phones have a battery life of 9 hours?

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.

The durations are examples of continuous data. A possible value for a battery life can be a fraction of an amount of time. These values are not countable.

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.

No. It is possible but unlikely for all E-Phone phones and all Unlimited phones to have a 10-hour battery life. It is more likely that some phones will have a battery life greater than 10 hours, some phones will have a battery life equal to 10 hours, and some phones will have a battery life less than 10 hours.

- If the average battery life is 10 hours, can some phones have a battery life of 11 hours?
- For every cell phone that has a battery life of 8 hours, it is likely that there is another cell phone that has a battery life of how many hours?
- For every cell phone that has a battery life of 12 hours, it is likely that there is another cell phone that has a battery life of how many hours?

Guiding Questions for Share Phase, Question 3

- How is the relative frequency determined for each interval of time?
- What is the total number of occurrences in this situation?
- What is the sum of all of the relative frequencies for E-Phone?
- What is the sum of all of the relative frequencies for Unlimited?
- Why is the sum of the relative frequencies for each company equal to 1, or 100%? Explain.
- Is it possible for the sum of the relative frequencies to equal a value other than 1, or 100%? Explain.

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.

Recall that relative frequency is the ratio of occurrences within an interval to the total number of occurrences.



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	0.01
8.5–9.0	2	0.02
9.0–9.5	17	0.17
9.5–10.0	30	0.30
10.0–10.5	32	0.32
10.5–11.0	15	0.15
11.0–11.5	3	0.03
11.5–12.0	0	0.00

Unlimited		
Battery Life (hours)	Number of Phones	Relative Frequency
8.0–8.5	0	0.00
8.5–9.0	1	0.01
9.0–9.5	14	0.14
9.5–10.0	37	0.37
10.0–10.5	36	0.36
10.5–11.0	11	0.11
11.0–11.5	0	0.00
11.5–12.0	1	0.01

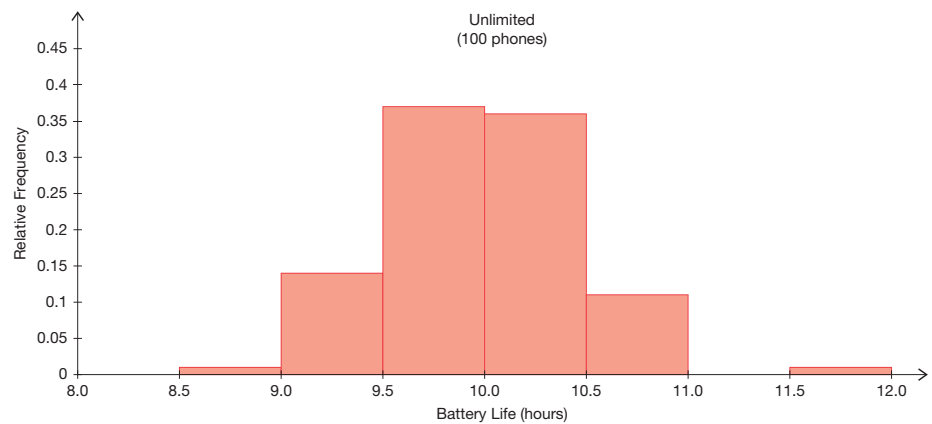
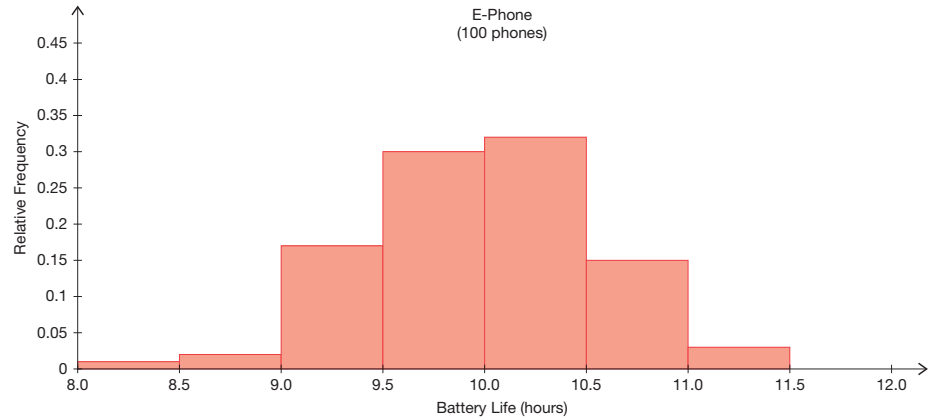
I divided the number of phones in each interval by the number of phones in the sample, 100, to determine each relative frequency.

Guiding Questions for Share Phase, Question 4

- How was the information in the tables used to create the graphs?
- What is the difference between a histogram and a bar graph?
- How did you determine which data values to place on the x -axis of the histogram?

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.

Guiding Questions for Share Phase, Question 5

- Do both histograms have a peak near the middle?
- What is the mean value on the histograms?
- Do both histograms have a peak near the mean of 10.0 hours?
- Which histogram appears a little more spread out?
- What histogram seems taller in the center?

Grouping

Have students complete Questions 6 through 8 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Question 6

- Which histograms are more symmetrical, the graph created using a sample space of 100 cell phones, or the graph created using a sample size of 10,000 cell phones?
- Which histograms have a more definite center, the graph created using a sample space of 100 cell phones, or the graph created using a sample size of 10,000 cell phones?
- As the sample size increases, does the shape of the relative frequency histogram change?



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

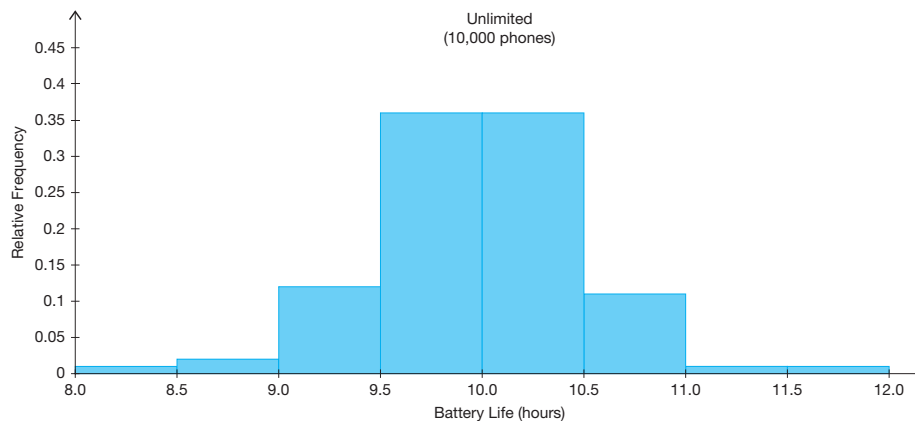
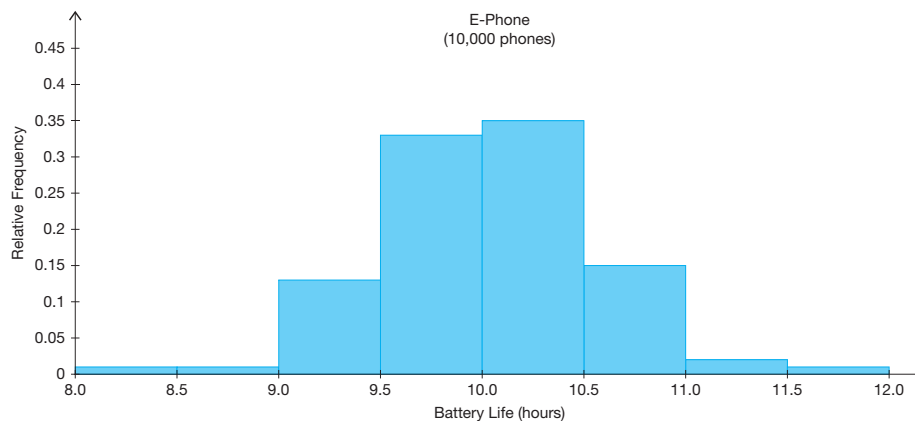
Answers will vary.

Both histograms have a peak near the middle, near the mean of 10.0 hours. E-Phone's histogram is a little more spread out, and Unlimited's histogram is a little taller in the center.

E-Phone's wider distribution of data indicates that its cell phones' battery lives are less consistent with the mean battery life of 10 hours than are the battery lives of Unlimited's phones. Both data distributions appear to be close to symmetric.



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?



Answers will vary.

The histograms for the 10,000 phones show more of the data closer to the mean of 10 hours than the histograms with 100 phones.

- As the sample size increases, does the shape of the relative frequency histogram approach a bell-shaped curve?
- As the sample size increases, does the shape of the relative frequency histogram approach a symmetrical distribution?

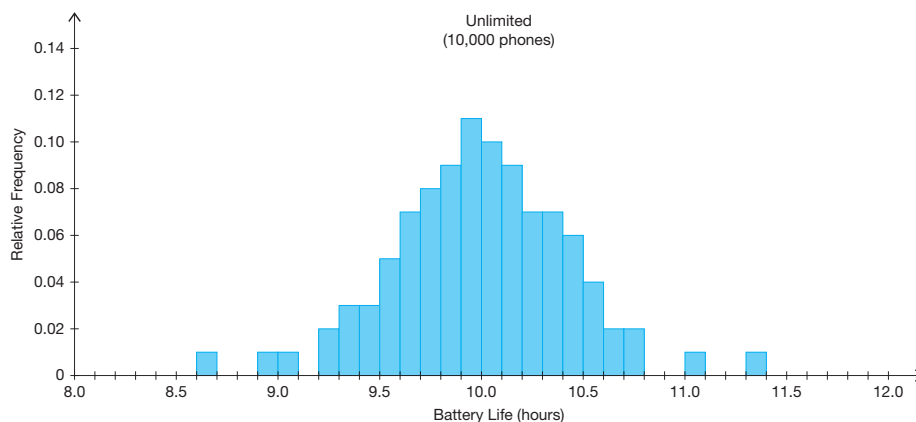
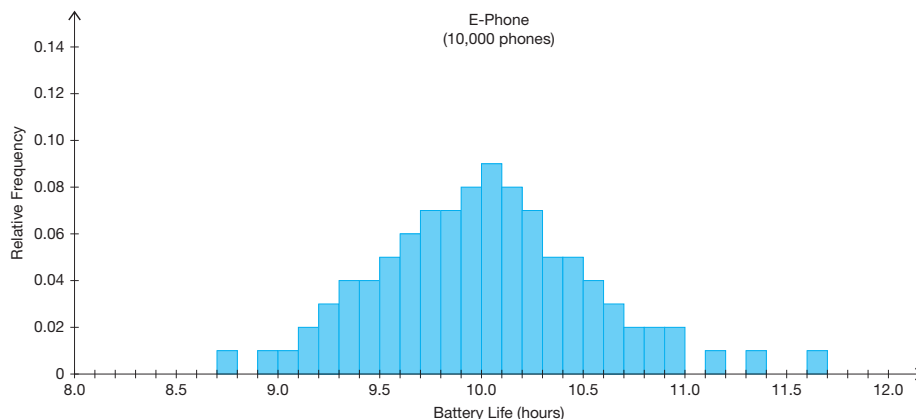
Guiding Questions for Share Phase, Questions 7 and 8

- How do the intervals change?
- Do the number of bars in the histogram increase?
- Is the path across the top of the bars smoother?
- As the bars get thinner, will the top of the distribution get closer?
- Will the top of the bars eventually form a smooth curve if the bars are extremely thin?
- What information on the histogram provides information about the mean?
- What information on the histogram provides information about the standard deviation?
- Where on the graph is the mean located?
- Which cell phone has a mean battery life of 10 hours?
- Which cell phone has a standard deviation of 0.5 hours?
- Which cell phone has a standard deviation of 0.4 hours?
- Is the data distribution bell-shaped?
- Is the data distribution symmetrical?

