**Algebra 2 B**

**Unit 6: Modeling Data Distributions**

**Unit Summary:**

In this unit, you will learn about a type of data distribution that compares each data point to the mean of the dataset and visually represents how many values can be expected to fall within a specific range.

You will use the number of standard deviations away from the mean to describe what percentage of data can be expected to fall within a certain range.

Some data distributions will resemble each other by following a normal distribution, and others will vary, showing a skewed distribution.

In either case, you will be able to:

* calculate how far a data value is from the mean
* estimate the area under the curve showing the distribution
* explain how the distribution models a given data set

**Lesson 2 – Shape, Center, and Spread**

**Key Words:**

* **box and whisker plot** – a graphic representation of a distribution by a rectangle, the ends of which mark the maximum and minimum values, and in which the median and first and third quartiles are marked by lines parallel to the ends
* **frequency table** – an arrangement of statistical data that exhibits the frequency of the occurrence of the values of a variable
* **histogram** – a representation of a frequency distribution by means of rectangles whose widths represent class intervals and whose areas are proportional to the corresponding frequencies
* **mean** – the average of a set of data, found by adding all numbers in the data set and dividing the sum by the number of numbers
* **mode** – the value or values that occur most often in a data set
* **multimodal** – having more than one modal value
* **skewed** – data that has a longer tail on one side than on the other, making the distribution asymmetric
* **skewed distribution** – a type of data distribution in which the data is clustered on the right or left ends when displayed in a histogram or dot plot
* **standard deviation** –a measure of the spread of a data set indicating the typical distance a value in the data set is from the mean
* **symmetrical distribution** – a type of data distribution in which the left side of the distribution mirrors the right side when displayed in a histogram or dot plot
* **uniform** – a distribution with equally likely outcomes, also called a rectangular distribution
* **unimodal** – having a single mode

**Objective 1:** In this lesson, you will distinguish between symmetric and skewed distributions.

*Mathematical Practice Standard: Make sense of problems and persevere in solving them.*

**Big Ideas**:

* A distribution is called symmetric when the data on either side of a distribution has approximately the same shape.
  + This frequency table and histogram show symmetrical distributions:



Notice that if the histogram is folded in half along its middle vertical line, two halves of the histogram would roughly match.

The graph shows that the data set has a **single mode** and that it is in the center of the graph.

The data in this set cluster near the mode.

* Some symmetric sets of data do not cluster around a single, central mode. There may be several peaks in the data, the data might be multi-modal and there might be gaps in the dispersion of data in the data set.

|  |  |
| --- | --- |
|  | This graph is symmetric because if you folded it in half along its middle vertical line, two halves of the graph would roughly match.  It is multi-modal, having more than one modal value.  There is a gap in the dispersion of data because the number cube only has the numbers 1,3,5. Therefore there cannot be any outcomes of 2 and 4. |
|  | This graph is symmetric because if you folded it in half along its middle vertical line, two halves of the graph would roughly match.  This graph has a single mode at the center of the graph.  There are gaps in the dispersion of the data set because the results show the sum of two number cubes with 1,3,5. You’ll notice the gaps are odd numbers because no results of the two number cubes can add up to 3, 5, 7 or 9. |
|  | This graph is symmetric because if you folded it in half along its middle vertical line, two halves of the graph would roughly match.  This graph has a uniform peak with equally likely outcomes.  It can also be called a rectangular distribution. |

* Skewed data is not symmetric.

|  |  |
| --- | --- |
|  | This graph is not symmetric.  This graph is **skewed right** because the distribution of data has an elongated tail on the right.  The distribution of the data is mostly on the left. |
|  | This graph is not symmetric.  This graph is **skewed left** because the distribution of data has an elongated tail on the left.  The distribution of the data is mostly on the right. |

* When describing the shape of a distribution, look for the following characteristics:
  + mode – single mode, multi-modal, uniform
  + symmetry
  + skewed

**Objective 2: In** this section, you will interpret the mean as a typical value of a distribution and the standard deviation as a typical distance a value is from the mean.

