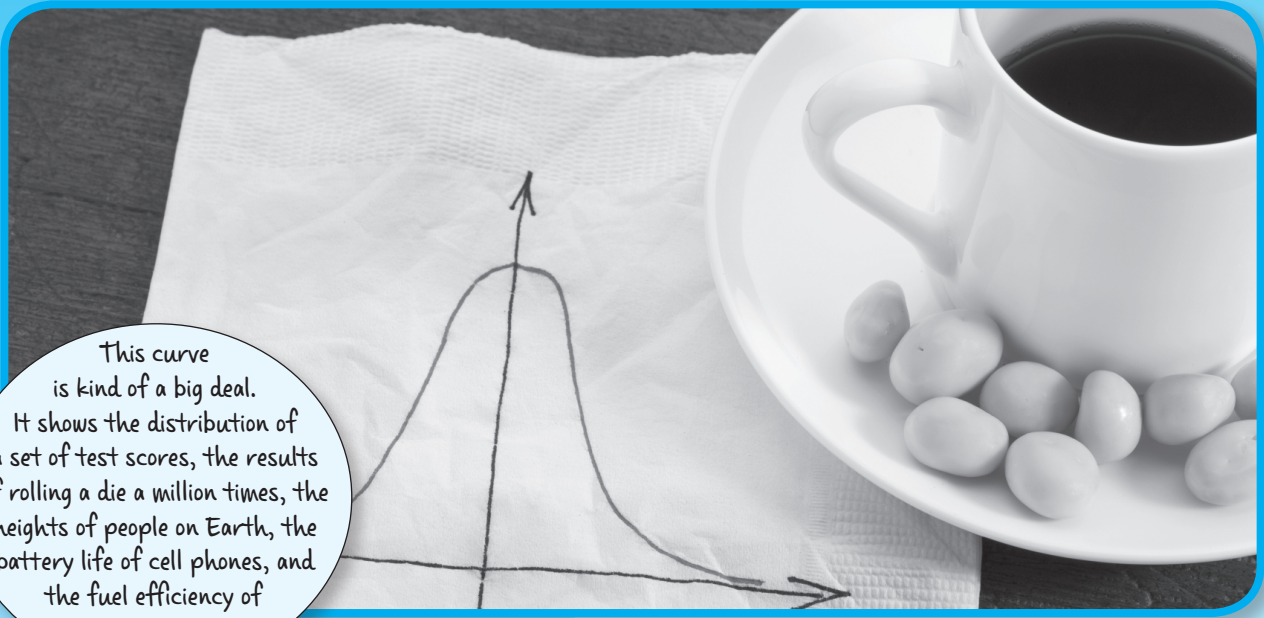


# Interpreting Data in Normal Distributions



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



# 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 ( $\mu$ )
- standard deviation ( $\sigma$ )

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.



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.



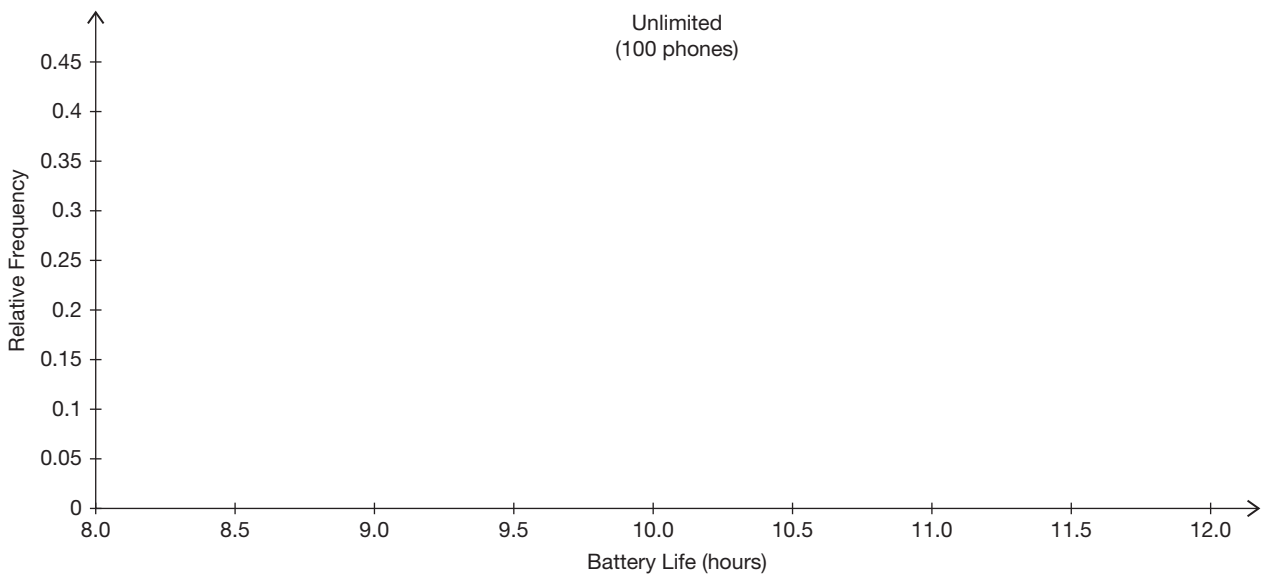
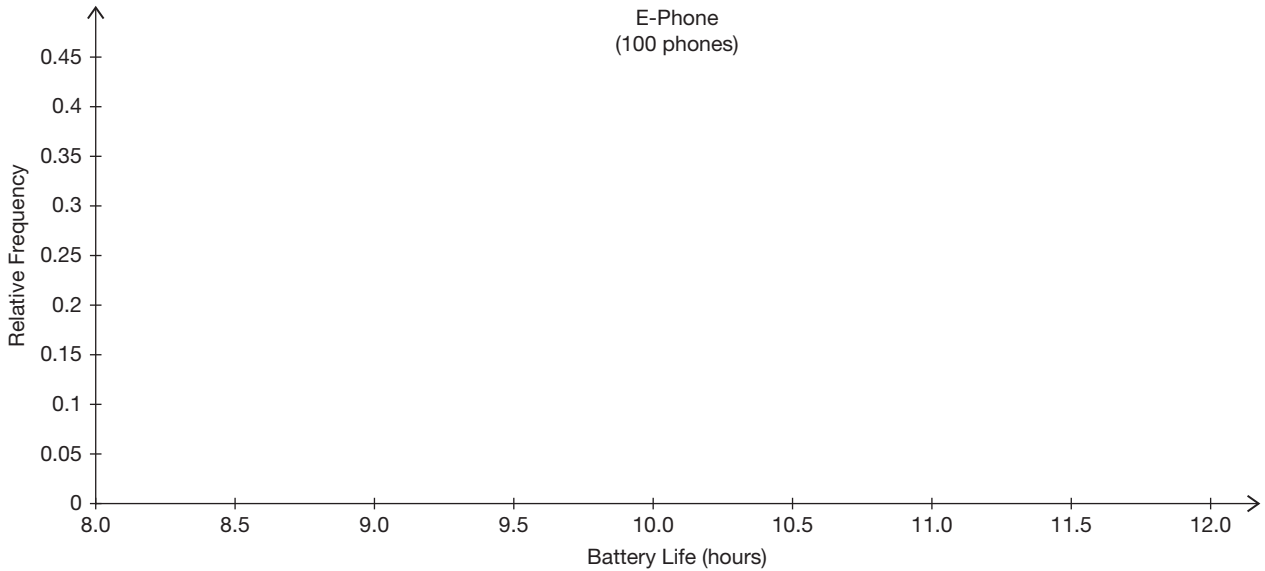
3. Complete the tables by calculating the relative frequency of phones in each interval. Explain how you determined the relative frequencies.