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?

Answers will vary.

There are more bars in each histogram, which gives a more distinct path across the top of the bars.



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

For each histogram, there are more data in intervals of larger size, so when the interval size decreased, the amount of data in each interval decreased as well.

Grouping

Ask a student to read the information and definitions. Discuss as a class.

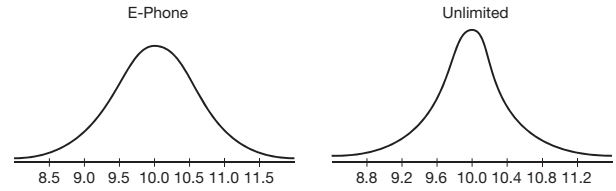


As the sample size continues to increase and the interval size continues to decrease, the shape of each relative frequency histogram will likely start to resemble a *normal curve*. A **normal curve** is a bell-shaped curve that is symmetric about the mean of the data.

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

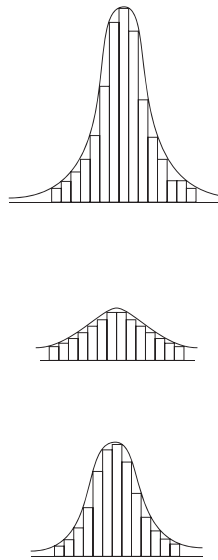
The normal curves for the E-Phone and Unlimited cell phone battery lives are shown. In order to display normal curves for each data set, different intervals were used on the horizontal axis in each graph.

The vertical axis for a graph of a normal curve represents relative frequency, but normal curves are often displayed without a vertical axis.

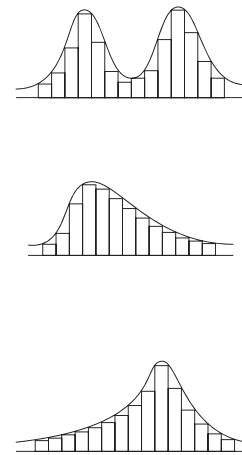


Although normal curves can be narrow or wide, all normal curves are symmetrical about the mean of the data.

Normal Distributions



Not Normal Distributions



Problem 2

Two sets of normal curves are given and students compare the means, and standard deviations in each set. They also discuss what could account for differences in the graphs.

Grouping

- Ask a student to read the information and definitions. Discuss as a class.
- Have students complete Questions 1 and 2 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Question 1

- Is the mean value the same or different for all three companies?
- If all three companies have a common peak value, does that result in the same mean value?
- Does a higher curve result in a larger mean value?
- Is the mean value related to the height of the curve?
- How is the mean value related to the height of the curve?
- Is the standard deviation the same or different for all three companies?
- Does a higher curve result in a greater or smaller standard deviation?
- Is the standard deviation related to the height of the curve?

PROBLEM 2 Deviating Slightly



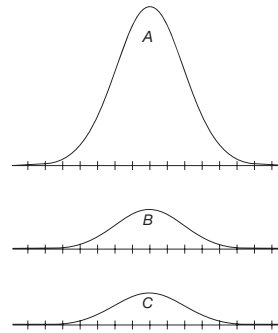
You already know a lot about the mean. With normal curves, the **mean** of a population is represented with the symbol μ . The mean of a sample is represented with the symbol \bar{x} . The **standard deviation** of data is a measure of how spread out the data are from the mean. The symbol used for the 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.

The symbol for mean, μ , is spelled mu and pronounced "myoo."



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.

For Normal Curve A, $\sigma = 0.1$. The data are the least spread out from the mean in this normal curve, so the least standard deviation value matches this curve.

For Normal Curve C, $\sigma = 0.5$. The data are the most spread out from the mean in this normal curve, so the greatest standard deviation value matches this curve.

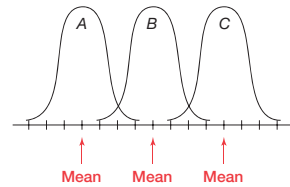
For Normal Curve B, $\sigma = 0.4$. The data are spread out more than in Graph A and less than in Graph C, so the middle value for the standard deviation matches this curve.

- How is the standard deviation related to the height of the curve?
- Which company is associated with a battery life that is more consistent and closer to the mean?
- Why does Company A appear more efficient?
- Which company is associated with a battery life that is most inconsistent and varies the most from the mean?
- Why does Company C appear less efficient?

Guiding Questions for Share Phase, Question 2

- If all three companies have different centers, does that result in different mean values?
- Why does Company A have the smallest mean?
- Why does Company C have the greatest mean?
- Is the shape of the normal curves the same?
- Is the shape of the normal curve related to its standard deviation?
- How is the shape of the normal curve related to its standard deviation?
- Which company could represent a company that uses older battery technology?
- Which company is associated with a shorter mean battery life?
- Which company could represent a company that uses newer battery technology?
- Which company is associated with a higher mean battery life?

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.

Company A has the smallest mean.

Company C has the greatest mean.

Company B's mean is between the mean of Companies A and C.

- b. Compare the standard deviation of each distribution.

The standard deviation is the same for all three companies because the shape of the normal curves is the same.



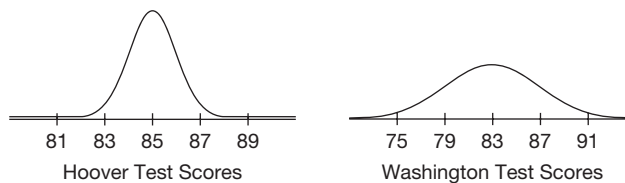
Be prepared to share your solutions and methods.

Check for Students' Understanding

The standardized test scores at Washington Elementary School produced a bell-shaped curve. The mean score was 83 and the standard deviation was 4 points.

The standardized test scores at Hoover Elementary School produces a bell-shaped curve. The mean score was 85 and the standard deviation was 2 points.

Analyze both graphs and describe their similarities and differences.



Answers will vary.

Hoover's test scores form a narrower and taller curve. Washington's test scores form a flatter and shorter curve.

#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.

ESSENTIAL IDEAS

- The standard normal distribution is a normal distribution with a mean value of 0 and a standard deviation of 1.
- Positive integers represent standard deviations greater than the mean and negative integers represent standard deviations less than the mean.
- The Empirical Rule of Normal Distributions states:
 - Approximately 68% of the data in a normal distribution is within one standard deviation of the mean.
 - Approximately 95% of the data in a normal distribution is within two standard deviations of the mean.
 - Approximately 99.7% of the data in a normal distribution is within three standard deviations of the mean.
- The Empirical Rule of Normal Distributions can be used to estimate the percent of data within specific intervals of a normal distribution.

KEY TERMS

- standard normal distribution
- Empirical Rule for Normal Distributions

COMMON CORE STATE STANDARDS FOR MATHEMATICS

S.ID Interpreting Categorical and Quantitative Data

Summarize, represent, and interpret data on a single count or measurement variable

1. Represent data with plots on the real number line (dot plots, histograms, and box plots).
4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

Overview

Two histograms from the previous lesson are used again. Students estimate the percent of data within each standard deviation on the histograms, which leads to the Empirical Rule for Normal Distributions. The Empirical Rule for Normal Distributions is depicted with an illustration of a normal curve. In different problem situations, students use the Empirical Rule for Normal Distributions to estimate the percent of data within specific intervals.

Warm Up

1

1. Describe the similarities and differences between two normal curves that have the same mean but different standard deviations.

The center of each normal curve will have the same value because their means are equal.

The distribution with the smaller standard deviation will form a narrower and taller curve.

The distribution with the larger standard deviation will form a flatter and shorter curve.

2. Describe the similarities and differences between two normal curves that have the same standard deviation but different means.

The center of each normal curve will have a different value because their means are different.

The shape of both curves will be the same because their standard deviations are equal.

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

Two histograms from the previous lesson that display cell phone battery life from two companies are given. Students use the histograms to estimate the percent of data within each standard deviation. They conclude that similar percents of data lie within intervals based on the same number of standard deviations. The histograms serve as an introduction and connection to the Empirical Rule for Normal Distributions. The term standard normal distribution is defined.