*Mathematical Practice Standard: Make sense of problems and persevere in solving them.*

**Big Ideas:**

* The standard deviation of a distribution is a measure of how much the values in a data set are spread out from the mean.
  + It can be interpreted as the distance a typical value or data point is from the mean.
  + A distribution with a relatively **small standard deviation** will have a large number of its values close to the mean. This results in a narrower histogram that has a high number of values around the mean and fewer at the edges.
  + A distribution with a relatively **large standard deviation** will instead have more values which are farther from the mean. This results in a wider histogram, as the data is more spread out.
* Before you can determine the standard deviation of a distribution, you must first know its mean.
  + The mean of a data set is calculated by adding all its values together and dividing the sum by the number of values.
  + Use this formula to calculate the mean:

|  |
| --- |
| is the mean |
| is the sum of all values |
| represents the number of values |

* Use this formula to calculate the standard deviation:
  + For each value in the data set, subtract the mean from the value, then square the difference.
  + Add all the results together and divide by the number of values in the set.
  + Take the square root to find the standard deviation.
* Example: Find the standard deviation of the following data set: 12,14,14,15,15,15,16,19

|  |  |
| --- | --- |
| Step 1 – find the mean | The mean for this data set is 15. |
| Step 2 – Use the standard deviation formula ; start by subtracting the mean from each value |  |
| Step 3 – square the results |  |
| Step 4 – find the sum |  |
| Step 5 – complete the division | = |
| Step 6 – find the square root | The standard deviation for this set is 1.87. This is interpreted to mean that a typical value of the data set is within 1.87 units of the mean, which is 15. |

**Objective 3:** In this section, you will discuss estimates of the mean and standard deviation, and their reasonableness, for symmetric and skewed distributions.

*Mathematical Practice Standard: Use appropriate tools strategically.*

**Big Ideas:**

* For perfectly symmetrical distribution, the mean, , is the number for which the shape of distribution to the left of exactly mirrors the shape of the distribution to the right of The mean is the number halfway between the minimum and maximum values and can be found by the formula:
  + This formula can also be used when a distribution is *nearly* symmetrical, especially if outliers are ignored. The mean can be estimated by trying to judge the value for which the shape of the distribution to the left of the estimated mean is an approximate mirror image of the shape of the distribution to the right of the estimated mean.
* When estimating a left-skewed distribution, the following statements are true:
  + Data is clustered on the right side of the display and tails off to the left.
  + The mode is the value with the greatest frequency.
  + The median is typically to the left of the mode.
  + The mean is typically to the left of the median.
* When estimating a right-skewed distribution, the following statements are true:
  + Data is clustered on the left side of the display and tails off to the right.
  + The mode is the value with the greatest frequency.
  + The median is typically to the right of the mode.
  + The mean is typically to the right of the median.
* When estimating the standard deviation for a symmetrical distribution:
  + First estimate the mean,.
  + Next determine a distance, , from either side of the estimated mean so that the distance contains about one-half to two-thirds of the data values, ex . This distance is the estimated standard deviation.
    - The number of data values contained in the intervals on either side of the estimated mean should be equal or approximately equal.
* When estimating the standard deviation for a skewed distribution:
  + First estimate the mean,.
  + Next consider different values for the standard deviation until you are reasonably convinced that the interval contains one-half to two-thirds of the data values.
    - Since the data set is skewed, the number of data values contained in the intervals to either side of the estimated mean will not be equal. They should represent about one-half to two-thirds of the data values.