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

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

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

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



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

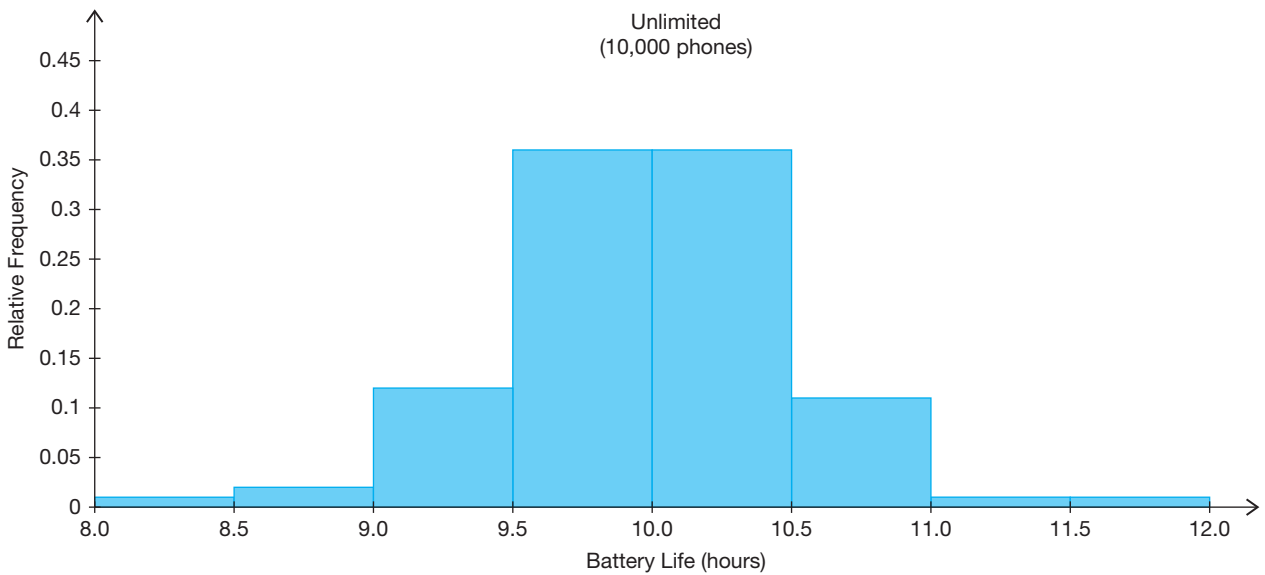
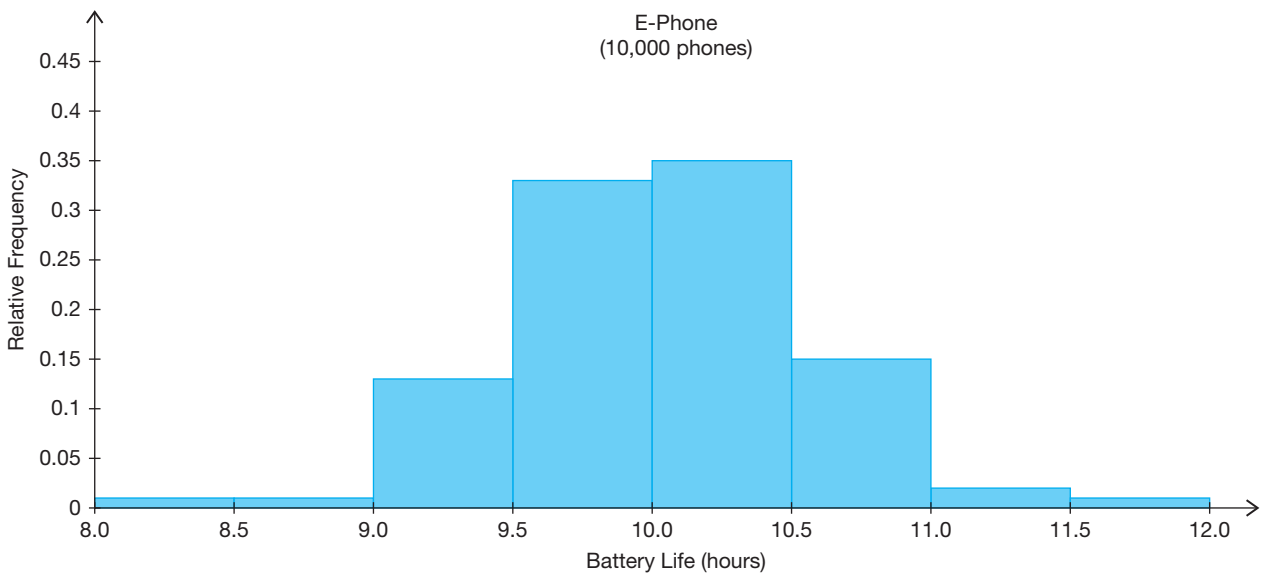