Grouping

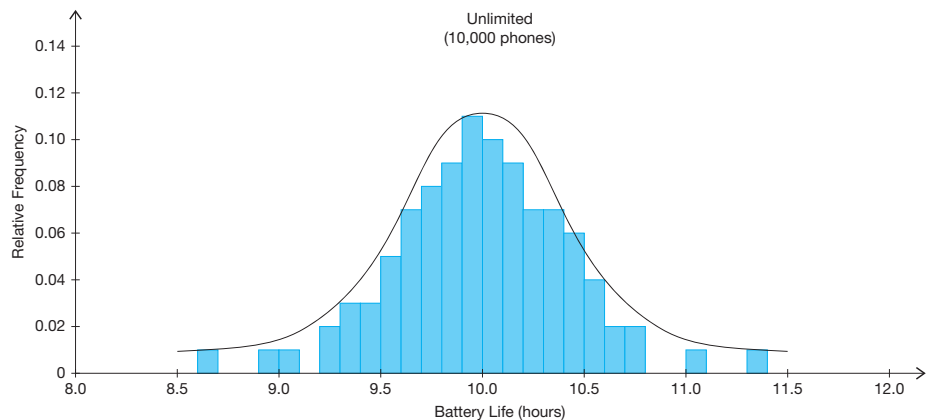
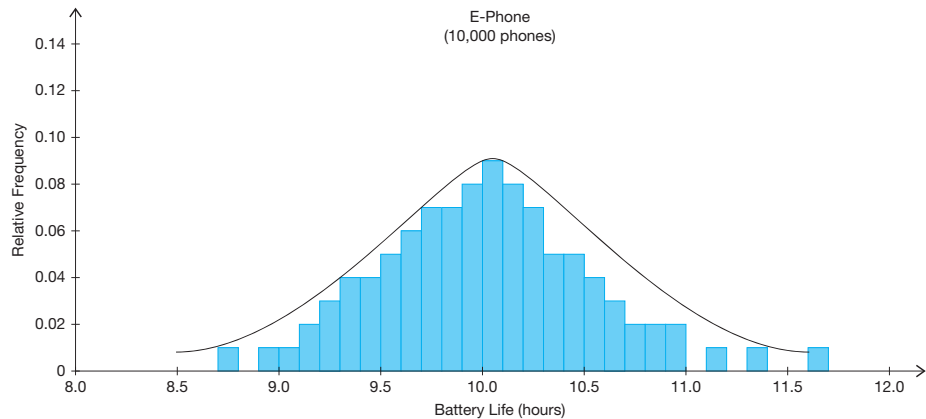
- Ask a student to read the information. Discuss as a class.
- Have students complete Questions 1 through 4 with a partner. Then have students share their responses as a class.

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.



Guiding Questions for Share Phase, Questions 1 and 2

- How are the histograms used to estimate the percent of data within each standard deviation?
- Are the percents in each standard deviation interval for both companies similar?
- Do you notice a pattern for the percents in each interval?



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.

See graphs.

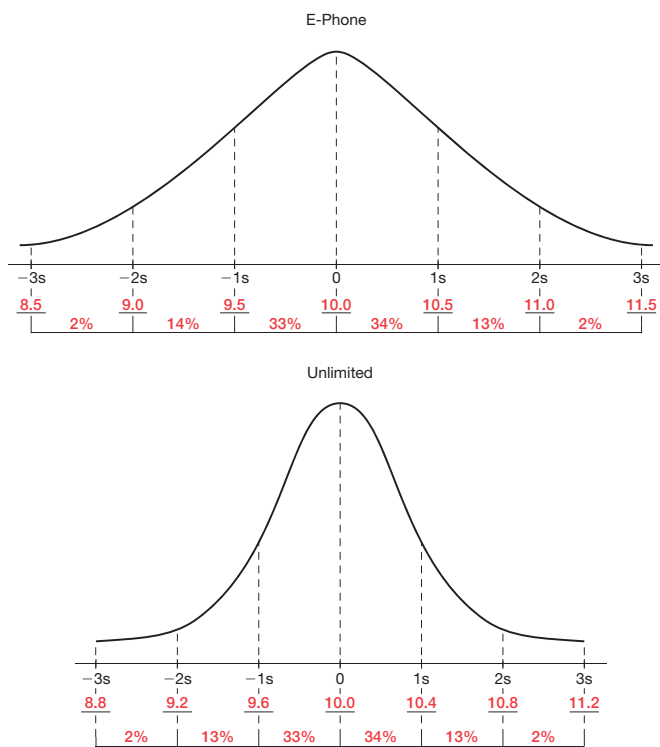
- b. What value is represented at $s = 0$ for both graphs?

The value of $s = 0$ represents the mean of the data, which is 10.0 for each data set.

Notice that different symbols are used to represent the mean and standard deviation of a sample as opposed to a population.



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.



Guiding Questions for Share Phase, Questions 3 and 4

- Which percents are added to determine the data within one standard deviation of the mean?
- Is the sum of the percents from the middle two intervals within one standard deviation of the mean?
- Which percents are added to determine the data within two standard deviations of the mean?
- Is the sum of the percents from the four middle intervals within two standard deviations of the mean?
- Which percents are added to determine the data within three standard deviations of the mean?
- Is the sum of the percents from the six intervals within three standard deviations of the mean?

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?

Answers will vary.

The percents in each standard deviation interval for E-Phone are very similar to the percents in each standard deviation interval for Unlimited.

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

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

Answers will vary.

Approximately 67% of the data is within 1 standard deviation of the mean. I determined the answer by adding the percents from the two middle intervals.

"Within one standard deviation" means between -1σ and 1σ , or between -1σ and 1σ .



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

Answers will vary.

Approximately 94% of the data is within 2 standard deviations of the mean. I determined the answer by adding the percents from the four middle intervals.



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

Answers will vary.

Approximately 98% of the data is within 3 standard deviations of the mean. I determined the answer by adding the percents from the six intervals.

Grouping

- Ask a student to read the information and definitions. Discuss as a class.
- Have students complete Question 5 with a partner. Then have students share their responses as a class.

Guiding Questions for Discuss Phase

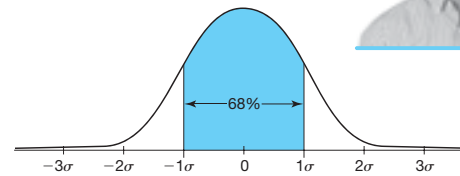
- What values are used to label the horizontal axis of a normal distribution?
- What values are used to label the horizontal axis of a standard normal distribution?
- What is the difference between a normal distribution and a standard normal distribution?
- What is similar between a normal distribution and a standard normal distribution?



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:

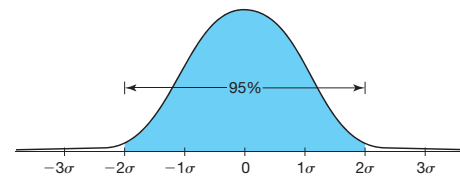
- Approximately 68% of the data in a normal distribution for a population is within 1 standard deviation of the mean.



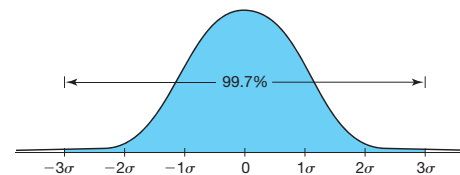
The Empirical Rule for Normal Distributions is often summarized using a standard normal distribution curve because it can be generalized for any normal distribution curve.



- 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.



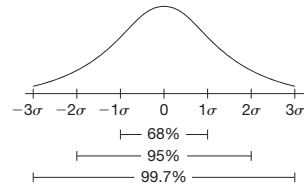
The Empirical Rule applies most accurately to population data rather than sample data. However, the Empirical Rule is often applied to data in large samples.

Guiding Questions for Share Phase, Question 5

- On a box-and-whisker plot, what do Q1 and Q3 represent?
- On a box-and-whisker plot, how do the minimum, Q1, median, Q3, and maximum values divide up the data?
- What measure of center is used for a box-and-whisker plot?
- What measure of center is used for a standard normal curve?
- How can you determine a measure of spread for a data set that is displayed with a box-and-whisker plot?
- How can you determine a measure of spread for a data set that is displayed with a standard normal curve?



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?

Answers will vary.

Both models have a constant percent of data for a particular region. A box-and-whisker plot displays the data in quartiles. Each quartile contains approximately 25% of the data values. A standard normal distribution displays the data according to the Empirical Rule for Normal Distributions.

The center for a box-and-whisker plot is represented by the median, but the center for a standard normal distribution is represented by the mean.

A box-and-whisker plot does not have to be symmetric and balanced, but a standard normal curve is bell-shaped and symmetric.



Problem 2

Students use the Empirical Rule for Normal Distributions to estimate the percent of data within specific intervals of a normal distribution. Situations such as history test scores, the height of adult males and the height of adult females are used. Students determine the percent of data greater than the mean, between the mean and two standard deviations below the mean, between one and two standard deviations above the mean, and the percent of data that is more than two standard deviations above the mean.

Grouping

Have students complete Questions 1 and 2 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Question 1, parts (a) and (b)

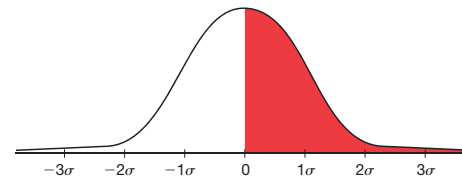
- Where on the graph is the mean?
- Is the mean a measure of center?
- Where on the graph is the data greater than the mean?
- Is 50% of the data greater than the mean?
- What percent of the data is less than the mean?
- Why is half of the 68% between the mean and one standard deviation above the mean?

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?

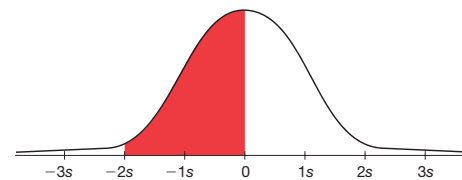


Approximately 50% of the data is greater than the mean.

The mean is a measure of center. Fifty percent of the data is greater than the mean and 50% of the data is less than the mean.

The sigma values indicate that this distribution represents data from a population.

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



Approximately 47.5% of the data is between the mean and 2 standard deviations below the mean.

I know that 95% of the data are within 2 standard deviations of the mean and the normal curve is symmetric. So, half of the 95% are between the mean and 2 standard deviations below the mean.

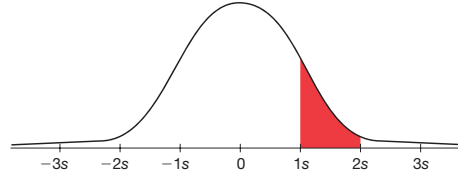
The s-values indicate that this distribution represents data from a sample.

- Why is half of the 68% between the mean and one standard deviation below the mean?
- What is half of 68%?
- Why is half of the 95% between the mean and two standard deviations above the mean?
- Why is half of the 95% between the mean and two standard deviations below the mean?
- What is half of 95%?

Guiding Questions for Share Phase, Question 1, parts (c) and (d)

- How can you use the Empirical Rule for Normal Distributions to answer these questions?
- What percent of data are within two standard deviations of the mean?
- What percent of the data is within the mean and one standard deviation above the mean?
- What percent of the data is within the mean and two standard deviations above the mean?
- Which interval contains approximately 34% of the data?
- Which interval contains approximately 47.5% of the data?
- What is the difference between the symbol s and the symbol σ ?

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

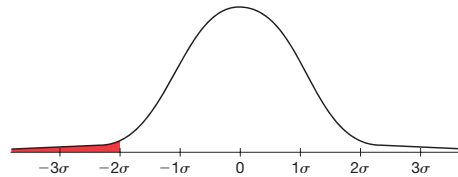


Approximately 13.5% of the data is between 1 and 2 standard deviations above the mean.