**Lesson 4 – The Normal Curve**

**Key Words:**

* **bimodal** – a data set with two modes
* **mean** – the average of a dataset found by adding all of the data values and dividing by the number of data values in the dataset
* **mode** – the value or values that occur most often in a data set
* **multimodal** – a data set with two or more modes
* **normal distribution** – a type of probability distribution in which the graph is a bell-shaped curve that is symmetric about the mean
* **standard deviation** – a measure of the spread of a data set indicating the typical distance a value in the data set is from the mean
* **unimodal** – a data set with only one mode

**Objective 1:** In this section, you will use a smooth curve to model a data distribution.

*Mathematical Practice Standard: Reason abstractly and quantitatively.*

**Big Ideas**:

* Drawing a smooth curve over the tops of bars in a histogram allows us to find patterns in the shape of the data.

|  |  |
| --- | --- |
|  | * The curve has one peak near the center indicating there is one mode for the data. * The curve has a tail on each end where the frequency is lower. * The curve appears to be nearly symmetric on each side. This tells us that the mean, median and mode are all close to each other and located at the peak of the curve. |
|  | * The curve has two peaks, indicating the data is bimodal.   + (If there are more than two peaks, the data would be multimodal.) |
| Skewed right | * The curve shows the peak is closer to the left side. This is called a right-skewed distribution. * The mode is still at the peak. * The mean will fall to the right of the peak. This is because the measure of center is pulled to the right by the additional data to the right of the mode, also referred to as the elongated tail. |
|  | * The curve shows the peak is closer to the right side. This is called a left-skewed distribution. * The mode is still at the peak. * The mean will fall to the left of the peak. This is because the measure of center is pulled to the left by the additional data to the left of the mode, also referred to as the elongated tail. |

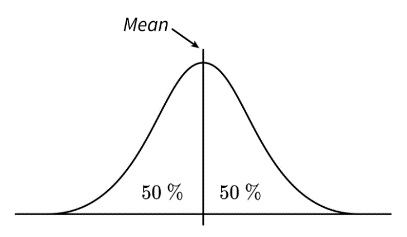
* Modeling a distribution as a smooth curve allows for a deeper analysis of the distribution. Curves can be compared to each other to draw conclusions about relative modes and standard deviations.

**Objective 2:** In this section, you will learn about the attributes of a normal curve and how to use them to describe different probability distributions.

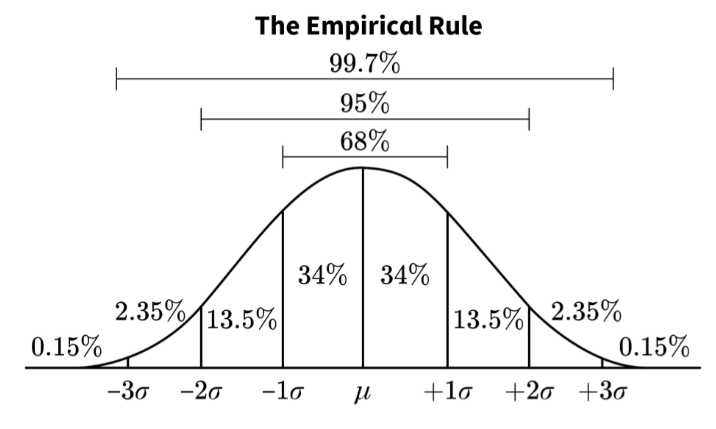
*Mathematical Practice Standard: Construct viable arguments and critique the reasoning of others.*

**Big Ideas:**

* Here are some attributes of a normal curve:
  + The curve has one peak, or is unimodal, and symmetric about the center.
  + The curve is defined by the mean and standard deviation of the data.
  + The total area under the curve is equal to 100%. Each side of the mean is 50%.



* The Empirical Rule, sometimes called the 68-95-99.7 rule is a guideline to how data is spread out in a normal distribution. The rule states:
  + Approximately 68% of the data values will lie within one standard deviation of the mean.
  + Approximately 95% of the data values will lie within two standard deviations of the mean.
  + Approximately 99.7% of the data will lie within three standard deviations of the mean.



**Objective 3:** In this section, you will recognize when it is reasonable to use a normal curve as a model for a distribution.