5. Describe the shape and spread of the histograms. What might these characteristics reveal about the data for each company?

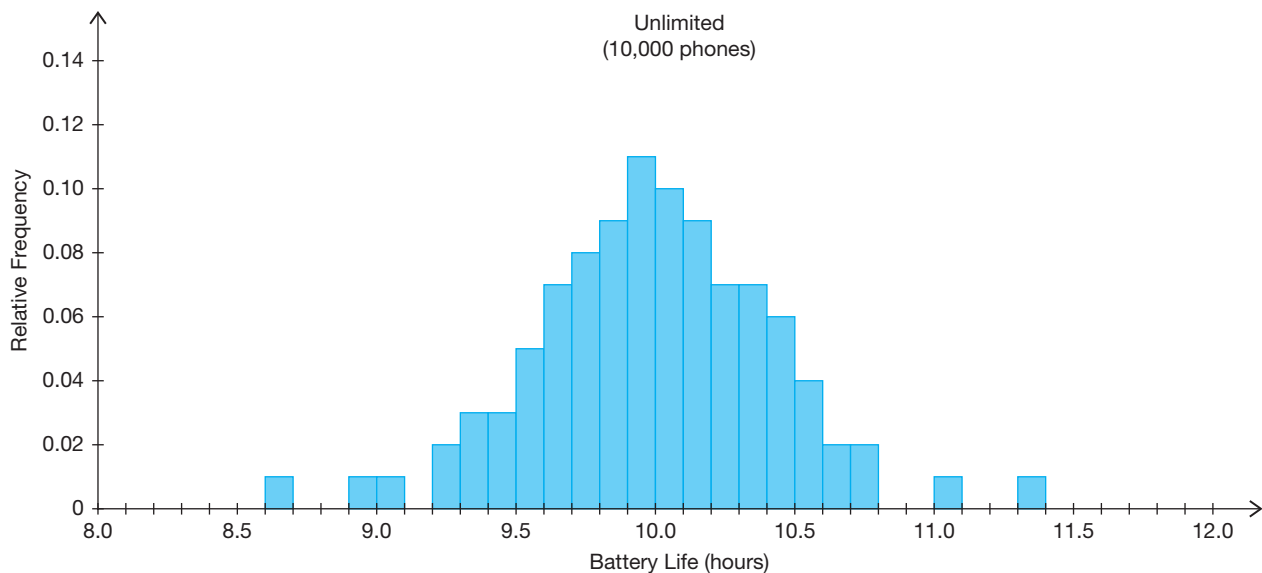
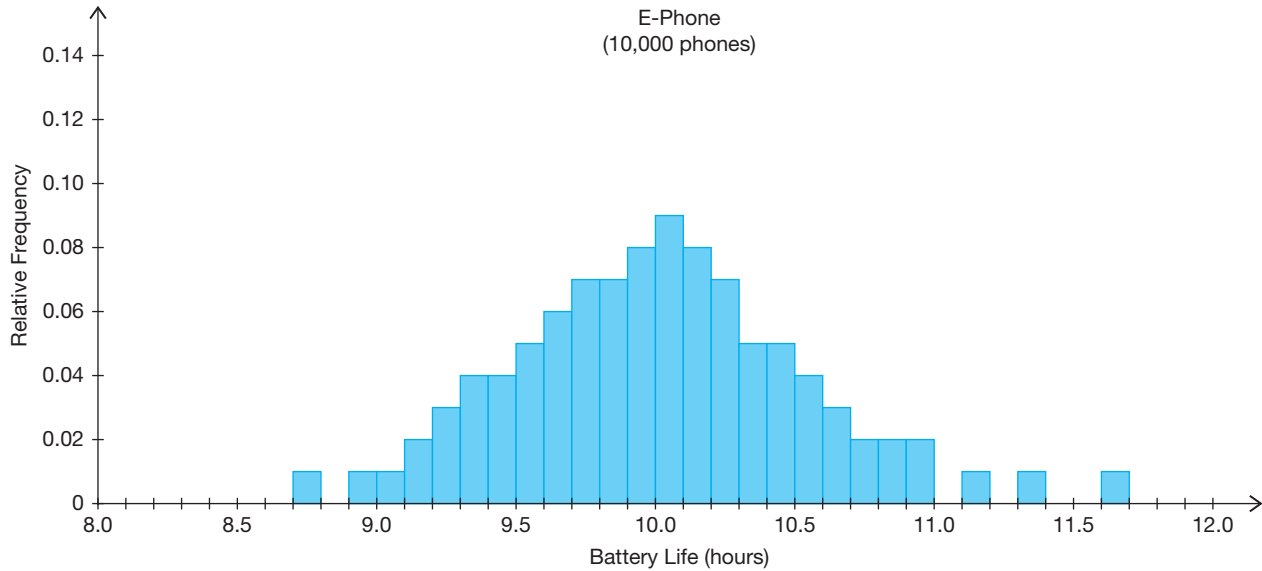
1