I know that 34% of data is within the mean and 1 standard deviation above the mean. I also know that 47.5% of the data is within the mean and 2 standard deviations above the mean. So, $47.5 - 34$, or 13.5% of the data, is between 1 and 2 standard deviations above the mean.

The s -values indicate that this distribution represents data from a sample.

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



Approximately 2.5% of the data is more than two standard deviations below the mean.

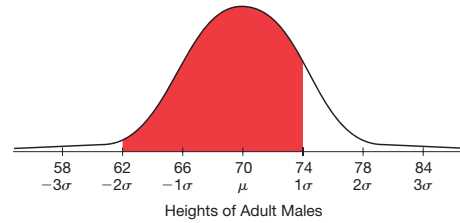
I know that 50% of the data is below the mean. I also know that 47.5% of the data is between the mean and 2 standard deviations below the mean. So, $50 - 47.5$, or 2.5% of the data is more than 2 standard deviations below the mean.

The sigma values indicate that this distribution represents data from a population.

Guiding Questions for Share Phase, Question 2, parts (a) and (b)

- How is the percent of data between one and two standard deviations above the mean used to determine the percent of adult males that have a height between 62 and 74 inches?
- How is the percent of data within two standard deviations of the mean used to determine the percent of adult males that have a height between 62 and 74 inches?
- Is the percent of adult females taller than 68.5 inches described by the data above one or two standard deviations?
- Are 16% of adult female heights taller than 68.5 inches?
- How did you use the Empirical Rule for Normal Distributions to answer parts (a) and (b)?

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?

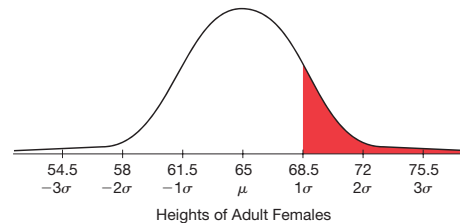


Approximately 81.5% of adult males have a height between 62 inches and 74 inches. I know that 13.5% of data is between 1 and 2 standard deviations above the mean. I also know that 95% of the data is within 2 standard deviations of the mean. So, $95 - 13.5$, or 81.5% of the data is between 2 standard deviations below the mean and 1 standard deviation above the mean. The sigma values indicate that this distribution represents data from a population.

Keep in mind that 1σ corresponds to a data value that is one standard deviation greater than the population mean and -1σ corresponds to a data value that is one standard deviation less than the mean.



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

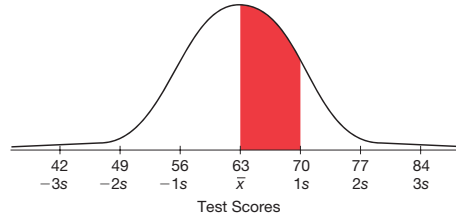


Approximately 16% of adult female heights are taller than 68.5 inches. I know that 50% of the data is above the mean. I also know that 34% of the data is between the mean and 1 standard deviation above the mean. So, $50 - 34$, or 16% of the data is more than 1 standard deviation above the mean. The sigma values indicate that this distribution represents data from a population.

Guiding Questions for Share Phase, Question 2, part (c)

- What percent of data are within one standard deviation of the mean?
- How can you use the symmetry of a normal curve to calculate percentages on either side of the mean?
- What other interval of test scores is that same size as the interval between 63 and 70 points?

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



Approximately 34% of the history test scores are between 63 points and 70 points.

I know that 68% of the test scores are within 1 standard deviation of the mean and the normal curve is symmetric. So, half of the 68% are between the mean and 1 standard deviation above the mean.

The s-values indicate that this distribution represents data from a sample.



Be prepared to share your solutions and methods.

Check for Students' Understanding

1

Most days Teresa eats dinner at 6:00 PM. Her dinner time can be modeled using a normal distribution curve with a standard deviation of 15 minutes.

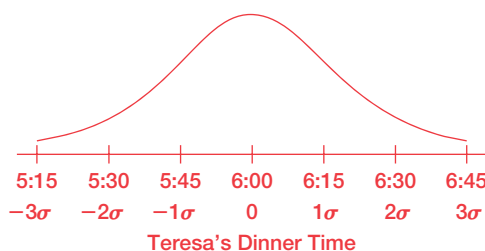
1. Identify the mean in this situation.

The mean is 6:00 PM.

2. What is the meaning of one standard deviation in this situation?

One standard deviation in this situation is 15 minutes, so 6:15 PM is one standard deviation above the mean and 5:45 PM is one standard deviation below the mean.

3. Sketch a standard normal curve for this situation that includes 3 standard deviations above and below the mean.



4. What does zero represent on the horizontal axis?

A zero on the horizontal axis represents the mean, 6:00 PM.

5. What does one represent on the horizontal axis?

A one on the horizontal axis represents the data value that is one standard deviation above the mean, 6:15 PM.

6. What does two represent on the horizontal axis?

A two on the horizontal axis represents the data value that is two standard deviations above the mean, 6:30 PM.

7. What does negative one represent on the horizontal axis?

A negative one on the horizontal axis represents the data value that is one standard deviation below the mean, 5:45 PM.

8. What does negative two represent on the horizontal axis?

A negative two on the horizontal axis represents the data value that is two standard deviations below the mean, 5:30 PM.

Below the Line, Above the Line, and Between the Lines

Z-Scores and Percentiles

LEARNING GOALS

In this lesson, you will:

- Use a z-score table to calculate the percent of data below any given data value, above any given data value, and between any two given data values in a normal distribution.
- Use a graphing calculator to calculate the percent of data below any given data value, above any given data, and between any two given data values in a normal distribution.
- Use a z-score table to determine the data value that represents a given percentile.
- Use a graphing calculator to determine the data value that represents a given percentile.

ESSENTIAL IDEAS

- A z-score is a number that describes a specific data value's distance from the mean, in terms of standard deviation units.
- A percentile is a data value for which a certain percentage of the data is below the data value.
- A z-score table or a graphing calculator is used to compute data values, the percent of data values in a given interval, and percentiles.

KEY TERMS

- z-score
- percentile

COMMON CORE STATE STANDARDS FOR MATHEMATICS

S.ID Interpreting Categorical and Quantitative Data

Summarize, represent, and interpret data on a single count or measurement variable

4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

Overview

The fuel efficiency of hybrid cars is modeled by a normal curve. Students use the model to answer questions about the percent of data in given intervals based on integer multiples of the standard deviation from the mean. In the next activity, within the same context, the data values are not aligned with integer multiples of the standard deviation from the mean and students use a z-score table to determine percents. A worked example uses a graphing calculator to determine percents. The number of text messages teens send and receive every day is modeled using a normal distribution and students use the mean and standard deviation to calculate various percentiles. A z-score table and graphing calculators can be used to answer questions related to percentiles.

Warm Up

1

The battery life of a laptop computer has a mean of 16 hours with a standard deviation of 2.5 hours.

Explain what this means to someone who is unfamiliar with the terms mean and standard deviation and apply the Empirical Rule of Normal Distributions.

The mean implies the average battery life for this type of laptop is 16 hours.

About 68% of these laptop computers will have a battery life between 13.5 and 18.5 hours.

About 95% of these laptop computers will have a battery life between 11 and 21 hours.

About 99.7% of these laptop computers will have a battery life between 8.5 and 23.5 hours.

1.3

Below the Line, Above the Line, and Between the Lines

Z-Scores and Percentiles

LEARNING GOALS

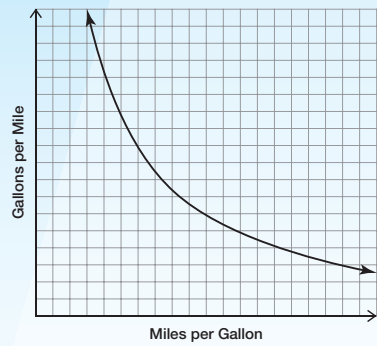
- In this lesson, you will:
- Use a z-score table to calculate the percent of data below any given data value, above any given data value, and between any two given data values in a normal distribution.
 - Use a graphing calculator to calculate the percent of data below any given data value, above any given data, and between any two given data values in a normal distribution.
 - Use a z-score table to determine the data value that represents a given percentile.
 - Use a graphing calculator to determine the data value that represents a given percentile.

KEY TERMS

- z-score
- percentile

In 2013, the labels on new vehicles sold in the U.S. got a little different. Instead of just showing how many miles per gallon the vehicle gets, the label also shows the number of gallons per mile it gets.

Why is that important? The reason can be seen on this graph.



When a car gets a low number of miles per gallon (say, 12 mpg), switching to a slightly higher number (say, 15 mpg) represents a large decrease in the number of gallons per mile, which is a big savings—even bigger than switching from 30 mpg to 50 mpg!

How many gallons per mile does your car get?

Problem 1

The fuel efficiency of hybrid cars, in miles per gallon, is modeled by a standard normal distribution. Students use the Empirical Rule for Normal Distributions to estimate the percent of data within specific intervals. The concept of z-scores is introduced in order to determine the percent of data in an interval that is not aligned with integer multiples of the standard deviation from the mean. Students use a z-score table to determine the percent of data less than a given data value. A worked example uses the graphing calculator to determine the percent of data below a given z-score. Different student work provide additional opportunity to further analyze using a z-score table and a graphing calculator to determine the percent of data in a specified interval.

Grouping

Have students complete Question 1 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Question 1

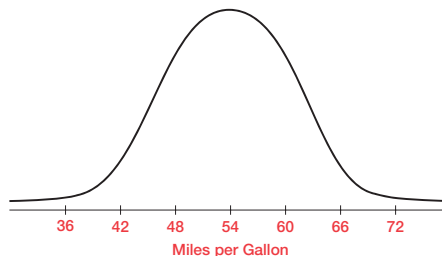
- What mpg is associated with the value under the peak of the curve?
- What percent of the data is between the mean and one standard deviation above the mean?

PROBLEM 1 Off the Mark



1. The fuel efficiency of a sample of hybrid cars is normally distributed with a mean of 54 miles per gallon (mpg) and a standard deviation of 6 miles per gallon.

- a. Use the mean and standard deviation to label the intervals on the horizontal axis of the normal curve in miles per gallon.



So,
1s = 6 mpg.



- b. Determine the percent of hybrid cars that get less than 60 miles per gallon. Explain your reasoning.

Approximately 84% of hybrid cars get less than 60 miles per gallon.

I know that 34% of the data is between the mean and 1 standard deviation above the mean. I also know that 50% of the data is below the mean. So, $34 + 50$, or 84% of the data, is below 1 standard deviation above the mean.