*Mathematical Practice Standard: Look for and express regularity in repeated reasoning.*

**Big Ideas:**

* A normal distribution has three key characteristics:
  1. It has an approximately symmetric bell shape.
  2. The mean is near the center of the distribution.
  3. The Empirical Rule closely fits.
* Let’s use those three key characteristics to determine when it’s reasonable to use a normal curve as a model for a distribution:

|  |  |  |
| --- | --- | --- |
| 1. It has an approximately symmetric bell shape. |  | 1. Yes, the histogram is approximately bell shaped. |
| 2.The mean is near the center of the distribution. | 2. Yes, the mean is 68.08 and is located at the center of the distribution. |
| 3. The Empirical Rule closely fits. | The standard deviation is 2.40. This means that on average, the heights in the data set vary by 2.40 inches from the mean of 68.08.  Calculate three standard deviations from the mean to the left and to the right:  1 standard deviation to the left is: 68.08- 1(2.40) =65.68  2 standard deviations to the left is: 68.08-2(2.40) =63.28  3 standard deviations to the left is: 68.08-3(2.40) =60.88  1 standard deviation to the right is: 68.08+ 1(2.40) =70.48  2 standard deviations to the right is: 68.08+2(2.40) =72.88  3 standard deviations to the right is: 68.08+3(2.40) =75.28 | 3. Yes, the Empirical Rule closely fits.  100% of the data falls within 3 standard deviations of the mean. (Empirical Rule states 99.7%)  92% of the data falls within 2 standard deviations of the mean. (Empirical Rule states 95% and 92% is close to 95%)  68% of the data falls within 1 standard deviations of the mean. (Empirical Rule states 68%) |

**Practice Questions and Answers**

|  |  |  |
| --- | --- | --- |
|  | Question | Answer |
| P 1 | When a smooth curve is used to model a data distribution, which measure describes the spread of the curve? | the standard deviation |
| P 2 | *Use the image to answer the question.*    Which curve has the lowest mode? | A |
| P 3 | Within a population, IQ score has a normal distribution with a mean of and a standard deviation of . Apply the Empirical Rule to find the probability that an individual has an IQ score above 130. | 2.5% |
| P 4 | *Use the images to answer the question.*  Which of these data distributions can be modeled by a normal curve?  Image A:    Image B: | Neither A nor B |
| P 5 | *Use the images to answer the question.*  Which data distribution can be modeled by a normal curve?  Option #1    Option #2: | 1 |

**Quick Check Questions and Answers**

|  |  |  |
| --- | --- | --- |
|  | Question | Answer |
| Q 1 | *Use the image to answer the question.*    Which curve has the highest standard deviation? | C |
| Q 2 | *Use the image to answer the question.*    Janet is using two curves to model data from two different data distributions. The curve labeled A represents data from group A, and the curve labeled B represents data from group B. Based on the shapes of the curves, what conclusions can she draw about the two groups? | Group A has a higher standard deviation, and group B has a higher mode. |
| Q 3 | Which of the following describes an attribute of the normal curve? | The normal curve is symmetric about the center. |
| Q 4 | *Use the image to answer the question.*    The graph shows the normal curve describing the distribution of the heights of 10-story buildings from cities around the world. What is the probability that a 10-story building is shorter than 35 meters? | 84% |
| Q 5 | Which of the following lists the criteria for using a normal curve as a model for a distribution? | The data values follow the Empirical Rule, the mean is in the center of the data, and the distribution is bell-shaped. |

**Lesson 5 – Area Under a Normal Curve**

**Key Words:**

* **normal distribution** – a type of probability distribution in which the graph is a bell-shaped curve that is symmetric about the mean
* **z-score** – a statistical measurement that describes the number of standard deviations a given data point lies above or below the mean

**Objective 1:** In this section, you will calculate z-scores.

*Mathematical Practice Standard: Use appropriate tools strategically.*

**Big Ideas**:

* A z-score is a statistical measurement that describes the number of standard deviations a given data point lies above or below the mean.
  + z-scores are a way to compare results to a normal population.
* Use this formula when calculating a z-score:
  + is the data point; is the mean; is the standard deviation
* Here is how to interpret z-scores:
  + A z-score of less than 0 represents a data point less than the mean.
  + A z-score greater than 0 represents a data point greater than the mean.
  + A z-score equal to 0 represents a data point equal to the mean.
  + A z-score equal to 1 represents a data point which is 1 standard deviation greater than the mean; a z-score equal to 2 represents a data point which is 2 standard deviations greater than the mean; and so on.
  + A z-score equal to -1 represents a data point which is 1 standard deviation less than the mean; a z-score equal to -2 represents a data point which is 2 standard deviations less than the mean; and so on.
  + Any data greater than or less than two standard deviations from the mean is considered an unusual data value from the population.