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



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



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



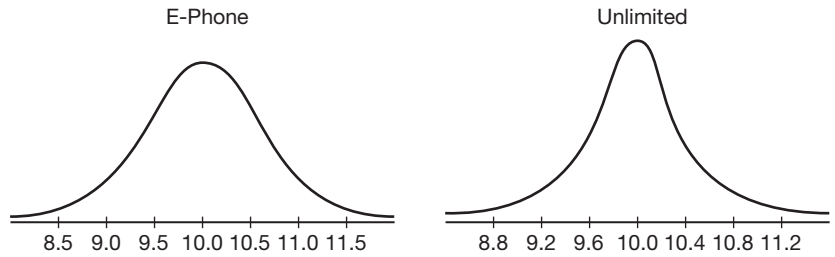


As the sample size continues to increase and the interval size continues to decrease, the shape of each relative frequency histogram will likely start to resemble a *normal curve*. A **normal curve** is a bell-shaped curve that is symmetric about the mean of the data.

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

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

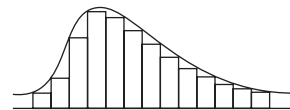
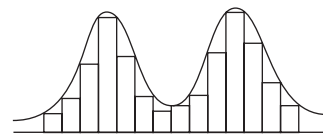
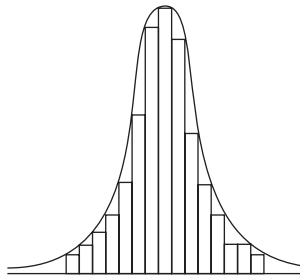
The normal curves for the E-Phone and Unlimited cell phone battery lives are shown. In order to display normal curves for each data set, different intervals were used on the horizontal axis in each graph.



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

**Normal Distributions**

**Not Normal Distributions**

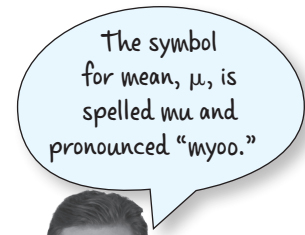


## PROBLEM 2 Deviating Slightly

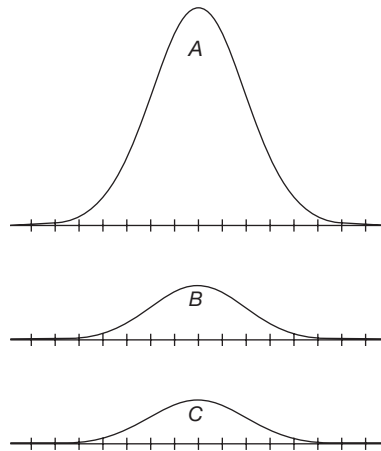


You already know a lot about the mean. With normal curves, the **mean** of a population is represented with the symbol  $\mu$ . The mean of a sample is represented with the symbol  $\bar{x}$ . The **standard deviation** of data is a measure of how spread out the data are from the mean. The symbol used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

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

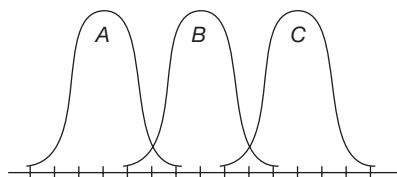


1. Normal curves  $A$ ,  $B$ , and  $C$  represent the battery lives of a population of cell phones of comparable models from three different companies.



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

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



- a. Compare the mean of each company.
- b. Compare the standard deviation of each distribution.



Be prepared to share your solutions and methods.



# #I'mOnline

## The Empirical Rule for Normal Distributions

### LEARNING GOALS

In this lesson, you will:

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

### KEY TERMS

- standard normal distribution
- Empirical Rule for Normal Distributions

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

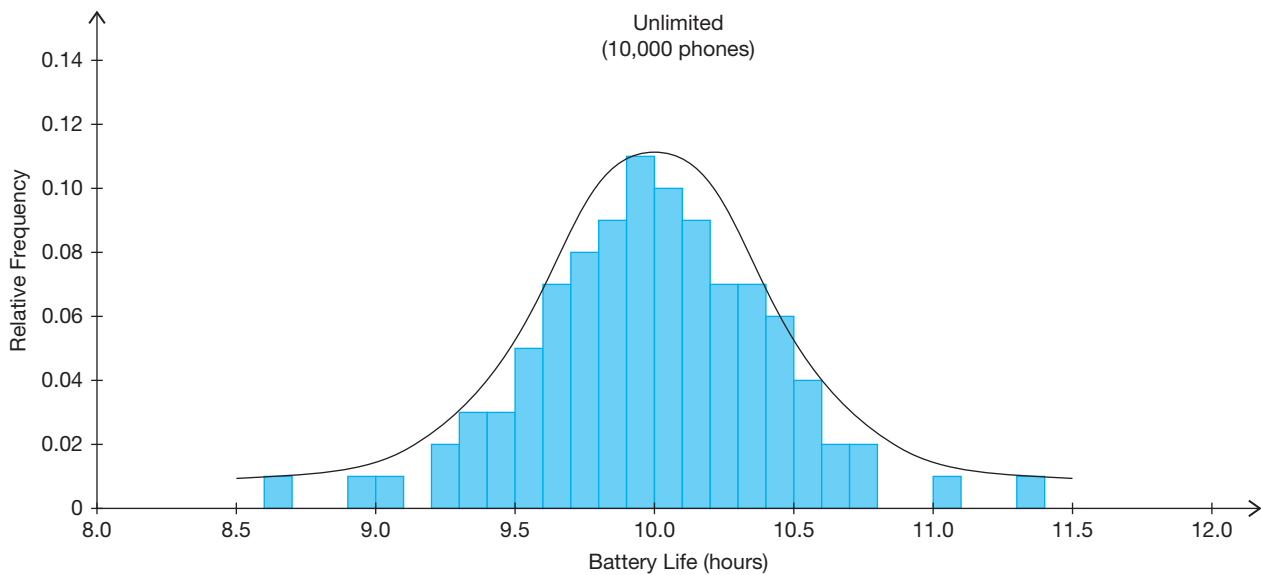
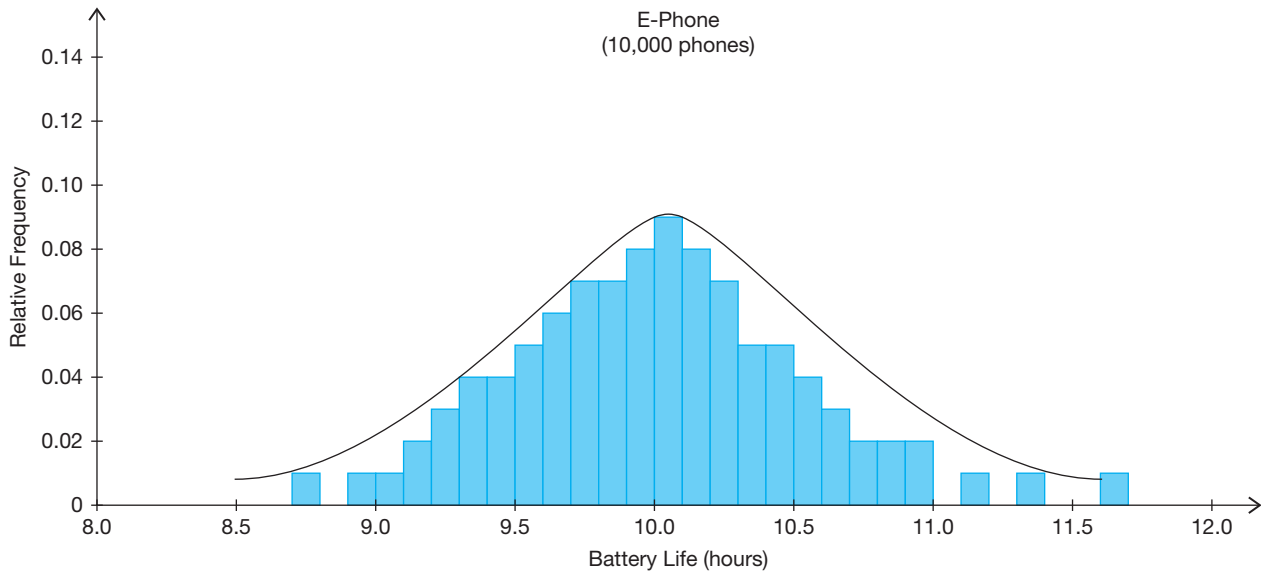
According to some economic models, the crash that occurred on Black Monday represented an event that was 20 standard deviations away from the normal behavior of the market. Mathematically, the odds of a Black Monday-type event occurring were 1 in  $10^{50}$ .

## PROBLEM 1 Count 'em Up



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

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



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

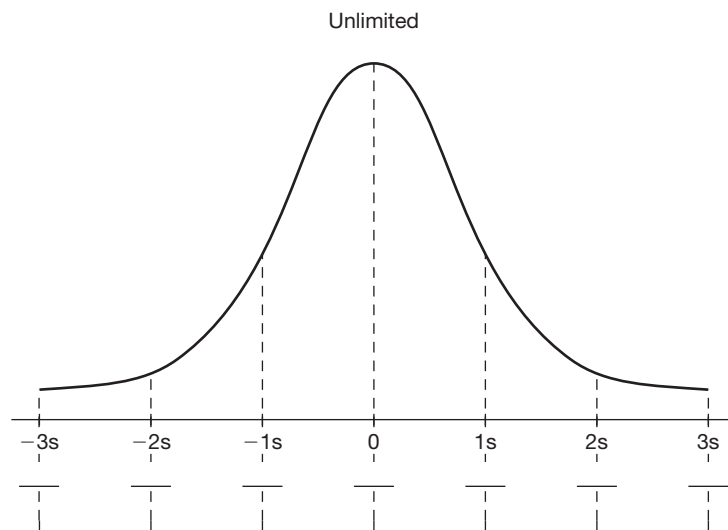
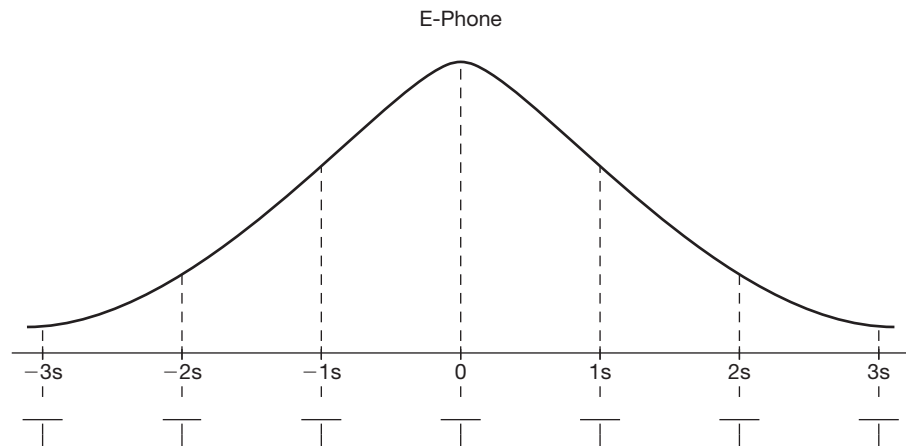


1. Study the graphs shown.
  - a. For each graph, label each standard deviation unit with its corresponding battery life.
  - b. What value is represented at  $s = 0$  for both graphs?