- c. Determine the percent of hybrid cars that get less than 66 miles per gallon. Explain your reasoning.

Approximately 97.5% of hybrid cars get less than 66 miles per gallon.

I know that 47.5% of the data is between the mean and 2 standard deviations above the mean. I also know that 50% of the data is below the mean. So, $47.5 + 50$, or 97.5% of the data is below 2 standard deviations above the mean.



- d. Determine the percent of hybrid cars that get less than 72 miles per gallon. Explain your reasoning.

Approximately 99.85% of hybrid cars get less than 72 miles per gallon.

I know that 49.85% of the data is between the mean and 3 standard deviations above the mean. I also know that 50% of the data is below the mean. So, $49.85 + 50$, or 99.85% of the data is below 3 standard deviations above the mean.

- What percent of the data is below the mean?
- In terms of the mean and standard deviations, how is the percent of hybrid cars that get less than 60 mpg described?
- What sum represents the data that is below one standard deviation above the mean?
- In terms of the mean and standard deviations, how is the percent of hybrid cars that get less than 66 mpg be described?

continued on the next page

- What percent of the data is between the mean and two standard deviations above the mean?
- What sum represents the data that is below two standard deviations above the mean?
- What percent of the data is between the mean and three standard deviations above the mean?
- In terms of the mean and standard deviations, how is the percent of hybrid cars that get less than 72 mpg be described?
- What sum represents the data that is below three standard deviations above the mean?

Grouping

- Ask a student to read the information. Discuss as a class.
- Have students complete Question 2 with a partner. Then have students share their responses as a class

Guiding Questions for Share Phase, Question 2

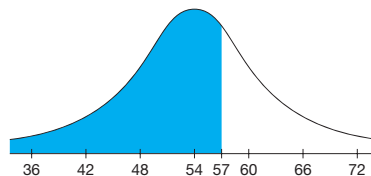
- If an interval does not align with integer multiples of the standard deviations, why is it challenging to determine the percent of data in that interval?
- Is the percent of data in each of the intervals 54–57 and 57–60 the same or different? Explain.



When data values are aligned with integer multiples of the standard deviation from the mean, you can use the Empirical Rule for Normal Distributions to calculate the percent of data values less than that value. But what if a data value does not align with the standard deviations?



2. Let's consider the fuel efficiency of hybrid cars again. The mean is 54 miles per gallon and 1 standard deviation is 6 miles per gallon. What percent of cars get less than 57 miles per gallon?



- a. How many standard deviations from the mean is 57 miles per gallon? Explain how you determined your answer.

Fifty-seven miles per gallon is 0.5 standard deviation greater than the mean.

I determined the answer by subtracting the mean from 57, then dividing the result by the value of 1 standard deviation in miles per gallon.

One standard deviation is 6 miles per gallon, so 2 miles per gallon represents $\frac{3}{6}$, or $\frac{1}{2}$, of a standard deviation.

$$\frac{57 - 54}{6} = \frac{3}{6} = 0.5$$



- b. Greg incorrectly estimated the percent of hybrid cars that get less than 57 miles per gallon.

Greg

Approximately 67% of hybrid cars get less than 57 miles per gallon.

$$50\% + \frac{1}{2}(34\%) \approx 67\%$$

Explain why Greg's reasoning is incorrect.

Multiplying $\frac{1}{2}$ by 34% is not reasonable because the percent of data in the intervals 54–57 and 57–60 is different. Of the two intervals, 54–57 has the larger percent of data and 57–60 has the smaller percent.

The actual percent will be larger than Greg's estimate.

- Which interval has the largest percent of data? How do you know?
- Which interval has the smallest percent of data? How do you know?

Grouping

- Ask a student to read the information, definition, and worked example. Discuss as a class.
- Have students complete Questions 3 through 5 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Question 3

- Compared to the mean, what data values produce a positive z-score?
- Compared to the mean, what data values produce a negative z-score?
- What do you know about a data value with a z-score of -1.2 and a data value with a z-score of 1.2 ?



The number you calculated in Question 2 is a z-score. A **z-score** is a number that describes a specific data value's distance from the mean in terms of standard deviation units.

For a population, a z-score is determined by the equation

$$z = \frac{(x - \mu)}{\sigma}$$

where x represents a value from the data.

You can use a z-score table to determine the percent of data less than a given data value with a corresponding z-score.

A z-score table is provided at the end of this lesson.

So a z-score is just how many standard deviations the data value is from the mean.



To determine the percent of hybrid cars that get less than 57 miles per gallon with a z-score table, first calculate the z-score for 57 miles per gallon. In Question 2, you calculated the score for 57 miles per gallon as 0.5.

z	0.0	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5433	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088

Next, locate the row that represents the ones and tenths place of the z-score. For a z-score of 0.5, this is the row labeled 0.5. Also locate the column that represents the hundredths place of the z-score. For a z-score of 0.5, this is the column labeled 0.0. Note that the table represents z-scores only to the hundredths place.

Finally, locate the cell that is the intersection of the row and column. The numbers in each cell represent the percent of data values below each z-score. For a z-score of 0.5, the corresponding cell reads 0.6915.

This means that 69.15% of hybrid cars get less than 57 miles per gallon.



3. What would a negative z-score indicate? Explain your reasoning.

A negative z-score would mean that the data value being considered is below the mean of the normal distribution.

A graphing calculator can determine a percent of data below a z-score. This function is called the normal cumulative density function (normalcdf). The function determines the percent of data values within an interval of a normal distribution.



You can also use a graphing calculator to determine the percent of data below a given z-score.

Step 1: Press **2nd** and then **VARΣ**.
Select **2:normalcdf(**

Step 2: Enter the lower bound of the interval, the upper bound of the interval, the mean, and the standard deviation, including commas between each number.

Step 3: Press **ENTER**. Calculator results will vary from results obtained using a z-score table.

In this problem, the lower bound is negative infinity, the upper bound is 57, the mean is 54, and the standard deviation is 6. Most calculators do not have an infinity button. You can use -1×10^{99} , a very small number, for the lower bound.

Guiding Questions for Share Phase, Questions 4 and 5

- What number did you use for the lower bound of the interval?
- Could you use a different value for the lower bound of the interval? Explain.
- Do you prefer to use a z-score table or a calculator?

4. Use a graphing calculator to determine the approximate percent of hybrid cars that get less than 57 miles per gallon.

0.6914624678

Approximately 69.15% of hybrid cars get less than 57 miles per gallon.



5. How does this answer compare to the answer you got by using the z-score table?

Both the table and the calculator give the same z-score.

Grouping

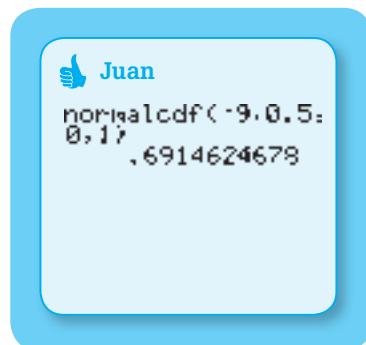
Have students complete Questions 6 and 7 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Question 6

- Which student used actual data values for the normalcdf command?
- Which student used actual z-score values for the normalcdf command?
- Do you prefer to use actual data values or z-score values? Explain.
- What is the z-score for 0 miles per gallon?
- Why did Juan and Michael use -9 and 0 for their lower bounds?
- How does the lower bound value for Question 6 compare to the lower bound value for Question 5?
- If you used a lower bound value of zero for Question 5, would that produce an accurate result? Explain.



6. Juan and Michael used a graphing calculator to determine that the percent of hybrid cars that get less than 57 miles per gallon is approximately 69.15%. Juan entered values in standard deviation units, and Michael entered values in terms of miles per gallon.



- a. Explain why Juan and Carlos used different values but still got the same result.
- Juan used values based on the standard normal distribution curve, which has a mean of 0 and a standard deviation of 1.
- Michael used values based on the normal distribution curve which has the actual values for miles per gallon.
- b. Explain why Michael used 0 for the lower bound of the interval. Explain why Juan used -9 for the lower bound of the interval.
- Michael used 0 because that is the lowest possible value for miles per gallon. Negative values for miles per gallon do not make sense.
- Juan used -9 because the standard deviation is 6 and -9 standard deviations is equivalent to $6 \cdot 9$, or 54 miles per gallon below the mean of 54. This is equivalent to the lower bound of 0 that Michael used.

Guiding Questions for Share Phase, Question 7

- Does using a graphing calculator produce an accurate answer?
- Does using a z-score table produce an accurate answer?
- Does the difference between the two answers make one right and one wrong? Explain.
- Does estimation come into play when calculating the percent of data in an interval of a normally distributed data set?
- In the real world, should you expect a calculation for the percent of data in a particular interval to be exact?

Problem 2

The fuel efficiency of hybrid cars is the context again. Students determine the percent of data greater than a specified value and in between two specified values. The questions can be answered using a z-score table or a graphing calculator.

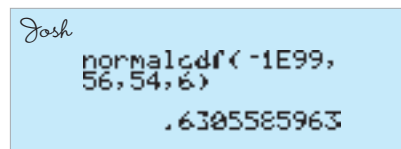
Grouping

Have students complete Questions 1 through 4 with a partner. Then have students share their responses as a class.



7. Josh calculated the percent of hybrid cars that get less than 56 miles per gallon using the z-score table and a graphing calculator.

z	0.0	0.01	0.02	0.03
0.0	0.5000	0.5040	0.5080	0.5120
0.1	0.5398	0.5438	0.5478	0.5517
0.2	0.5793	0.5832	0.5871	0.5910
0.3	0.6179	0.6217	0.6255	0.6293



Explain why Josh received different results from the z-score table and the graphing calculator.

The z-score for 56 miles per gallon is $\frac{56 - 54}{6}$, or $\frac{1}{3}$.

The z-scores in the table are only given to the hundredths place, but the actual z-score is 0.3. The graphing calculator gave a larger percent because it calculated the percent of data below $\frac{1}{3}$ standard deviation above the mean, whereas the table shows the percent of data below $\frac{33}{100}$ standard deviation above the mean.

PROBLEM 2 More or Less . . .



1. Calculate the percent of hybrid cars that get less than 50 miles per gallon.

About 25.14% of hybrid cars get less than 50 miles per gallon.

Calculate the z-score for 50 miles per gallon.

$$\begin{aligned} z &= \frac{50 - 54}{6} \\ &= \frac{-4}{6} \\ &\approx -0.67 \end{aligned}$$

In the z-score table, the percent of data that is less than a z-score of -0.67 is 0.2514.

2. Use your answer to Question 1 to calculate the percent of hybrid cars that get more than 50 miles per gallon. Explain your reasoning.