**Objective 2:** In this section, you will use z-scores to estimate the area under a normal curve.

*Mathematical Practice Standard: Use appropriate tools strategically.*

**Big Ideas:**

* The areas under all normal curves are related. Area refers to “area percentage”.
  + Ex: the area percentage to the right of one standard deviation above the mean is the same for all normal curves.

Figure 1:

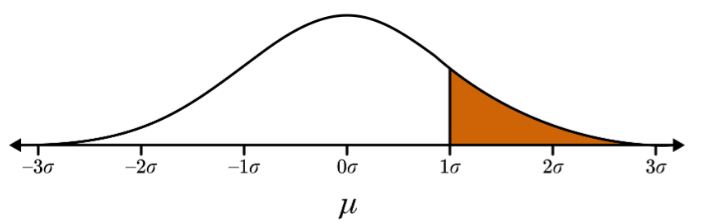
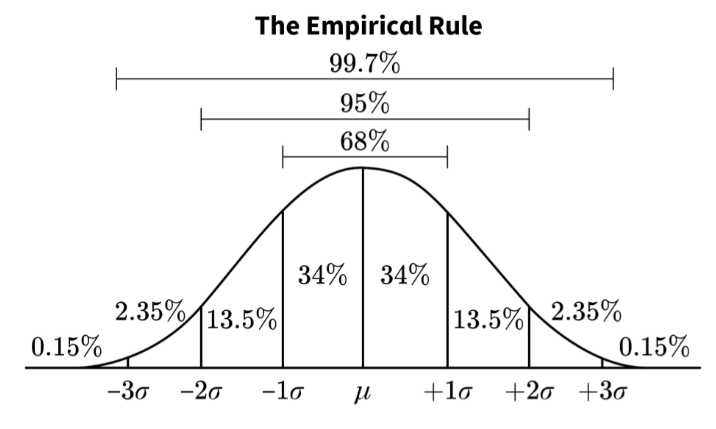
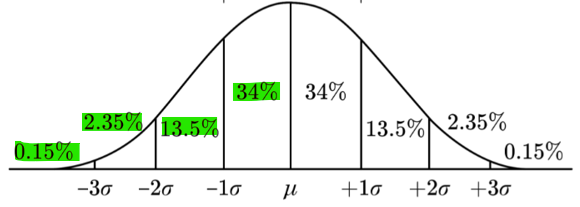
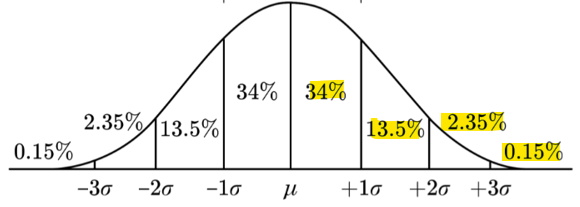


Figure 2:



* + Using the percentage of the Empirical Rule image, the shaded area under the curve in Figure 1 can be calculated.
    - 13.5% + 2.35% + 0.15% = 16%
    - Therefore, 16% of the curve is shaded to the right of one standard deviation above the mean.
* Not all areas under a normal distribution will fall within the 68% - 95% - 99.7% subdivisions.
  + This is where z-scores come in handy. Z-scores are used to calculate the area percentages anywhere along a standard normal distribution curve.
* The area percentage calculated using a z-score will be a decimal value between 0 and 1. The total area under any normal curve is 1, which is equivalent to 100%.
  + The area on either side of the mean is .5, which is equivalent to 50%.
    - 34% + 13.5% + 2.35% + 0.15% = 50%
    - written in decimal form: .34 + .135 + .0235 + .0015 = .5