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.



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

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

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



“Within one standard deviation” means between  $-1\sigma$  and  $1\sigma$ , or between  $-1\sigma$  and  $1\sigma$ .

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



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



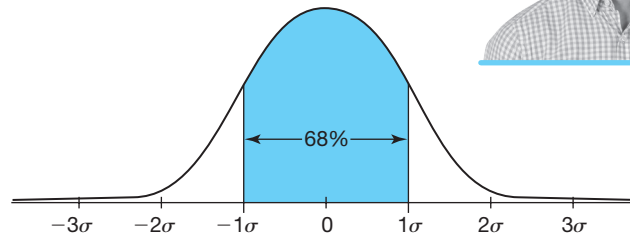


The **standard normal distribution** is a normal distribution with a mean value of 0 and a standard deviation of  $1\sigma$  or  $1s$ . In a standard normal distribution, 0 represents the mean. Positive integers represent standard deviations greater than the mean. Negative integers represent standard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

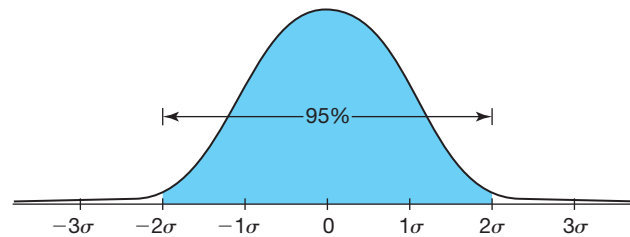
The Empirical Rule for Normal Distributions is often summarized using a standard normal distribution curve because it can be generalized for any normal distribution curve.

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

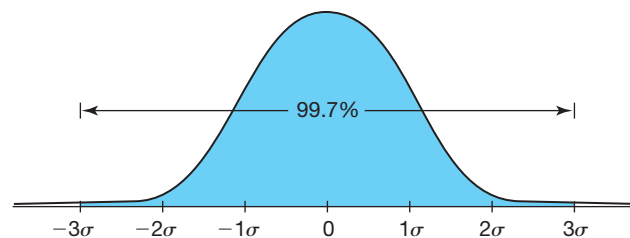
- Approximately 68% of the data in a normal distribution for a population is within 1 standard deviation of the mean.



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



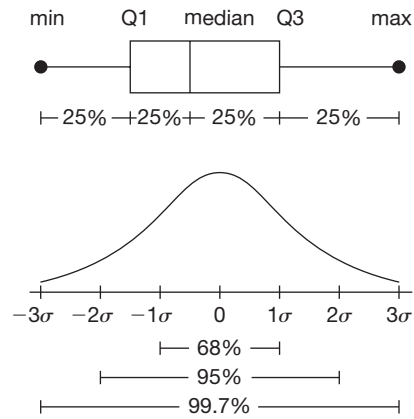
- Approximately 99.7% of the data in a normal distribution for a population is within 3 standard deviations of the mean.



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



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



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



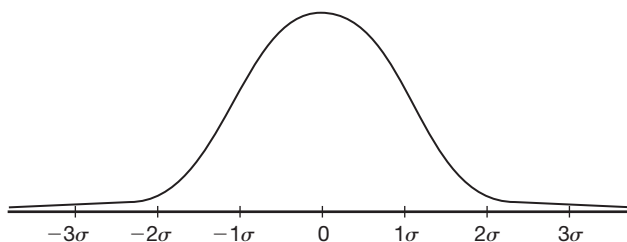
## PROBLEM 2 Interval Training

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

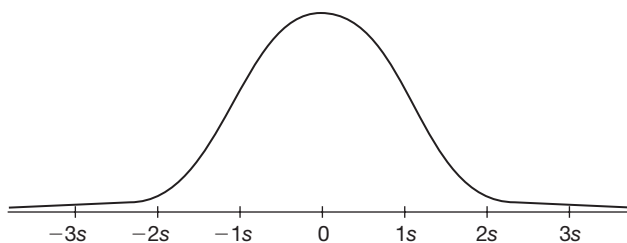


1. Determine each percent and explain your reasoning. Shade the corresponding region under each normal curve. Then tell whether the distribution represents population data or sample data.

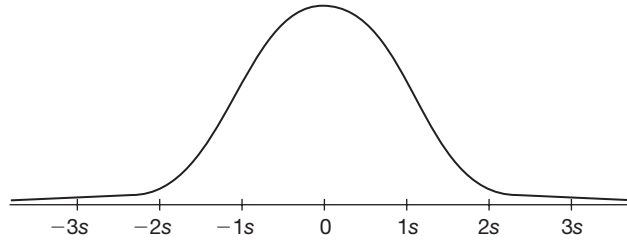
- a. What percent of the data is greater than the mean?



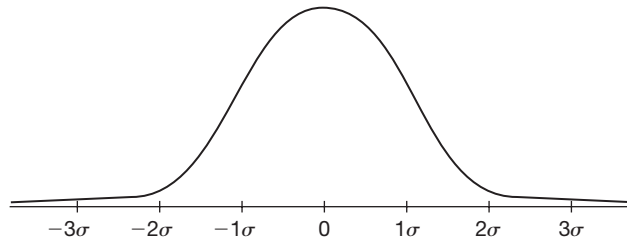
- b. What percent of the data is between the mean and 2 standard deviations below the mean?



