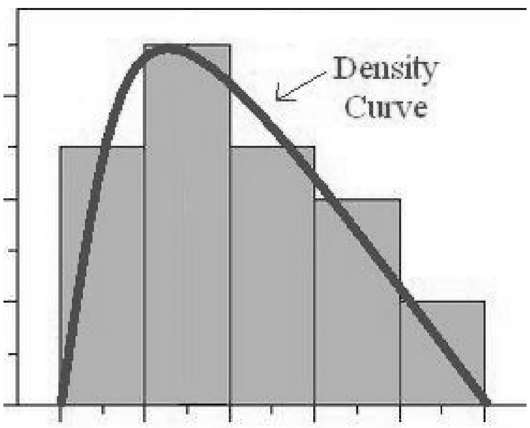
Chapter 7 - The Normal Distribution

7.1 Introduction to the Normal Curve

Learning Objectives

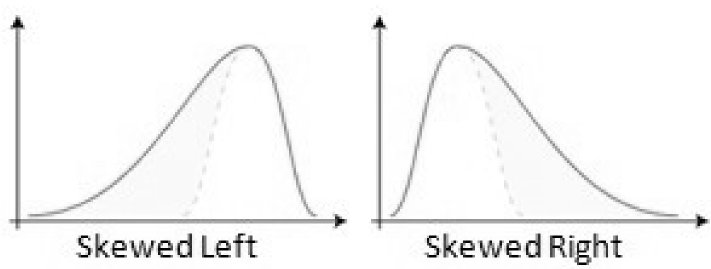
- Understand how a density curve can be used to approximate the data in a histogram
- Understand how to visually identify the mean and standard deviation of a normal distribution
- Be able to tie the concepts of percentages in the 68-95-99.7 Empirical Rule to normal distributions

https://bit.ly/probstatsUnit7
(3 sections in this unit)



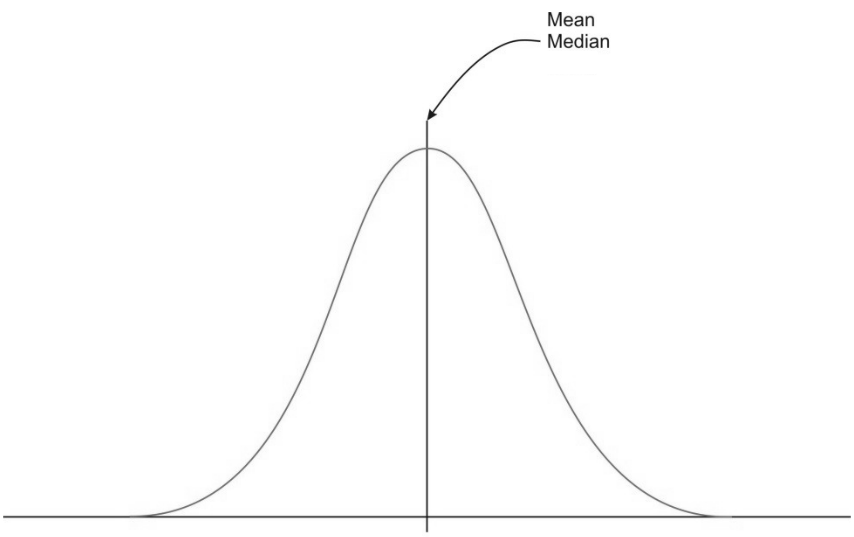
In previous chapters we have seen how data can be represented by histograms. A **density curve** is a curve that gives an approximate description of a distribution. The curve is smooth, so any small irregularities in the data are ignored. An approximate density curve for one particular histogram is shown to the left. Perhaps the most important thought to remember about a density curve is that it represents 100% of the data. In other words, the area under any density curve is equal to 1. This is important because it allows us to ask probability questions about a population. For example, we might ask how likely is it that a teenager has a shoe size of 8 or larger.

In this chapter, we will focus on a special density curve called the normal curve. Have you ever wondered if you are 'normal'? You probably are normal in most ways, but there may be some things about you that might not be considered normal by the mathematical definition. If you are on the high school baseball team, do you throw the baseball at a 'normal' speed? Is your hair a 'normal' length? Do you drive at a 'normal' speed on the freeway? Our goal this chapter is to gain an understanding of what 'normal' really is and how to properly calculate within the Normal Distribution. We have seen skewed distributions before. The density curves in the figure to the left show one density curve that is skewed left and one that is skewed right.



A **normal curve** is neither skewed left nor right and is often referred to as 'the bell curve' because of its shape. It is symmetrical. In addition, as you get closer and closer to the middle of the curve, there is a higher frequency of results. The **mean** (along with the median and mode) always lands at the center of a normal distribution. When dealing with the mean in previous chapters, we have used the symbol $\bar{x}$ because that calculation came from sample data. Normal distributions deal with an entire population instead of just a sample and we will use the symbol μ (Greek letter mu) to mark the mean of a normal distribution for an <u>entire</u> population. The mean is one of two key values needed to make a proper sketch and analysis of a normal distribution. The curve shown below represents a normal distribution and is a good representation of what a normal curve looks like.

https://bit.ly/probstatsSection7-1
(Intro to the normal curve)



Normal Distribution Center

Note that the amount of data to the right of the mean is the same as the amount of data to the left of the mean. Thinking about the definition of the median, this suggests that the mean and median are located at the same point. The other key component used to construct and analyze a normal distribution is the standard deviation. The **standard deviation** is a measure of spread and can be loosely thought of as a 'typical' distance from the mean. You may have calculated the standard deviation before for data sets either by hand or by using your calculator and looked for the s_x in the statistical calculations summary screen. The symbol s_x is used for the standard deviation whenever data is collected through the use of a sample from a population. When dealing with the normal distribution, we will use the symbol σ (Greek letter sigma) to represent the standard deviation. The σ symbol indicates that the standard deviation of the <u>entire</u> population is known. Visually, the standard deviation can be seen as the distance from the mean to an **inflection point**. An inflection point is located on a curve at the point where the curve changes from **concave up** (bent up) to **concave down** (bent down) or vice versa. On the normal curve in Figure 7.4, the mean is 23 and the standard deviation is 3.