* A z-score table shows the percentage of values to the left and right of a given z-score on a normal distribution.
* How to use a z-score table:
  + Step 1: find the z-score of the data value.
    - Sometimes this is provided, if not use the formula:
  + Step 2: use the z-score table to find the area.
  + Step 3: convert the decimal from the table into a percentage.
  + Step 4: check to confirm your answer is reasonable.
  + Ex: Estimate the area percentage under a normal curve to the left of a z-score of -1.63

|  |  |
| --- | --- |
| Step 1: find the z-score; this is provided in the question. | -1.63 |
| Step 2: use the z-score table to find the area: | Start by finding -1.6 row in the first z-score column.  Next find .03 in the columns along the top of the table.  Lastly, follow the -1.6 row over to the .03 column to find the area value.  The area value is .0516 |
| Step 3: convert the area decimal into a percentage |  |
| Step 4: check to confirm your answer is reasonable. | A z-score equal to -1 represents a data point which is 1 standard deviation less than the mean.  You should be visualizing the area of a bell shaped curve that is more than 1 standard deviation to the left of the mean.    5.16% is a reasonable answer |

**Objective 3:** In this section, you will use the area under normal curves to calculate probabilities of events.

*Mathematical Practice Standard: Reason abstractly and quantitatively.*

**Big Ideas:**

* The Empirical Rule can be used to calculate the probability of an event when an event falls exactly one, two or three standard deviations from the mean.
* Using z-scores, the probability of events at any distance from the mean can be calculated.
  + To find out how likely it is for a number from a group of data to be in a certain range, we use z-scores. These z-scores help us compare the numbers in the group.
  + No matter the z-score or the section of the area under the curve you are calculating, area always accumulates from the left.
  + Follow these steps to determine the probability:
    - Step 1: Sketch the normal model
    - Step 2: find the z-score using the formula:
    - Step 3: Use the z-score table to find the area
    - Step 4: Find the probability (by converting the decimal from the table into a percentage
* Calculating probabilities with two events, or z-scores follows this general form: area left of the greater z-score – area left of the lesser z-score
  + To calculate probabilities with two events, use the first 3 steps listed above to calculate the z-scores of each event.
  + After calculating the z-scores of the events separately, calculate the z-score of the events combined by subtracting the area left of the greater z-score from the area left of the lesser z-score.
  + Lastly, convert the result after subtracting the z-scores into a percentage to find the probability of the combined event.

**Practice Questions and Answers**

|  |  |  |
| --- | --- | --- |
|  | Question | Answer |
| P 1 | The average playing time of albums in a certain collection is 27 minutes, with a standard deviation of 4 minutes. Calculate the z-score, to the nearest thousandth, for an album that has a playing time of 29.5 minutes. | 0.625 |
| P 2 | Which of the following z-score values would indicate a data point that is greater than the population mean?  Option #1:  Option #2:  Option #3: | 3 |
| P 3 | Estimate the area percentage under a normal curve to the left of a z-score of .  \_\_\_\_\_ percent | 34.46 |
| P 4 | *Use the table to answer the question.*  The top-left cell of a z-score table is labeled z. The remaining cells in the header row have values ranging from 0.00 to 0.09 in increments of 0.01. Below z, the first column has values ranging from 0.3 to 2.1 in increments of 0.1.  A factory makes aluminum cans that will be filled with soda. The walls of the cans have a mean thickness of 0.1 mm and a standard deviation of 0.02 mm. What is the probability that a randomly selected can will have walls that are less than 0.118 mm in thickness?  The probability that the thickness of the walls of a randomly selected can will be less than 0.118 mm is \_\_\_\_\_ percent. | 81.59 |
| P 5 | *Use the table to answer the question.*    The mean height of adult males in the United States is 69 inches with a standard deviation of 2.5 inches. What is the probability that a randomly-selected adult male will have a height between 69.3 inches and 71.2 inches? | 26.28% |