About 74.86% of hybrid cars get more than 50 miles per gallon.

If 25.14% of hybrid cars get less than 50 miles per gallon then $100 - 25.14$, or 74.86% of hybrid cars get more than 50 miles per gallon.

Guiding Questions for Share Phase, Questions 1 and 2

- What information do you need in order to calculate a z-score?
- What information does a z-score provide?
- Did you use a z-score table or a graphing calculator to answer Question 1?
- Could you answer Question 2 without determining the z-score first? Explain.

Guiding Questions for Share Phase, Questions 3 and 4

- Do you have to use a z-score table to calculate the percent of hybrid cars that get less than 60 miles per gallon?
- How do you calculate the percent of hybrid cars that get less than 60 miles per gallon?
- How can you calculate the answer to Question 4 without using the answer from Question 3?
- Is it more efficient to use a z-score table or a graphing calculator to determine the percent of data in a given interval? Explain.

Problem 3

The number of text messages teens send and receive every day is modeled using a normal distribution. Given the mean and standard deviation, students calculate the 50th percentile and answer questions related to the 90th percentile, the 10th percentile, and the 20th percentile. A z-score table and graphing calculator are used to determine the total number of text messages represented by specific percentiles.

Grouping

- Ask a student to read the information, definition, and calculator example. Discuss as a class.

3. Calculate the percent of hybrid cars that get less than 50 miles per gallon and the percent of hybrid cars that get less than 60 miles per gallon.
From Question 1, I know that about 25.14% of hybrid cars get less than 50 miles per gallon.
The z-score for 60 is 1 because it is 1 standard deviation above the mean. About 84.13% of hybrid cars get less than 60 miles per gallon.



4. Use your answer to Question 3 to calculate the percent of hybrid cars that get between 50 and 60 miles per gallon. Explain your reasoning.
About 58.99% of hybrid cars get between 50 and 60 miles per gallon.
I subtracted the percent of hybrid cars that get less than 50 miles per gallon from the percent of hybrid cars that get less than 60 miles per gallon.

PROBLEM 3 Top Texters



You may have heard someone say, “My baby’s weight is in the 90th percentile” or, “My student scored in the 80th percentile in math.” What do these phrases mean?

A **percentile** is a data value for which a certain percentage of the data is below the data value in a normal distribution. For example, 90% of the data in a set is below the value at the 90th percentile, and 80% of the data is below the value at the 80th percentile.

The number of text messages teens send and receive every day can be represented as a normal distribution with a mean of 100 text messages per day and a standard deviation of 25 texts per day.



1. Calculate the 50th percentile for this data set. Explain your reasoning.
The 50th percentile is 100 text messages per day because the 50th percentile represents the center of the distribution, the mean.
2. Would a teen in the 90th percentile send and receive more or fewer than 100 text messages per day? Explain your reasoning.
Teens in the 90th percentile would send more than 100 text messages per day. Ninety-percent of the teens send and receive fewer text messages than a teen in the 90th percentile and 10% of teens receive and send more text messages.
3. Would a teen in the 10th percentile send and receive more or fewer than 100 text messages per day? Explain your reasoning.
Teens in the 10th percentile would send fewer than 100 text messages per day. Ten percent of the teens send and receive fewer text messages than a teen in the 10th percentile and 90% of teens receive and send more text messages.

- Have students complete Questions 1 through 5 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Questions 1 through 3

- Does the 50th percentile represent the center of the distribution?
- Is the 50th percentile associated with the mean of the data set?
- Is 100 text messages per day the mean in this situation?

continued on the next page

- Would teens in the 40th percentile send more or less text messages than teens in the 50th percentile?
- Would teens in the 60th percentile send more or less text messages than teens in the 50th percentile?
- What percent of teens send and receive fewer text messages than a teen in the 90th percentile?
- What percent of teens send and receive more text messages than a teen in the 90th percentile?
- If 90% of the teens receive and send more text messages than a particular student, what percentile is the student associated with?
- If 10% of the teens receive and send more text messages than a particular student, what percentile is the student associated with?

Guiding Questions for Share Phase, Questions 4 and 5

- What is the decimal equivalent of 90%?
- Why is 1.28 the closest z-score to 90%?
- How was the z-score closest to 90% determined?
- Is the data value associated with the 90th percentile 1.28 standard deviations greater than the mean?
- How is the value of 1.28 used to calculate the data value that represents the 90th percentile?

4. Use a z-score table to determine the 90th percentile for teen text messages.
- a. Determine the percent value in the z-score table that is closest to 90%. Explain what information the z-score provides.

The percent value closest in the z-score table that is closest to 90% is 0.8997. Its z-score is 1.28, which means the actual data value associated with the 90th percentile is 1.28 standard deviations greater than the mean.

- b. Calculate the 90th percentile. Show your work.

The data value that represents the 90th percentile is 132 texts.

$$1.28 = \frac{x - 100}{25}$$

$$32 = x - 100$$

$$132 = x$$



You can also use a graphing calculator to calculate a percentile. To calculate a percentile, you can use the inverse of the normal cumulative density function (invNorm). The invNorm function takes a percent as input and returns the data value.



You can use a graphing calculator to determine the total number of text messages that correspond to the 90th percentile.

Step 1: Press **2nd** and then **VARs**.
Select **3:invNorm(**

Step 2: Enter the percentile in decimal form, the mean, and the standard deviation, including commas between each number.

Step 3: Press **ENTER**.

In the texting situation, the percentile is 0.90, the mean is 100, and the standard deviation is 25.

5. Determine the total number of text messages that represent the 20th percentile.

Approximately 80 total text messages represent the 20th percentile.

I used the invNorm feature on a graphing calculator.

$$\text{invNorm}(0.20, 100, 25) \approx 79$$



Be prepared to share your solutions and methods.

- What values would you enter for the invNorm command to calculate the total number of text messages that represent the 75th percentile?

z-scores and percent of data below the z-score										
z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.0
-3.4	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
-3.2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
-3.1	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
-3.0	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
-2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
-2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
-1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
-1.3	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
-1.2	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151
-1.1	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
-1.0	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
-0.9	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
-0.8	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
-0.7	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420
-0.6	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
-0.5	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
-0.4	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
-0.3	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
-0.2	0.3829	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
-0.1	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
-0.0	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

z	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5967	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7068	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8269	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8960	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9666	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Check for Students' Understanding

The battery life of a laptop computer has a mean of 16 hours with a standard deviation of 2.5 hours.

1. Is the z-score associated with 17 hours positive or negative? Explain your reasoning.

The z-score associated with 17 hours is positive because this value is greater than the mean.

2. Is the z-score associated with 15 hours positive or negative? Explain your reasoning.

The z-score associated with 15 hours is negative because this value is less than the mean.

3. What is the z-score associated with 17 hours? 15 hours?

The z-score associated with 17 hours is 0.4 and the z-score associated with 15 hours is -0.4 .

$$z = \frac{17 - 16}{2.5} = \frac{1}{2.5} = 0.4$$

$$z = \frac{15 - 16}{2.5} = \frac{-1}{2.5} = -0.4$$

4. What percentage of laptops have a battery life that is less than 17 hours?

Approximately 66% of laptops have a battery life that is less than 17 hours.

I used a z-score table to look up the percent of data that is less than a z-score of 0.40, 0.6554.

5. What percentage of laptop batteries' charges will last more than 17 hours?

Approximately 34% of laptops have a battery life that is greater than 17 hours.

I subtracted the answer to Question 4 from 1.

$$1 - 0.6554 = 0.3446$$

You Make the Call

Normal Distributions and Probability

LEARNING GOALS

In this lesson, you will:

- Interpret a normal curve in terms of probability.
- Use normal distributions to determine probabilities.
- Use normal distributions and probabilities to make decisions.

ESSENTIAL IDEAS

- Normal distribution curves are used to model real world situation.
- Normal distributions are used to determine probabilities and make decisions.

COMMON CORE STATE STANDARDS FOR MATHEMATICS

S.MD Using Probability to Make Decisions

Use probability to evaluate outcomes of decisions

6. (+) Use probabilities to make fair decisions.
7. (+) Analyze decisions and strategies using probability concepts.

Overview

Teen texting, pizza shop deliveries, and a home grown tomato contest are situations modeled by a normal distribution. Students use a normal curve to determine the probabilities of randomly selected students sending and receiving text messages, the likelihood of delivering a pizza within a given time frame, and selecting a prize-winning tomato.

Warm Up

1

The mean battery life of a video camera is 10 hours with a standard deviation of 1.5 hours.

1. Determine the percent of video cameras with a battery life greater than 10 hours.

If the mean is 10 hours, then 50% of the camera batteries have a battery life greater than 10 hours.

2. Use a z-score table to determine the percent of camera batteries with a battery life greater than 12 hours.

Approximately 9% of video cameras have a battery life greater than 12 hours.

To determine the answer, I calculated the z-score for 12 hours, then looked up the percent in the z-score table, and subtracted that value from 1.

$$z = \frac{12 - 10}{1.5} = \frac{2}{1.5} = 1.33$$

The percent of data less than a z-score of 1.33 is 0.9082. So, the percent of data greater than a z-score of 1.33 is $1 - 0.9082$, or 0.0918.

3. Use the graphing calculator to determine the percent of camera batteries with a battery life greater than 12 hours.

Approximately 9% of video cameras have a battery life greater than 12 hours.

$$\text{normalcdf}(1.33, 1E99, 0, 1) \approx 0.0918$$

You Make the Call

Normal Distributions and Probability

LEARNING GOALS

In this lesson, you will:

- Interpret a normal curve in terms of probability.
- Use normal distributions to determine probabilities.
- Use normal distributions and probabilities to make decisions.

You can grow both tomatoes and potatoes easily in a home garden—separately— but what about a plant that can grow both vegetables (or is it a fruit and a vegetable?) at the same time?

A company based in the United Kingdom created just that—a hybrid plant that produces both tomatoes above the ground and potatoes below. This remarkable plant was not created through genetic engineering, but rather by grafting the two types of plants together at the stem.

Now for the important question: What would you call this hybrid plant?

Problem 1

The teen texting situation from the previous lesson is revisited. The situation is modeled by a normal distribution curve with a mean of 100 text messages per day and a standard deviation of 25 text messages per day. Students interpret the normal distribution curve in terms of probabilities and calculate the probability that a randomly selected teen sends and receives a specified number of text messages per day.