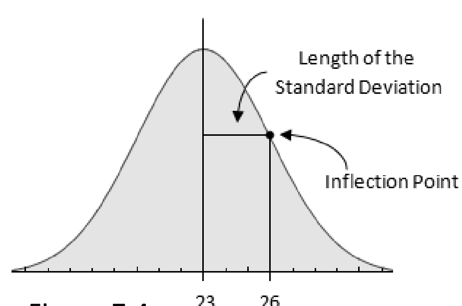
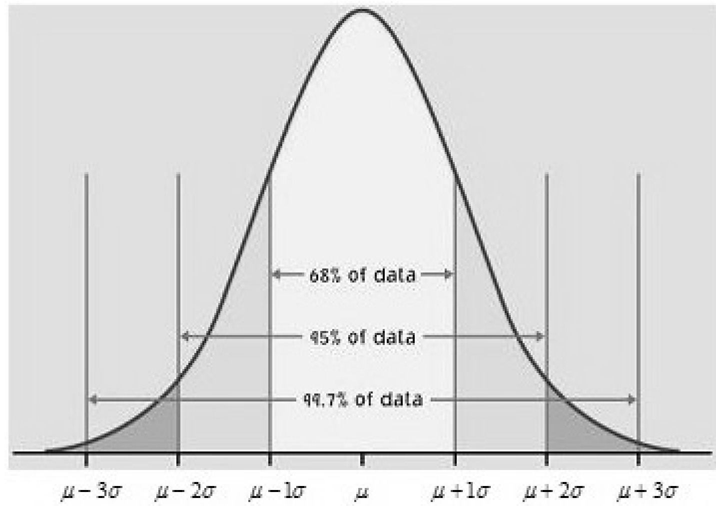


Figure 7.4

The Empirical Rule (68-95-99.7 Rule)

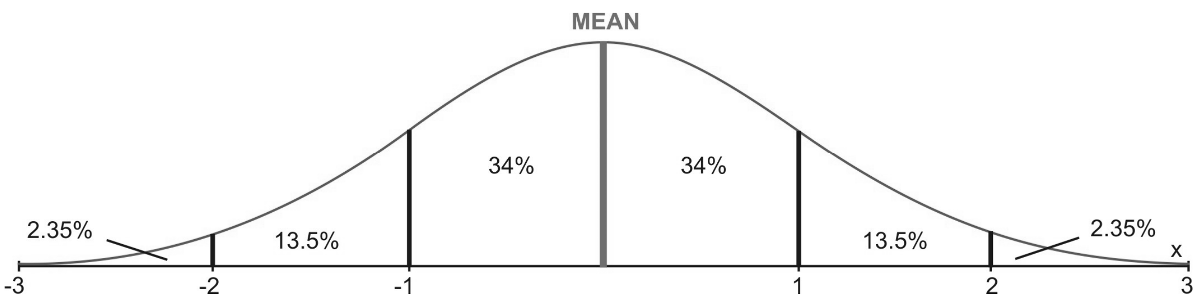
It is now time to make use of some of the special characteristics of the normal curve. By definition, 100% of all results of a normal distribution, fall somewhere under the normal curve. It turns out that approximately 68% of all results are within one standard deviation of the mean, 95% of all results are within two standard deviations of the mean, and 99.7% of all results land within three standard deviations of the mean. These percentages are illustrated in Figure 7.5 to the right.

Figure 7.5



The numbers on the bottom represent the number of standard deviations from the mean. For example, the $\mu - 1\sigma$ marks the point one standard deviation below the mean. Some simple addition and subtraction allows us to be very specific in the percentages of the data that land in the sections of the normal curve as shown below.

Because 99.7% of all results lie within three standard deviations of the mean, both the area above 3 standard deviations and below -3 standard deviations would each contain 0.15%.



Can you see the 68-95-99.7 rule here?

Figure 7.6

Example 1

Suppose the mathematics portion of the SAT exam is normally distributed with a mean of 500 points and a standard deviation of 100 points.

a) Sketch a normal curve for this situation marking the mean and the values 1, 2, and 3 standard deviations above and below the mean.

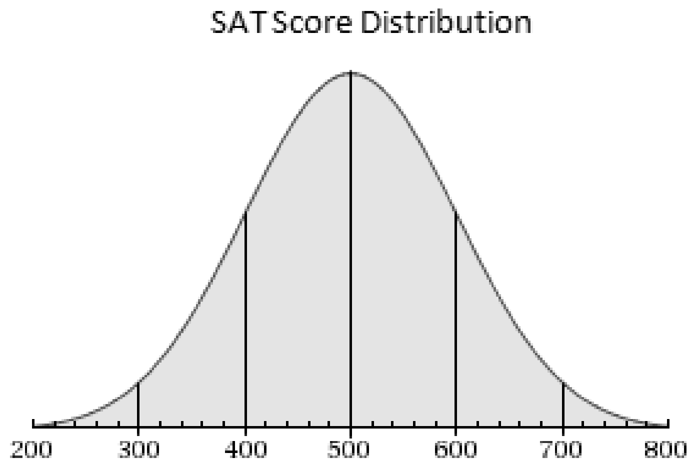
b) Using the 68-95-99.7 Rule, approximately what percent of students scored at least 600 points on this test?

c) Between approximately which two scores did the middle 95% of students score?

d) Suppose that 4600 students take the exam this month. Approximately how many of those students should we expect to obtain a score of at least 700 points?