**Quick Check Questions and Answers**

|  |  |  |
| --- | --- | --- |
|  | Question | Answer |
| Q 1 | The average number of red blood cells in an adult male is 6.2 million cells per microliter, with a standard deviation of 0.6 million cells per microliter. If Sergio’s blood contains 6.7 million cells per microliter, which option correctly calculates and interprets his z-score? | Sergio’s red blood cell count is 0.833 standard deviations above the mean. |
| Q 2 | Given the mean of a data set is 254 and has a standard deviation of 12, which of the following data points would result in a z-score that indicates the data point is 2 standard deviations below the mean? | 230 |
| Q 3 | Which of the following percentages accurately estimates the area under a normal curve to the left of a z-score of 1.23? | 89.07 percent |
| Q 4 | Which of the following percentages accurately estimates the area under a normal curve between a z-score of −0.11 and 2.43? | 53.63 percent |
| Q 5 | *Use the table to answer the question.*  The top-left cell of a z-score table is labeled z. The remaining cells in the header row have values ranging from 0.00 to 0.09 in increments of 0.01. Below z, the first column has values ranging from negative 1.7 to negative 0.0 in increments of 0.1.  The mean temperature during the summer in a certain city is 80 degrees Fahrenheit, with a standard deviation of 5 degrees Fahrenheit. What is the probability that a randomly selected day has a temperature below 73 degrees Fahrenheit? | 8.08% |

**Lesson 6 – Modeling with Normal Distributions**

**Key Words:**

* **normal curve** – a frequency curve in which the graph is a bell-shaped curve that is symmetric about the mean
* **normal distribution** – a type of probability distribution in which the graph is a bell-shaped curve that is symmetric about the mean
* **normal model** – a model of a normal distribution that gives the mean and standard deviation of the set of data, expressed as N (mean, standard deviation)

**Objective 1:** In this section, you will use technology to estimate the area under a normal curve.

*Mathematical Practice Standard: Use appropriate tools strategically.*

**Big Ideas**:

* Use a graphing calculator to find probabilities when:
  + a *z*-table is not available
  + the values are beyond the limits of the *z*-table
  + the x-value is beyond one, two or three standard deviations away from the mean
  + precision is needed and estimation is not helpful
* Use the Probability Calculator within the lesson to find the probability that a random x-value is
  + between two values
  + greater than a value
  + less than a value

|  |  |
| --- | --- |
| Step 1: Identify the mean, standard deviation, the lower bound and the upper bound of a dataset. (These **must** be identified in order to use the calculator.) | If the lower bound and/or upper bound values are not stated, using a very large or very small number for the bound will be close enough in terms of giving an accurate probability.  Ex: for the lower bound and for the upper bound |
| Step 2: Click on the calculator at the top of the lesson page | This will appear after clicking on the Probability Calculator:    The calculator is automatically set to:   * a normal distribution * mean ( of 0 * standard deviation ( of 1 * lower bound value of -1 * upper bound value of 1 |
| Step 3: Enter values for mean, standard deviation, lower bound and upper bound | Click in one of the areas where the values need to be added. A keyboard icon will appear in the bottom left of the calculator box.    Click on the icon and the keyboard will appear:    Enter each value: |
| Step 4: Calculate the probability  The calculator automatically calculates and the bell curve updates as each value is entered. After all values are entered, the probability is displayed and needs to be converted into a percent by multiplying it by 100. | This image shows the answer to Example 1 a: The number of hours people spend traveling for the holidays is normally distributed, with a mean of 4.5 hours and a standard deviation of 0.75 hours. Find the probability that a person chosen at random spends from 2.5 to 5 hours traveling for the holidays.  The probability of this scenario is 74.62% |

**Click on any part of the lesson page outside of the calculator to close it.**

**Objective 2:** In this section, you will model data distributions with appropriate normal distributions.

*Mathematical Practice Standard: Make sense of problems and persevere in solving them.*

**Big Ideas:**

* When a dataset can be modeled using a normal model, write the model in this form: (mean, standard deviation)
* Remember, if no technology is available, you can use these formulas to find the mean and standard deviation:
  + mean:

|  |
| --- |
| is the mean |
| is the sum of all values |
| represents the number of values |

* + standard deviation:
    - For each value in the data set, subtract the mean from the value, then square the difference.
    - Add all the results together and divide by the number of values in the set.
    - Take the square root to find the standard deviation.
* How to verify the normal model fits the given data (when the data is provided.)
  + Create a histogram of the data.
  + If the histogram is described by these three characteristics, it is normal:

1. There is a single peak.
2. The mean is near the center of the data distribution.
3. The distribution is symmetric about the mean.