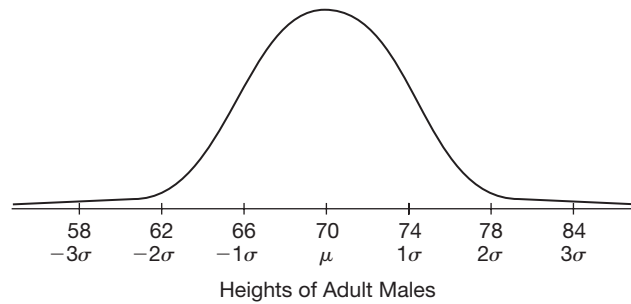
- c. What percent of the data is between 1 and 2 standard deviations above the mean?



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



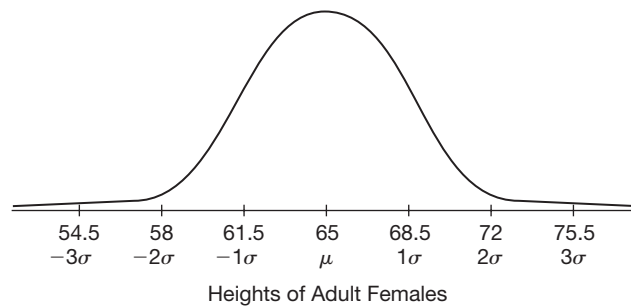
2. Use the normal curve to answer each question and explain your reasoning. Shade the region under each normal curve to represent your answer. Then tell whether the distribution represents population data or sample data.
- a. What percent of adult males have a height between 62 inches and 74 inches?



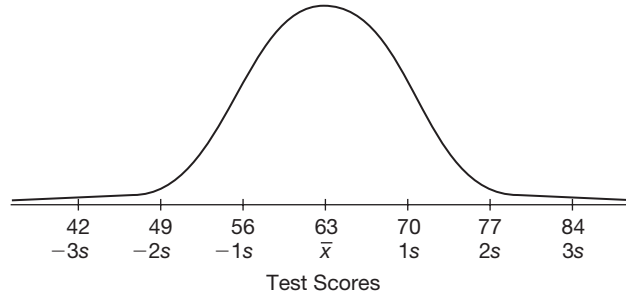
Keep in mind that  $1\sigma$  corresponds to a data value that is one standard deviation greater than the population mean and  $-1\sigma$  corresponds to a data value that is one standard deviation less than the mean.



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



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



Be prepared to share your solutions and methods.

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

## Z-Scores and Percentiles

### LEARNING GOALS

In this lesson, you will:

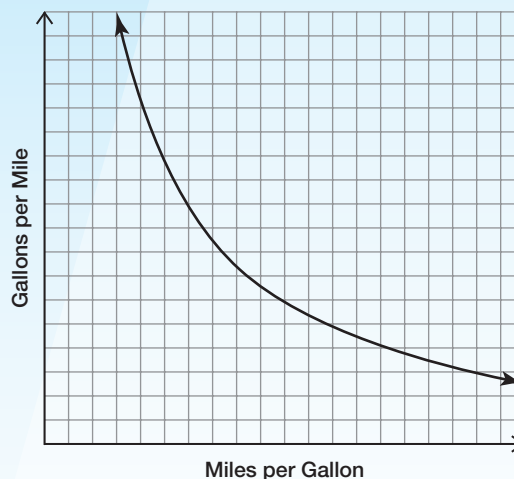
- Use a z-score table to calculate the percent of data below any given data value, above any given data value, and between any two given data values in a normal distribution.
- Use a graphing calculator to calculate the percent of data below any given data value, above any given data value, and between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

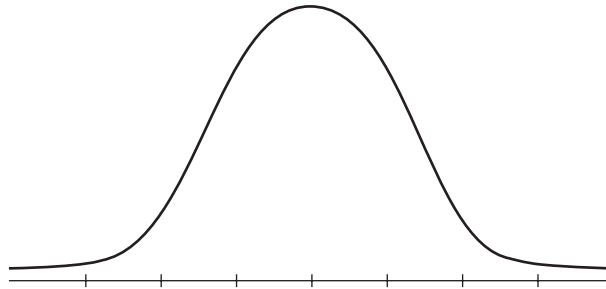


When a car gets a low number of miles per gallon (say, 12 mpg), switching to a slightly higher number (say, 15 mpg) represents a large decrease in the number of gallons per mile, which is a big savings—even bigger than switching from 30 mpg to 50 mpg!

How many gallons per mile does your car get?

**PROBLEM 1** 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.



- b. Determine the percent of hybrid cars that get less than 60 miles per gallon. Explain your reasoning.
- c. Determine the percent of hybrid cars that get less than 66 miles per gallon. Explain your reasoning.
- d. Determine the percent of hybrid cars that get less than 72 miles per gallon. Explain your reasoning.



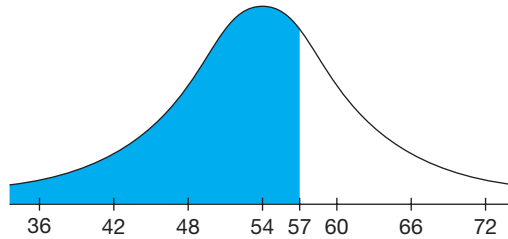




When data values are aligned with integer multiples of the standard deviation from the mean, you can use the Empirical Rule for Normal Distributions to calculate the percent of data values less than that value. But what if a data value does not align with the standard deviations?



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.



- 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.



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.