Grouping

Have students complete Questions 1 through 3 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Questions 1 through 3

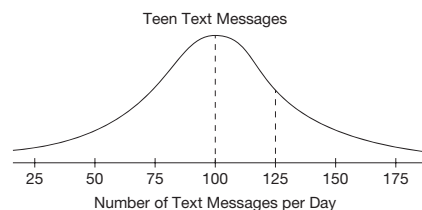
- What is the mean in this situation?
- What is the value of one standard deviation above the mean?
- What percent of the data is between the mean and one standard deviation above the mean?
- What is the value of one standard deviation below the mean?
- What percent of the data is between the mean and one standard deviation below the mean?
- What percent of the data is below the mean?

PROBLEM 1 Teens and Texting

So far, you have explored the percent of data values that fall within specified intervals. However, you can also interpret a normal distribution in terms of probabilities.



Based on a survey, the number of text messages that teens send and receive every day is a normal distribution with a mean of 100 text messages per day and a standard deviation of 25 text messages per day.



You randomly select a teen from the survey. Calculate each probability.

1. Determine the probability that the randomly selected teen sends and receives between 100 and 125 text messages per day.

The probability that a randomly selected teen sends and receives between 100 and 125 text messages per day is 34%.

The mean is 100 and 1 standard deviation above the mean is 125. I know that 34% of the data is between the mean and 1 standard deviation above the mean.

2. Determine the probability that the randomly selected teen sends and receives fewer than 75 text messages per day.

The probability that a randomly selected teen sends and receives fewer than 75 text messages per day is 16%.

The mean is 100 and 1 standard deviation below the mean is 75. I know that 34% of the data is between the mean and 1 standard deviation below the mean. I also know that 50% of the data is below the mean. So, $50 - 34$, or 16% of the data is more than 1 standard deviation below the mean.



3. Determine the probability that the randomly selected teen sends and receives more than 140 text messages per day.

The probability that a randomly selected teen sends and receives more than 140 text messages per day is approximately 5.48%.

I used my graphing calculator and entered $\text{normalcdf}(140, 1 \cdot 10^{99}, 100, 25)$. The lower bound is 140, the upper bound is infinity so I used a large positive number $1 \cdot 10^{99}$, the mean is 100, and the standard deviation is 25.

- Can a graphing calculator be used to determine the probability in this situation?
- How can a graphing calculator be used to determine the probability in this situation?
- What value was used for the lower bound?
- If the upper bound is positive infinity, what value can be used on a graphing calculator?

Problem 2

The mean and standard deviation describing the delivery time for two competing pizza shops are given. Students use this information to determine which shop has a better chance of delivering a pizza in a specified time. Using a graphing calculator, they compute the probabilities of delivering an order within a specified amount of time from each shop and make a recommendation based on the probabilities.

Grouping

Have students complete Questions 1 through 3 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Questions 1 through 3

- Could the size of an order influence the delivery time of an order?
- Could the number of delivery drivers influence the delivery time of an order?
- Could the distance between the customer and the pizza shop influence the delivery time of an order?
- On an average, which pizza shop will deliver an order more quickly? How do you know?
- Whose delivery times are more consistent and closer to the mean delivery time?

PROBLEM 2 Pizza Anyone?



You have collected data on the delivery times for two local pizza shops, Antonio's Pizza and Wood Fire Pizza. Based on your data, Antonio's Pizza has a mean delivery time of 30 minutes and a standard deviation of 3 minutes. Wood Fired Pizza has a mean delivery time of 25 minutes and a standard deviation of 8 minutes.

1. What factors could influence the delivery time of an order from either pizza shop?

Answer will vary.

Factors that could influence the delivery time from either pizza shop include:

- The size of an order
- The number of delivery drivers working
- The customers' distance from the pizza shop

2. What can you conclude based only on the mean and standard deviation for each pizza shop?

Answers will vary.

The mean delivery time of Antonio's Pizza is greater than the mean delivery time of Wood Fired Pizza. That means that on average, Wood Fire Pizza will deliver an order more quickly.

The standard deviation for Antonio's Pizza delivery time is less than the standard deviation for Wood Fired Pizza's delivery time. This means the Antonio's Pizza delivery times are more consistent and closer to the mean delivery time. Wood Fired Pizza delivery times are less consistent and further away from the mean delivery time.



3. A friend of yours is planning a party. She needs the pizza for the party delivered in 35 minutes or less or the party will be a complete disaster! Which pizza shop has a greater probability of delivering the order within 35 minutes?

The probability that Antonio's Pizza will deliver an order within 35 minutes is approximately 95.22%.

I used my graphing calculator and entered $\text{normalcdf}(0,35,30,3)$. The lower bound is 0 minutes, the upper bound is 35 minutes, the mean is 30 minutes, and the standard deviation is 3 minutes.

The probability that Wood Fired Pizza will deliver an order within 35 minutes is approximately 89.34%.

I used my graphing calculator and entered $\text{normalcdf}(0,35,25,8)$. The lower bound is 0 minutes, the upper bound is 35 minutes, the mean is 25 minutes, and the standard deviation is 8 minutes.

My friend should order from Antonio's Pizza because Antonio's Pizza has a higher probability of delivering an order within 35 minutes.

The probability is 95.22%, compared to 89.34% for Wood Fired Pizza.

- Whose delivery times are less consistent and further away from the mean delivery time?
- How can the graphing calculator be used to determine the probability in this situation?
- What value was used for the lower bound?
- What value was used for the upper bound?
- What mean and standard deviation was used to determine the probability in this situation?

Problem 3

The mean and standard deviation describing the diameters of a normal distribution of Brad's and Toby's home grown tomatoes are given. They are competing against each other for prize-winning tomatoes. Prize-winning tomatoes have a diameter between 4 and 4.5 inches. Students determine which person is more likely to win a contest if the judge randomly selects a tomato from each basket.

Grouping

Have students complete Question 1 with a partner. Then have students share their responses as a class.

Guiding Questions for Share Phase, Question 1

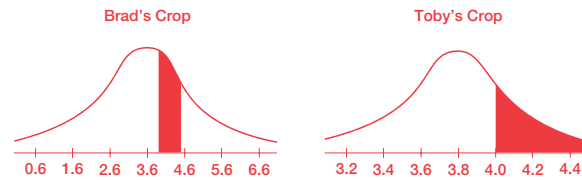
- Is it helpful to sketch the normal distribution curve for each situation? Why?
- What percent of Brad's distribution represents prize-winning tomatoes?
- What percent of Toby's distribution represents prize-winning tomatoes?
- Can a graphing calculator be used to determine the probability in this situation?
- How can a graphing calculator be used to determine the probability in this situation?

PROBLEM 3 You Say Tomato, I Say Prize-Winning Tomato!



- Brad and Toby both plan to enter the county tomato growing competition. Each person who enters the competition must submit a basket of tomatoes. The judges randomly select a tomato from each contestant's basket. According to the rules of the competition, a "golden" tomato has a diameter between 4 inches and 4.5 inches. The diameters of tomatoes in Brad's basket are normally distributed with a mean diameter of 3.6 inches and a standard deviation of 1 inch. The diameters of tomatoes in Toby's basket are also normally distributed with a mean diameter of 3.8 inches and a standard deviation of 0.2 inches. When the judges randomly select a tomato from Brad's and Toby's basket, whose is more likely to result in a "golden" tomato?

Brad has a slightly greater chance of having a "golden" tomato chosen from his basket by the judges. Brad's probability is 16.05% while Toby's is 15.84%. So, both have close to a 16% chance of having a "golden" tomato chosen from their basket. First, I sketched the a normal distribution curve for Brad and Toby,



Then I calculated the percent of tomatoes in each basket with a radius between 4 inches and 4.5 inches using my graphing calculator.

For Brad's basket, I entered $\text{normalcdf}(4,4.5,3.6,1)$. The lower bound is 4 inches, the upper bound is 4.5 inches, the mean is 3.6 inches, and the standard deviation is 1 inch. Approximately 16.05% of Brad's tomatoes have a radius between 4 inches and 4.5 inches.

For Toby's basket, I entered $\text{normalcdf}(4,4.5,3.8,0.2)$. The lower bound is 4 inches, the upper bound is 4.5 inches, the mean is 3.8 inches, and the standard deviation is 0.2 inch. Approximately 15.84% of Brad's tomatoes have a radius between 4 inches and 4.5 inches.



Be prepared to share your solutions and methods.

- What value can you use for the lower bound of the normalcdf command?
- What value can you use for the upper bound of the normalcdf command?

Check for Students' Understanding

1

The mean battery life of a video camera is 10 hours with a standard deviation of 1.5 hours.

1. Use a z-score table to determine the percent of video cameras with a battery life greater than 13 hours.

Approximately 2% of video cameras have a battery life greater than 13 hours.

To determine the answer, I calculated the z-score for 13 hours, then looked up the percent in the z-score table, and subtracted that value from 1.

$$z = \frac{13 - 10}{1.5} = \frac{3}{1.5} = 2$$

The percent of data less than a z-score of 2 is 0.9772. So, the percent of data greater than a z-score of 2 is $1 - 0.9772$, or 0.0228.

2. Verify your answer to Question 1 using a graphing calculator.

Approximately 2% of video cameras have a battery life greater than 13 hours.

$$\text{normalcdf}(13, 1E99, 10, 1.5) \approx 0.0228$$

3. What is the probability that a randomly selected video camera has a battery life between 8 and 9 hours?

The probability of randomly selecting a video camera with a battery life between 8 and 9 hours is approximately 16%.

I used the normalcdf command on a graphing calculator to determine the answer.

$$\text{normalcdf}(8, 9, 10, 1.5) \approx 0.1613$$

4. Use the Empirical Rule for Normal Distributions to determine the probability that a randomly selected video camera will have a battery life between 8.5 and 11.5 hours.

The probability that a randomly selected camera battery will need to be recharged between 8.5 and 11.5 hours is approximately 68% because 8.5 is one standard deviation below the mean and 11.5 is one standard deviation above the mean.

The Empirical Rule for Normal Distributions states that approximately 68% of the area under the normal curve is within one standard deviation of the mean.

5. Use a z-score table to compute the probability that a randomly selected video camera will have a battery life between 8.5 and 11.5 hours.

The probability of randomly selecting a video camera with a battery life between 8.5 and 11.5 hours is approximately 68%.

Because 8.5 hours is 1 standard deviation below the mean, its z-score is -1 . Using the same type of reasoning, I know the z-score for 11.5 hours is 1. To determine the answer, I looked up the percents in the z-score table, and subtracted them.

$$0.8413 - 0.1587 = 0.6826$$

6. Use the graphing calculator to compute the probability that a randomly selected camera battery will need to be recharged between 8.5 and 11.5 hours.