Solution

a)



b) We know that 50% of all results are below the 500 marker and that 34% of all results land between 500 and 600. We have used up 50% + 34% = 84% of all results. This tells us that 100% - 84% = 16% of all students scored above 600 on the mathematics portion of the SAT.

c) The middle 95% of all students scored within 2 standard deviations of the mean or between 300 and 700 points.

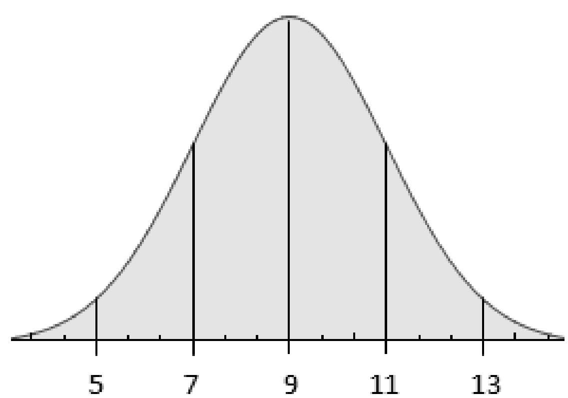
d) A score of 700 points marks the boundary two standard deviations above the mean such that only 2.5% of all test takers will score at least 700 points. Thus, 2.5% of 4600 is 115 students.

Example 2

The normal curve below represents the number of races that a typical racehorse will run in one calendar year.

a) Approximately what percent of racehorses will run between 5 and 11 races during a calendar year?

b) What are the values of the mean and standard deviation for the distribution shown?



Solution

a) Add 13.5% + 34% + 34% to get 81.5% so 81.5% of racehorses run between 5 and 11 races per year (see Figure 7.6).

b) The mean racehorse will run 9 races per year with a standard deviation of 2 races.

What is Normal?

Let's now go back and try to think about our original question "What is normal?" In mathematics, the middle 95% is often (but not always) considered our 'normal' group. For example, suppose the ACT exam is normally distributed with a mean of 18 and a standard deviation of 6. Our 'normal' group would be comprised of those students who scored anywhere within two standard deviations from the mean or from 6 to 30 on the exam. A student who scored 31 or higher on the exam would have achieved an exceptional score. We might say that this student was not normal with regards to their ACT score.

Normal distributions are not as common as you might think. What if we measured the lengths of shoes of teenagers? Many students think that this would be normal when in fact; there are a couple of contributing factors that might tip us off that the situation may not be normal. First of all, teenagers encompass a large population. Most of those who are in their upper teen years have finished growing into their adult shoe size length whereas many of the younger teens are still growing. This would tend to give us a slightly larger percentage of smaller shoe lengths than we might expect from a normal distribution. In addition, teenagers include males and females. This could lead to a situation which might be bimodal (having two modes). We might expect to see a peak at the most common male shoe lengths and at the most common female shoe lengths.

Example 3

Which situation below is most likely to produce a normal distribution?

[A] The heights of all adults.

[B] The wingspans of three year-old American eagles.

[C] The number of teeth that Americans adults have.

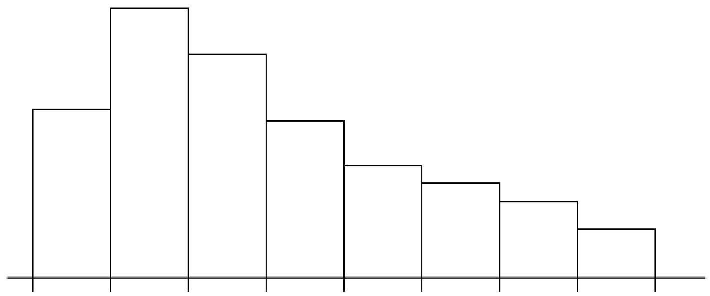
Solution

The correct answer is [B]. Three year-old American eagles have an average wingspan and we would expect that there are quite a few eagles at that wingspan or very close to it. As we move further and further up and down from that average, we would expect to see fewer and fewer eagles with those wingspans. Answer [A] could be ruled out quickly in that the heights here do not specify a particular group. For example, this data would include males and females. Answer [C] is out because the vast majority of American adults have 32 teeth. As we move away from 32, there are some people with fewer teeth due to a variety of reasons but there are virtually no people with more than 32 teeth. This distribution would be skewed left and therefore not a normal distribution.

Problem Set 7.1

Exercises

1) Consider the histogram shown to the right.

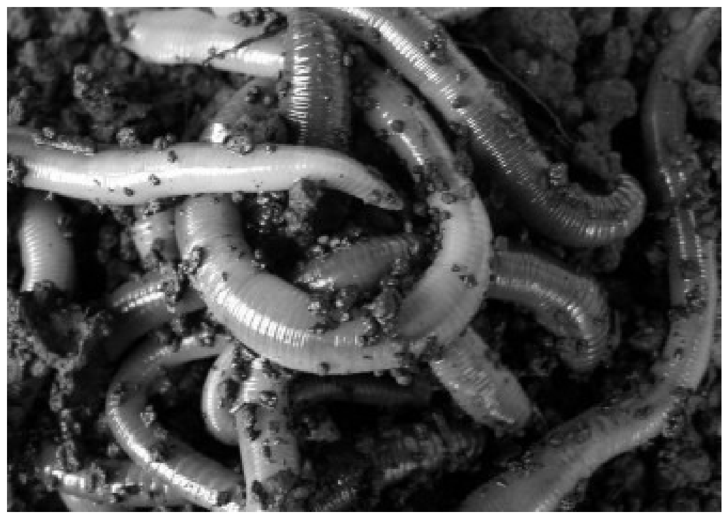


   a) Make a sketch of the histogram and overlay a sketch of a density curve for the histogram.

   b) What is the area under your density curve?

   c) What is the shape of the density curve?