**Practice Questions and Answers**

|  |  |  |
| --- | --- | --- |
|  | Question | Answer |
| P 1 | The height of men in the United States is normally distributed with a mean of 69 inches and a standard deviation of 4 inches. Complete the input for the spreadsheet program to calculate the probability that a man in the U.S. chosen at random will be shorter than 64 inches. Then identify this probability and round the answer to the nearest tenth. | 1. 69 2. 4 3. 64 4. 10.6 |
| P 2 | A bakery owner wants to ensure they make enough cookies each day to meet the demand from customers. On average, they sell 92 cookies a day with a standard deviation of 9. The baker makes 70 cookies each day. Using a calculator or a spreadsheet program, find the probability that the baker made enough cookies (no more than 70 cookies sold that day). Round the answer to the nearest tenth of a percent.  The probability that 70 cookies were enough for that day’s sales is \_\_\_\_\_%. | 7.3 |
| P 3 | Consider a dataset with the following characteristics:  mean = 37.8  standard deviation = 2.6  In a histogram, the data is roughly symmetric. Is a normal model a good fit for the dataset? | Not enough information to determine |
| P 4 | Jayla is given the model for a set of 75 data values and wants to determine if a normal model would be a good fit for her set. What are the and values for her set? How many data points should she expect to fall within the and values?   1. The value is \_\_\_\_\_. 2. The + value is \_\_\_\_\_. 3. Jayla can expect that roughly \_\_\_\_\_ values should fall between and | 1. 343 2. 379 3. 51 |
| P 5 | What is the mean of the dataset in the normal model ?  mean = \_\_\_\_\_ | 870 |

**Quick Check Questions and Answers**

|  |  |  |
| --- | --- | --- |
|  | Question | Answer |
| Q 1 | LeBron’s average bowling score for the season is 180 with a standard deviation of 28. Use technology to determine which of the following represents the probability that LeBron records a score higher than 200. | 23.8% |
| Q 2 | Given and , find the probability that a random variable, *x*, is between 3.6 and 6.1. | 98.4% |
| Q 3 | *Use the image to answer the following question.*  A normal curve is marked and labeled to show the values of a normal distribution. Three percentages are listed above the curve. The Mean and Standard Deviations are below the horizontal axis. Within the curve, 8 regions and percentages are defined. The graph is titled The Empirical Rule.  Given a dataset of 60 values with the normal model , approximately how many values should fall between 61 and 71 for the normal model to apply? | 29 |
| Q 4 | Marcus is examining a histogram based on a dataset. He notices that the data is roughly symmetrical around a single peak at 50. The mean of the data is at 50.5 and the point is at 60.5. What is the normal model for the data distribution? | (50.5, 10) |
| Q 5 | Lydia collected samples to see how many cars passed by her house in a period of 5 minutes. This is the data that she collected:  17, 8, 17, 21, 16, 15, 13, 10, 18, 17, 23, 17, 18, 10, 18, 13, 16, 21, 17, 16.  The normal model that Lydia calculated for this model is . Then, Lydia used the Empirical Rule to check whether this data fits the Empirical Rule. She checked the number of data that were to the left of the point, which is 19.83.  What is the correct conclusion?  A normal curve is marked and labeled to show the values of a normal distribution. Three percentages are listed above the curve. The Mean and Standard Deviations are below the horizontal axis. Within the curve, 8 regions and percentages are defined. The graph is titled The Empirical Rule. | The normal model is a good fit because 85% of the data are less than the value at the point, and the model predicts 84% |