The probability of randomly selecting a video camera with a battery life between 8.5 and 11.5 hours is approximately 68%.

$$\text{normalcdf}(8.5, 11.5, 10, 1.5) \approx 0.6827$$

7. Compare the answers to Questions 4, 5, and 6.

The prediction made using the Empirical Rule for Normal Distributions matched the computations done using the formula in conjunction with the z-score table, as well as the answer computed using a graphing calculator.

Chapter 1 Summary

KEY TERMS

- discrete data (1.1)
- continuous data (1.1)
- sample (1.1)
- population (1.1)
- normal curve (1.1)
- normal distribution (1.1)
- mean (μ) (1.1)
- standard deviation (σ) (1.1)
- standard normal distribution (1.2)
- Empirical Rule for Normal Distributions (1.2)
- z-score (1.3)
- percentile (1.3)

1.1 Differentiating Between Discrete Data and Continuous Data

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.

Example

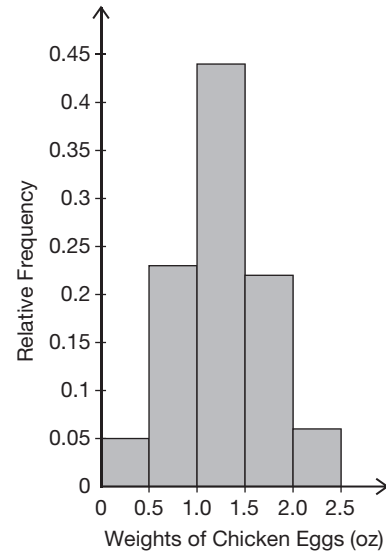
The heights of basketball players are examples of continuous data.

1.1 Drawing Distributions for Continuous Data

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

Example

Weights of Chicken Eggs (ounces)	Relative Frequency
0.0–0.5	0.05
0.5–1.0	0.23
1.0–1.5	0.44
1.5–2.0	0.22
2.0–2.4	0.06

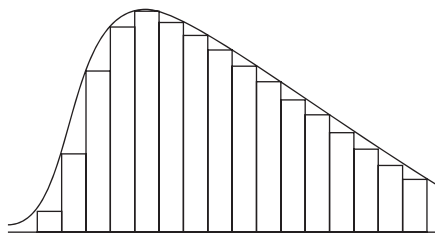


1.1 Recognizing the Difference Between Normal Distributions and Non-normal Distributions

A normal distribution is bell-shaped and symmetrical, and a non-normal distribution is neither bell-shaped nor symmetrical.

Example

The graph does not represent a normal distribution. It is neither bell-shaped nor symmetric, it is skewed.



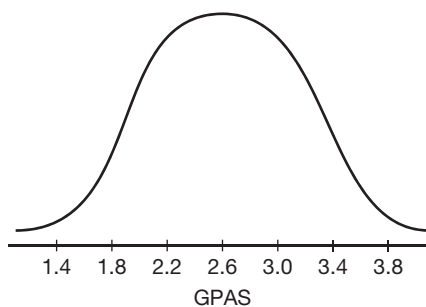
1.1 Recognizing and Interpreting Properties of a Normal Curve and a Normal Distribution

The mean of a normal curve is at the center of the curve. The standard deviation of a normal distribution describes how spread out the data are.

The symbol for the population mean is μ , and the symbol for the sample mean is \bar{x} . The standard deviation of a sample is represented with the variable s . The standard deviation of a population is represented with the symbol σ .

Example

The mean is 2.6 and the standard deviation is 0.4.

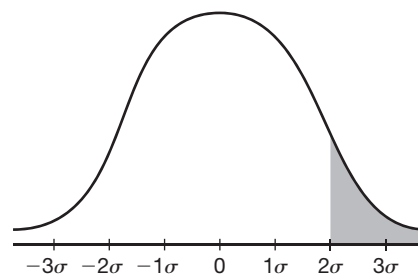


1.2 Recognizing the Connection Between Normal Curves, Relative Frequency Histograms, and the Empirical Rule for Normal Distributions

The standard normal distribution is a normal distribution with a mean value of zero and a standard deviation of 1. The Empirical Rule states that approximately 68% of the data in a normal distribution is within 1 standard deviation of the mean, 95% is within two standard deviations of the mean, and 99.7% is within three standard deviations of the mean.

Example

The percent of data that is more than 2 standard deviations above the mean for the standard normal curve is shaded.



1.2

Using the Empirical Rule for Normal Distributions to Determine the Percent of Data in a Given Interval

The Empirical Rule for Normal Distributions states that approximately 68% of the data in a normal distribution is within one standard deviation of the mean, 95% is within two standard deviations of the mean, and 99.7% is within three standard deviations of the mean. The percent of data for any normal distribution can be determined using the Empirical Rule.

Example

Determine the percent of commute times less than 36 minutes for a certain city, given that the commute times are normally distributed and the mean commute is 41 minutes with a standard deviation of 2.5 minutes.

A commute time of 36 minutes is 2 standard deviations below the mean. The Empirical Rule for Normal Distributions states that 50% of the data is below the mean and that 47.5% of the data is within 2 standard deviations below the mean. So, $50\% - 47.5\%$ or 2.5% of the data is below 2 standard deviations below the mean.

Approximately 2.5% of commute times are less than 36 minutes.

1.3

Using a z-score Table to Calculate the Percent of Data Below any Given Data Value, Above any Given Data Value, and Between any Two Given Data Values in a Normal Distribution

Data points can be converted into z-scores which represent the number of standard deviations the data value is from the mean. It is positive if above the mean and negative if below the mean. A z-score table can then be used to determine the percent of data you are looking for based on the z-scores.

Example

You can calculate the percent of adult men taller than 70 inches, given that adult men's heights are normally distributed and the mean height is 69.3 inches with a standard deviation of 2.8 inches.

$$\begin{aligned} z &= \frac{70 - 69.3}{2.8} \\ &= \frac{0.7}{2.8} \\ &= 0.25 \end{aligned}$$

About 59.87% of adult men are shorter than 70 inches, so $100 - 59.87$, or 40.13% of adult men are taller than 70 inches.

1.3 Using a Graphing Calculator to Calculate The Percent of Data Below any Given Data Value, Above any Given Data Value, and Between any Two Given Data Values in a Normal Distribution

To determine the percent of data between two scores on a normal curve, a graphing calculator can be used. The function and its arguments are entered as $\text{normalcdf}(\text{lower bound of the interval, upper bound of the interval, the mean, the standard deviation})$.

Example

You can determine the percent of adults with IQ scores between 102 and 132, given that IQ scores are normally distributed and the mean IQ score for adults is a 100 with a standard deviation is 15.

$\text{Normalcdf}(102, 132, 100, 15)$

Approximately 43.05% of adults have IQ scores between 102 and 132.

1.3 Using a Z-score Table to Determine The Data Value That Represents a Given Percentile

A percentile is the data value for which a certain percentage of the data is below the data value in a normal distribution. A z-score can be used to determine the data value. First, the percent value in the table closest to the percentage you are looking for should be found. Then the z-score for the percentile can be found from the table. This can be converted back to the original data value by using the formula for a z-score and solving for the x .

Example

You can determine the 80th percentile for SAT scores, given that SAT scores are normally distributed and the mean is 1500 with a standard deviation of 280.

The percent value in the z-score table that is closest to 80% is 0.0793. The z-score for this percent value is -1.41 .

$$-1.41 = \frac{x - 1500}{280}$$

$$-394.8 = x - 1500$$

$$1105.2 \approx x$$

The SAT score that represents the 80th percentile is approximately 1105.2.

1.3

Using a Graphing Calculator to Determine The Data Value That Represents a Given Percentile

To determine a percentile on a normal curve, a graphing calculator can be used. The function and its arguments are entered as $\text{invNorm}(\text{percentile in decimal form, the mean, the standard deviation})$.

Example

You can determine the 45th percentile for the length of time of a dance studio's recitals, given that the times are normally distributed and the mean is 145 minutes with a standard deviation of 7 minutes.

$$\text{Invnorm}(0.45, 145, 7) \approx 144.12$$

The length of time that represents the 45th percentile is approximately 144.12 minutes.

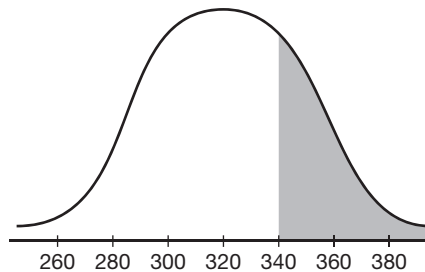
1.4

Interpreting a Normal Curve in Terms of Probability

The percent of data values that fall within specified intervals on a normal distribution can also be interpreted as probabilities.

Example

You can calculate the probability that a randomly selected annual precipitation amount in a city is more than 340 inches, given that the amounts are normally distributed with a mean of 320 inches and a standard deviation of 20 inches.



The probability that the randomly selected precipitation amount in a city is more than 340 inches is 16%.

The mean is 320 and one standard deviation above the mean is 340. I know that 34% of the data is between the mean and one standard deviation above the mean. I also know that 50% of the data is above the mean. So, $50 - 34$, or 16% of the data is more than one standard deviation above the mean.

1.4 Using Normal Distributions to Determine Probabilities

A normal distribution can be used to determine probabilities. The percent of data between specified intervals represents probabilities. The z-score table or a graphing calculator can be used to find the percents.

Example

You can determine the probability that a randomly selected student will score between a 74 and an 80 on an exam, if the exam scores are normally distributed and the mean is 82 with a standard deviation of 2.8

The probability that a randomly selected student will score between a 74 and an 80 is approximately 23.54%.

$$\text{Normalcdf}(74,80,82,2.8) \approx 0.2354.$$

1.4 Using Normal Distributions And Probabilities to Make Decisions

Determining probabilities of events occurring by using percentages from a normal distribution can help to make decisions about different products or situations.

Example

You can determine the factory that should be used to fill an order if it is needed in between 25 and 30 minutes, if Factory A has a mean order time of 27.5 minutes with a standard deviation of 0.8 minutes and Factory B has a mean order time of 28.2 minutes with a standard deviation of 1.2 minutes. Assume both order times can be represented by a normal distribution.

Factory A should be used because it has the best chance of filling an order between 25 and 30 minutes. The probability Factory A will fill the order between 25 and 30 minutes is 99.82%, while the probability that Factory B will fill the order between 25 and 30 minutes is only 92.94%.

$$\text{Factory A: } \text{normalcdf}(25, 30, 27.5, 0.8) \approx 0.9982$$

$$\text{Factory B: } \text{normalcdf}(25, 30, 28.2, 1.2) \approx 0.9294$$