2) A roadside bait salesman digs up worms to sell to fishermen. It turns out that the worms have a mean length of $\mu = 112$ mm and a standard deviation of $\sigma = 12$ mm.



   a) Draw and label a normal curve for this distribution. Include lines for the mean and for 1, 2, and 3 standard deviations above and below the mean.

   b) What percentage of the worms will have lengths longer than 112 mm?

   c) What percentage of the worms will have lengths between 100 and 124 mm long?

   d) What percentage of the worms will have lengths between 100 and 112 mm long?

   e) What percent of the worms are longer than 124 mm?

   f) What percent of the worms are shorter than 88 mm?

3) Sketch and label a normal curve which has a mean of 13 lbs. and a standard deviation of 3 lbs. Include lines for the mean and for 1, 2, and 3 standard deviations above and below the mean.

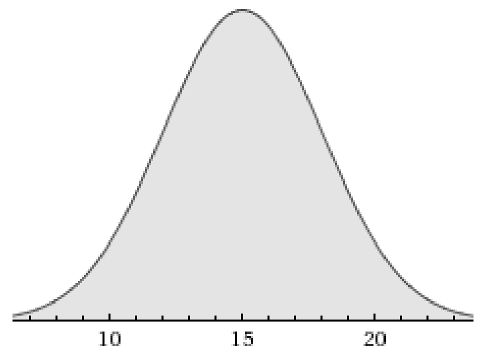
4) Not all 12-ounce cans of soda are the same. It turns out that the average 12-ounce can of soda does contain twelve ounces of soda, but the amount of soda is normally distributed with a standard deviation of 0.15 ounces. Fill in the blanks for each statement below.

   a) The middle 68% of all 12-ounce soda cans contain between _____ & _____ ounces of soda.

   b) The middle 95% of all 12-ounce soda cans contain between _____ & _____ ounces of soda.

   c) The middle 99.7% of all 12-ounce soda cans contain between _____ & _____ ounces of soda.

5) The graph to the right shows an approximate distribution of the number of fish caught by the competitors during a one hour pan-fishing contest. Give the approximate values of the mean and the standard deviation for the distribution.



Number of Fish Caught

6) Suppose the weights of adult males of a particular species of whale are distributed normally with a mean of 11,600 pounds and a standard deviation of 640 pounds.

   a) Draw and label a normal curve for this situation. Use vertical lines to mark and label the mean and 1, 2, and 3 standard deviations above and below the mean.

   b) Approximately what percent of these whales weigh less than 10,320 pounds?

   c) Between what two weights do the middle 99.7% of these whales weigh?

   d) About what percent of these whales weigh between 10,320 pounds and 12,240 pounds?

7) Which situation is most likely to be normally distributed? Explain your reasoning.

   a) The hair lengths for all the Statistics and Probability students who have Mr. Johnson as a teacher.

   b) The prices of all latest Samsung cell phones that are sold in Minnesota this week.

   c) The running times for all 4th grade males at Andover Elementary in the 50 yard dash.

8) Suppose a standard light bulb will run an average of 400 hours before burning out. Of course, some bulbs burn out sooner and some last longer. Suppose the lives of these bulbs are normally distributed with a standard deviation of 35 hours.

   a) Sketch and label a normal curve to illustrate this situation.

   b) What percent of these bulbs would we expect to burn out in 400 hours or less?

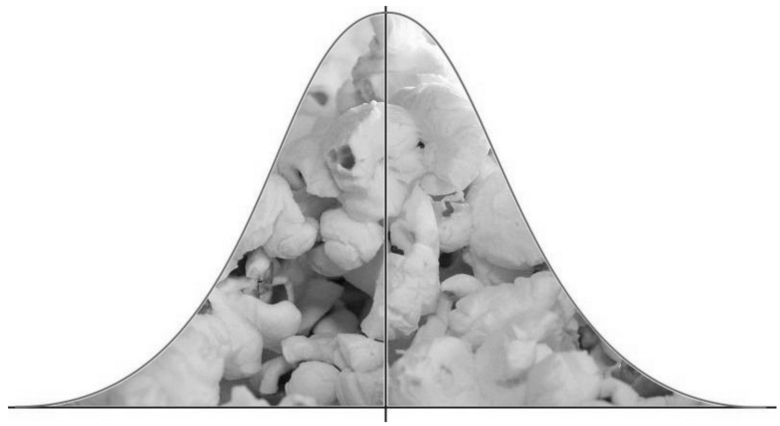
   c) Some bulbs will last longer than advertised. What percent of bulbs would we expect to last 435 hours or more? What percent of bulbs will last 470 hours or more?

   d) If there were 5000 bulbs needed for use in a large office building, how many would be expected to last at least 365 hours?

9) Suppose that the time that it takes for popcorn kernels to pop produces a normal distribution with a mean of 145 seconds and a standard deviation of 13 seconds for a standard microwave oven.

   a) It is usually not a good idea to let the microwave oven run until all the kernels are popped because some of the popcorn will start to burn. Suppose the ideal time to shut off the microwave oven is after about 97.5% of the kernels have popped. When will 97.5% of the kernels be popped?

   b) Between what two times will we see the middle 68% of kernels popped?



10) After a great deal of surveying, it is determined that the average wait times in the cafeteria line are normally distributed with a mean of 7 minutes and a standard deviation of 2 minutes. Suppose that 400 students are released to the cafeteria for 2nd lunch.

   a) Approximately how many students will have to wait more than 5 minutes for their food?

   b) Approximately how many students will have to wait more than 11 minutes for their food?