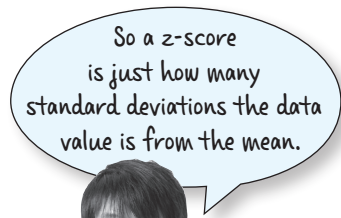
For a population, a z-score is determined by the equation

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

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.



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.

<b>z</b>	<b>0.0</b>	<b>0.01</b>	<b>0.02</b>	<b>0.03</b>	<b>0.04</b>	<b>0.05</b>
<b>0.0</b>	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
<b>0.1</b>	0.5398	0.5433	0.5478	0.5517	0.5557	0.5596
<b>0.2</b>	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
<b>0.3</b>	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
<b>0.4</b>	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
<b>0.5</b>	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 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 **2<sup>nd</sup>** 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  $-1 \times 10^{99}$ , a very small number, for the lower bound.

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



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



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.



**Juan**

```
normalcdf(-9, 0.5,  
0, 1)  
.6914624678
```



**Michael**

```
normalcdf(0, 57,  
54, 6)  
.6914624678
```

- a. Explain why Juan and Carlos used different values but still got the same result.
- 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.



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

```
Josh
normalcdf(-1E99,
56,54,6)
.6305585963
```

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

## PROBLEM 2 More or Less . . .



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

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.

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.



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.

### 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 50<sup>th</sup> percentile for this data set. Explain your reasoning.
2. Would a teen in the 90<sup>th</sup> percentile send and receive more or fewer than 100 text messages per day? Explain your reasoning.
3. Would a teen in the 10<sup>th</sup> percentile send and receive more or fewer than 100 text messages per day? Explain your reasoning.



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	0.5875	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



# 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.

**Y**ou 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.

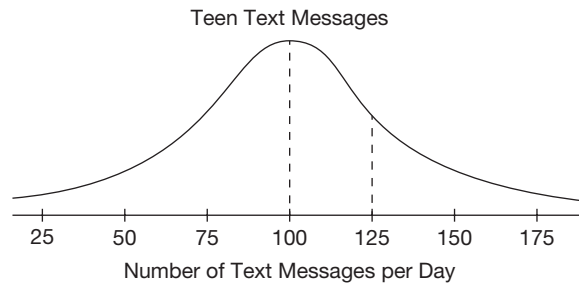
Now for the important question: What would you call this hybrid plant?

**PROBLEM 1** 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.
2. Determine the probability that the randomly selected teen sends and receives fewer than 75 text messages per day.



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



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

---



1. 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?



Be prepared to share your solutions and methods.

# 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 ( $\mu$ ) (1.1)
- standard deviation ( $\sigma$ ) (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

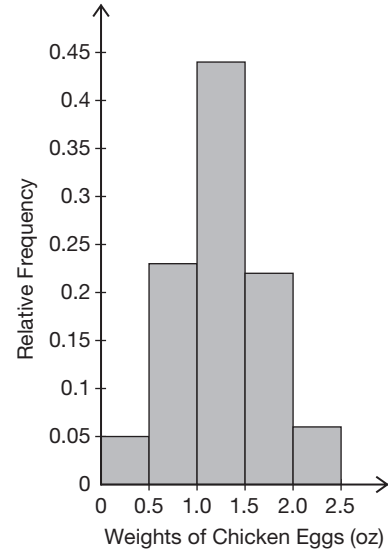
The heights of basketball players are examples of continuous data.

## 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

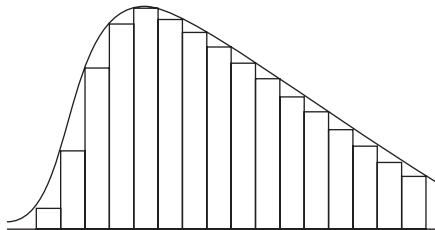


## 1.1 Recognizing the Difference Between Normal Distributions and Non-normal Distributions

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.





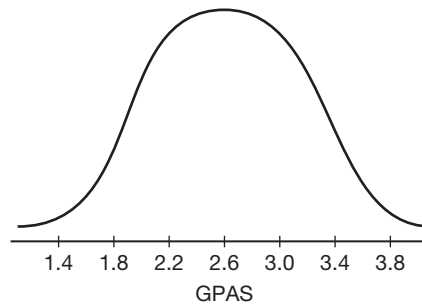
## Recognizing and Interpreting Properties of a Normal Curve and a Normal Distribution

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  $\mu$ , 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  $\sigma$ .

### Example

The mean is 2.6 and the standard deviation is 0.4.

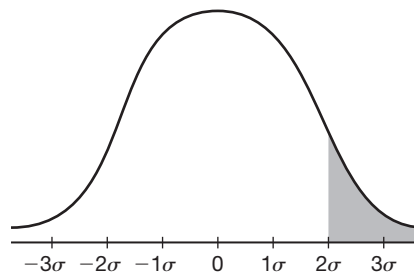


## Recognizing the Connection Between Normal Curves, Relative Frequency Histograms, and the Empirical Rule for Normal Distributions

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 `normalcdf(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.

`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 80<sup>th</sup> 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 80<sup>th</sup> 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 45<sup>th</sup> 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.

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

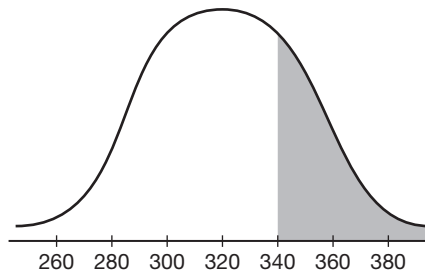
The length of time that represents the 45<sup>th</sup> percentile is approximately 144.12 minutes.

### 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.

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%.

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

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%.

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

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