11) Sudoku is a popular logic game of number combinations. It originated in the late 1800s in the French press, Le Siècle. The mean time it takes the average 11th grader to complete a particular Sudoku puzzle was found to be 19.2 minutes, with a standard deviation of 3.1 minutes.

   a) Draw and label a normal distribution curve to represent this data.

   b) Suppose Andover High School is going to put together a Sudoku team. The coach has decided that she will only consider players who score in the fastest 2.5% of the junior class as she puts together the team. How fast must a student solve a puzzle to be in the top 2.5% of puzzle solvers?

   c) If there are 400 kids in the Andover junior class, how many of them will be able to solve the Sudoku puzzle below in 16.1 minutes or less?

12) In order to qualify for undercover detective training, a police officer must take a stress tolerance test. Scores on this test are normally distributed with a mean of 60 and a standard deviation of 10. Only the top 16% of police officers score high enough on the test to qualify for the detective training. What is the cutoff score that marks the top 16% of all scores?

Review Exercises

13) The ages of the kids at the YMCA summer day camp on Tuesday ranged from 3 to 8 years old. Use your calculator to find the mean and standard deviation for the ages in the data set below.

    3, 4, 4, 5, 5, 5, 6, 6, 7, 7, 8

14) A pet store must select 2 dogs and 2 cats for display in their front window. In how many ways can this be done if there are 16 dogs and 12 cats available to choose from?

15) Several students were asked how many missing assignments they had in their math class. The results are reported below. Find the five number summary for the number of missing assignments for these students.

    3, 5, 5, 6, 8, 9, 10, 10, 12, 13, 13, 13, 14, 15, 17, 19, 19, 20

16) A student conducts a survey in which 100 tenth-graders are asked "What is your favorite item on the lunch menu at school today?" The student decides to conduct this survey by handing each tenth-grader a survey sheet while they are eating and asking them to fill it out and turn it in to room P202 by the end of the day. Why will this survey method have a problem with bias?

7.2 Z-Scores, Percentiles, and Normalcdf

https://bit.ly/probstatsSection7-2
(4 videos in this section)

Learning Objectives

- Be able to calculate and understand z-scores
- Understand the concept of a percentile and be able to calculate it for a particular result
- Be able to calculate percentages of data above, below, or in between any specific values in a normal distribution
- Be able to use z-scores to compare results for two different but related situations
- Be able to make all of these calculations by hand and with technology

In section 7.1, we analyzed normal distributions and specific situations in which analysis was done for data which followed the 68-95-99.7 Rule exactly. The truth of the matter is that most situations require us to answer questions that do not reference exact whole numbers of standard deviations above or below the mean. What if we asked a student what their actual score would be if they were in the top 10% of ACT test takers? We need a tool to help us deal with these types of situations.

Our first tool will be the z-score formula. The **z-score** is a measure of how many standard deviations above or below the mean a particular value is. If a z-score is negative, the result is below the mean and if it is positive, the result is above the mean. For example, if the ACT mathematics exam scores are normally distributed with a mean of 18 and a standard deviation of 6, then an ACT score of 30 would be equivalent to a z-score of 2 because 30 would be 2 standard deviations above the mean.

The formula below gives a quick way to calculate z-scores. In the formula, x is the observation, μ is the mean of the distribution, and σ is the standard deviation for the distribution.

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

$$z = \frac{observation - mean}{standard\ deviation}$$

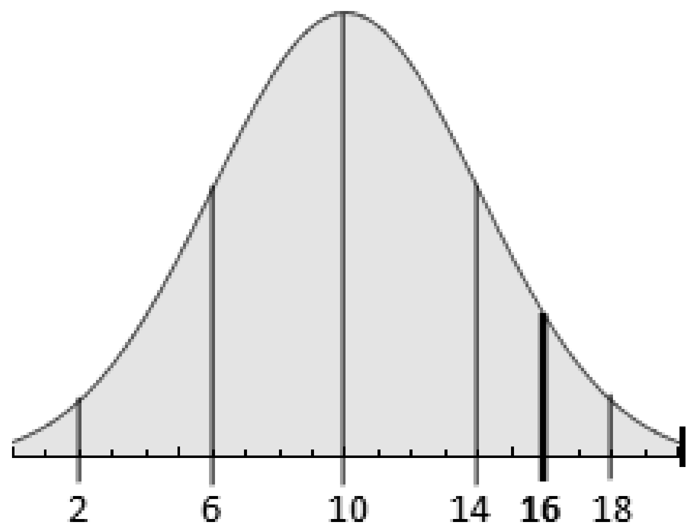
Example 1

Suppose the mean length of the hair of 10th grade girls is 10 inches with a standard deviation of 4 inches. What would be the z-score for hair length for a 10th grade girl whose hair is 16 inches long and what does it mean in terms of the normal curve?

Solution

It is often a good idea to draw a sketch for these sorts of situations so we can visualize what is happening.

Girls' Hair Lengths (inches)



Because 16 is located between 1 and 2 standard deviations above the mean, we expect a z-score between 1 and 2. Use the formula $z = \frac{x-\mu}{\sigma}$ to calculate the z-score. Our observation, x, is 16 inches while the mean is μ = 10 inches and the standard deviation is σ = 4 inches.

$z = \frac{16-10}{4}$ or z = 1.5. This tells us that a hair length of 16 inches will be 1.5 standard deviations above the mean.

Example 2

Suppose that the z-score for a particular 10th grade girl's hair length is z = -1.25. What is the length of the girl's hair?

Solution

We will use the z-score formula to find our answer again. Note that this time it is the observation that is unknown.

$$-1.25 = \frac{x-10}{4}$$

−5 = x − 10

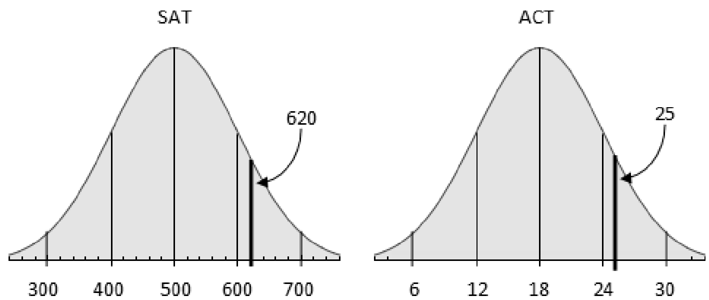
5 = x   The length of the hair for this girl would be 5 inches.

Example 3

Suppose a student can either submit only their SAT score or their ACT score to a particular college. Suppose their SAT score was 620 points and that the SAT has a mean of 500 points and a standard deviation of 100 points. Suppose also that the same student scored a 25 on their ACT exam and that the ACT exam has a mean of 18 points and a standard deviation of 6 points. Which score should the student submit?

Solution

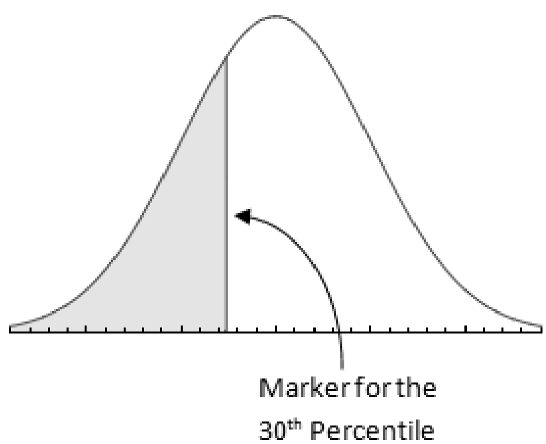
Looking at the diagram below, it is not exactly clear which score is better. They appear to be quite similar and we will need to do some calculations to make a distinction.



Calculate the z-score for each exam. For the SAT, $\mathrm{z} = \frac{620-500}{100} = 1.2$. For the ACT, $\mathrm{z} = \frac{25-18}{6} \approx 1.17$. Since the z-score is higher on the SAT, the student should submit their SAT exam score.

Percentiles

In order to understand how to apply z-scores beyond what we have already done, we must first understand percentiles. A **percentile** is a marker on a normal curve such that the marker is greater than or equal to that percentage of results. For example, suppose you are at the 30th percentile for how fast you type. This means that you can type faster than 30% of all people. The percentile can also be thought of as the percent of area to the left of its marker. The graphic below shows where the 30th percentile is located. The shaded area to the left of the marker represents 30% of the normal curve.



It is very common for colleges and universities to use percentiles for entrance criteria. For example, a rather elite university might require that you score at the 90th percentile or higher on your ACT exam to be considered for admission. Doctors often use percentiles to track the growth of babies. For example, can you picture what a baby would look like that is at the 70th percentile for weight and the 25th percentile for length?

Now we must ask what percentiles have to do with z-scores. Find the Normal Distribution Table in Appendix A, Part 2, in the back of your book. Let's examine the z-score of -1.25 from Example 2 to see how to use the table. Find the z-value of -1.2 and then go over until you are under the 0.05 column. A partial table is given in the graphic below. The value in the cell we are looking for is bold and underlined. The value of 0.1056 can be interpreted as a percentile. This means that the girl in Example 2 has hair that is longer than 10.56% of all girls. In other words, she is at about the 10th or 11th percentile for hair length for 10th grade girls.

Z	**0.09**	**0.08**	**0.07**	**0.06**	**0.05**	**0.04**	**0.03**	**0.02**	**0.01**	**0.00**
-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	0.1271	0.1271	0.1292	0.1314	0.1335	0.1357

Example 4

At what percentile for hair length is a 10th grade girl if her hair is 17 inches long? Recall that the mean is 10 inches and the standard deviation is 4 inches.

Solution

Start by determining her z-score which would be $z = \frac{17-10}{4} = \frac{7}{4} = 1.75$. We now go to the Normal Distribution Table in Appendix A, Part 2 of the book. We go across the row with z = 1.7 until we are under 0.05. This gives a value of 0.9599. This tells us the girl is at about the 96th percentile for hair length. In other words, this girl's hair is longer than 96% of all 10th grade girls.

'Between' and 'Above' Problems

While it is nice to find percentiles for certain situations, we are often asked for the percentage of results that are between two given parameters or above a given parameter. For example, we might be asked to find the percentage of all 10th grade girls that have hair lengths between 8 inches and 15 inches long. To find these types of results, we often must do multiple z-score calculations and some addition or subtraction.

Example 5

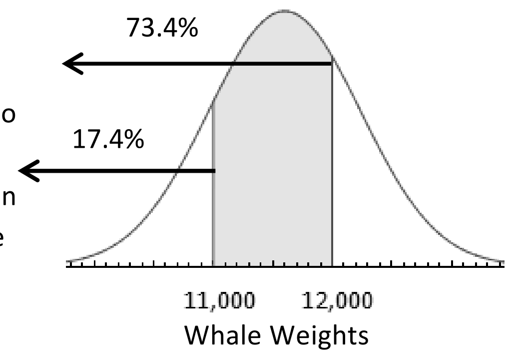
Suppose the weights of adult males of a particular species of whale are distributed normally with a mean of 11,600 pounds and a standard deviation of 640 pounds.

a) What percent of these adult male whales will weigh between 11,000 and 12,000 pounds?

b) What percent of these adult male whales will weigh more than 12,000 pounds?

Solution

a) Begin by finding the z-scores for both of the weights given and get $z = \frac{11,000-11,600}{640} = \frac{-600}{640} = -0.9375$ and $z = \frac{12,000-11,600}{640} = \frac{400}{640} = 0.625$. For z = -0.9375, our Normal Distribution Table from Appendix A, Part 2, gives us a value between 0.1736 and 0.1762. Since -0.9375 is closer to -0.94 than -0.93, we will use a value of 0.174. Likewise, we get a value between 0.7324 and 0.7357 for z = 0.625. We will split the difference on this and use 0.734. All that is left to do now is subtract 0.734 and 0.174 to get 0.56 or about 56% of all adult male whales of this species are between 11,000 and 12,000 pounds. The shaded region in Figure 7.4 represents about 56% of the normal curve.



b) Use z = 0.625 from part a) to get a value from the table of 0.734. This means that 73.4% of all whales weigh 12,000 pounds or less. Therefore, 100%-73.4%=26.6% of all whales weigh more than 12,000 pounds.

Technology

It is also important to note that graphing calculators can be used to quickly solve the types of problems discussed in this section by using the normalcdf command. Typically, this command requires that four values be entered, the lower bound, the upper bound, the mean, and the standard deviation. In Example 5, we can solve the problem in part a) simply by typing in the command string normalcdf(11000,12000,11600,640) and obtain a result of 0.5598 or 56%.

Be sure you know how to access this command if you have a graphing calculator. Appendix C on page 280 has some notes for common graphing calculators. An online calculator that is very similar to a graphing calculator and gives us the same information can be found at http://wolframalpha.com.

You might also be wondering how to solve a problem using the normalcdf command if only one parameter is given. Let's revisit Example 4 to see how this works.

Example 6

At what percentile for hair length is a 10th grade girl if her hair is 17 inches long? Recall that the mean hair length is 10 inches with a standard deviation of 4 inches.

Solution

There is only one boundary given in this problem. It is your job to come up with a second boundary. In this case, the percentile we want to calculate is found by finding the percentage of all girls whose hair is 17 inches or less. We will use a lower bound of -100 and an upper bound of 17. We use -100 simply because we are confident that we will not find any results any further left than this. Typically, choose your missing parameter as being so extreme that it will not be in the realm of possible results. Normalcdf(-100,17,10,4)=0.9599 so the length of the girl's hair is at about the 96th percentile.

Problem Set 7.2

Exercises

For problems 1) through 14) use the following information: On a particular stretch of road, the number of cars per hour produces a normal distribution with a mean of 125 cars per hour and a standard deviation of 40 cars per hour.

1) Sketch and label a normal curve for this situation. Be sure to label and mark the mean and 1, 2, and 3 standard deviations above and below the mean.
2) What is the z-score for an observation of 165 cars in one hour?
3) What is the z-score for an observation of 85 cars in one hour?
4) Calculate the z-score associated with an observation of 171 cars in one hour.
5) Suppose 135 cars are observed in one hour. At what percentile would this observation occur?
6) Suppose 70 cars are observed in one hour. At what percentile would this observation occur?

7) At what percentile would an observation of 125 cars occur?
8) What is the probability of observing at least 145 cars on the road in an hour?
9) What is the probability of observing between 100 and 150 cars on the road in an hour?
10) Determine the percentile for an observation of 140 cars on the road in one hour.
11) Determine the percentile for an observation of 65 cars on the road in one hour.
12) Determine the probability of observing between 90 and 130 cars on the road in one hour.
13) Determine the probability of observing at least 160 cars on the road in one hour.
14) Determine the probability of observing no more than 110 cars on the road in one hour.

For problems 15) through 20) use the following information: The number of ants found in a typical mature colony of leafcutter ants is normally distributed with a mean of 136 ants and a standard deviation of 14 ants.

15) One ant colony has 165 ants. At what percentile for size is this ant colony?
16) An ant colony has a z-score of $z = -1.35$ for size. Approximately how many ants would we expect to find in this colony?
17) Another ant colony has 131 ants. What is the z-score for this ant colony?
18) What is the probability of finding an ant colony with 160 ants or less?
19) What is the probability of finding an ant colony with 150 ants or more?
20) What is the probability of finding an ant colony that has between 120 and 155 ants in it?

21) Twin brothers Ricky and Robbie each took a college entrance exam. Ricky took the SAT which had a mean of 1000 with a Standard Deviation of 200 while Robbie took the ACT which had a mean of 18 with a standard deviation of 6. Which brother did better if Ricky scored a 1140 and Robbie scored a 22?

22) Suppose the average height of an adult American male is 69.5 inches with a standard deviation of 2.5 inches and the average height of an adult American female is 64.5 inches with a standard deviation of 2.3 inches. Who would be considered taller when compared to their gender, an adult American male who is 74 inches tall or an adult American female who is 68.5 inches tall? Explain your answer.

23) Professional golfer John Daly is one of the longest hitting golfers in history. Suppose his drives average 315 yards with a standard deviation of 12 yards. Will a drive of 345 yards be in his top 1% of his longest drives? Explain your answer.

Review Exercises

24) What is the area under any density curve?

25) In a standard deck of 52 cards, what is the probability of being dealt two queens if you are dealt two cards from the deck without replacement?

26) In a class competition, each grade (9-12) enters 10 students to run in a 500 meter race. Times for 9th grade males and 12th grade males are given in the table below in seconds. Build a back-to-back stem plot to compare data for the two groups of students.

9th Grade Times = 115, 118, 118, 121, 126, 127, 131, 134, 140

12th Grade Times = 106, 106, 109, 112, 114, 116, 116, 121, 122, 133

27) It turns out that countries that have higher percentages of people with computers also tend to have people who live longer. Is it logical to assume that shipping many computers to countries whose people have lower life-expectancies will help the people in those countries live longer? Answer the question including justification that references Cause and Effect, Common Response, Confounding, or Coincidence.

28) A sample survey at a local college campus asked 250 students how many textbooks they were currently carrying. The table below shows a summary of the findings. Use the table to determine the expected number of textbooks that an average college student at this campus would be carrying.

Textbooks Carried by Students

# of Books	0	1	2	3
Probability	0.21	0.37	0.32	0.1

7.3 Inverse Normal Calculations

https://bit.ly/probstatsSection7-3
(1 video in this section)

Learning Objectives

- Understand how to use the Normal Distribution Table and the z-score formula to find values for a particular normal distribution given a percentile
- Be able to use the Inverse Normal command on a graphing calculator to find values for a particular normal distribution given a percentile
- Be able to find values for a particular normal distribution given a 'middle' percentage range

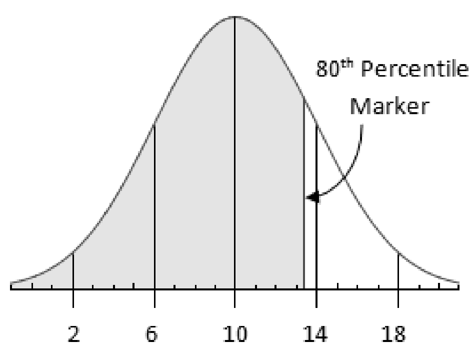
We can now comfortably calculate percentages, percentiles, and probabilities given key information about a normal distribution. It is possible to go the other direction. If you are told a certain result is at a specific percentile, you can figure out what the actual value is associated with that percentile. The process can be done using the Normal Distribution Table on page 265. Begin by identifying the percentile you are interested in and find its decimal equivalent in the table. From there, work backward to find the associated z-score. Finally, use that z-score in the z-score formula and solve it for the observation in question.

Example 1

Suppose that 10th grade girls have hair lengths that are normally distributed with a mean of 10 inches and a standard deviation of 4 inches. How long would a 10th grade girl's hair have to be in order to be at the 80th percentile for length?

Solution

The figure below shows the distribution of hair lengths and also marks where the 80th percentile is located.



Begin by finding the value closest to 0.8000 in the Normal Distribution Table. We find our closest value to be 0.7995 which corresponds to a z-score of 0.84. Put this value into the z-score formula to get

$$0.84 = \frac{x-10}{4}$$

$$0.84 = \frac{x-10}{4}$$

$$3.36 = x-10$$

$$13.36 = x$$

A 10th grade girl would have to have a hair length of about 13.4 inches to be at the 80th percentile. This looks to be right based upon comparison to the figure above.

Technology

Once again, it is important to note that technology can be used to solve these types of problems without having to reference the Normal Distribution Table. The command that is commonly used for these types of problems is the Inverse Normal command or invNorm. The Inverse Normal command requires users to enter the percentile in question, the mean, and the standard deviation. To solve the problem in Example 1, we could have typed in invNorm(0.80,10,4) and we would be given an answer of 13.366 or about 13.4 inches of hair.

Be sure you know how to access this command if you have a graphing calculator. Appendix C on page 280 has some notes for users of graphing calculators. An online calculator that can produce the same information can be found at http://wolframalpha.com.

https://bit.ly/probstatsSection7-3a
(Inverse Normal Calculation)

'Middle' and 'Top' Problems

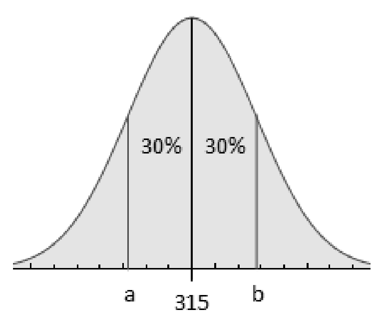
Sometimes we are in situations where we want to know what range of results are found in a middle percentage interval or what value one would have to be at in order to be in a specific top percentage. For example, a car salesman might wish to know what sales prices comprise the middle 50% of his sales to help him learn more about the type of customer that buys at his dealership. A student might wish to know what they need to score on a test in order to be in the top 10% of the class. Once again, this process can be done with either the Normal Distributions Table or by using technology.

Example 2

Professional golfer John Daly is known for his long drives off the tee. Suppose his drives have a mean distance of 315 yards with a standard deviation of 12 yards. What lengths of drives will constitute the middle 60% of all of his drives?

Solution

The sketch to the right is helpful in understanding what is happening here.



It is easy to calculate that marker line 'a' is at the 20th percentile and marker line 'b' is at the 80th percentile simply by noting their relationship to the 50th percentile marker. In addition, note that 'a' and 'b' clearly enclose the middle 60 percent of all data. From the Normal Distributions Table, we can see that the z-score associated with the 20th percentile is -0.84 and the z-score associated with the 80th percentile is 0.84. We now calculate $-0.84 = \frac{x-315}{12}$ or x = 304.9 yards. A similar calculation at the 80th percentile gives us x = 325.1 yards. We conclude that the middle 60% of John Daly's drives will travel between 304.9 yards and 325.1 yards.

We also could have used the Inverse Normal command once we knew the percentiles; invNorm(0.20,315,12) = 304.9 yards and invNorm(0.80,335,14) = 325.1 yards.

Example 3

In a weightlifting competition, the amount that the competitors can lift is normally distributed with $\mu = 196$ kg and $\sigma = 11$ kg. Only the top 20% of all competitors will be able to advance to the next phase of the competition. What amount must a competitor lift in order to move into the next phase of the competition?

Solution

The key to this problem is noticing that to be in the top 20%, a competitor would actually have to be at the 80th percentile. The z-score at the 80th percentile is z = 0.84.

$$0.84 = \frac{x-196}{11}$$

$$9.24 = x - 196$$

$$205.24 = x$$

The competitor would have to lift about 205 or 206 kg. Using a calculator, we get invNorm(0.8,196,11) = 205.26 kg.

Problem Set 7.3

Exercises

1) The Standard Normal Curve is defined as having a mean of 0 and a standard deviation of 1.
    a) What is the z-score associated with a result at the 84th percentile?
    b) What is the z-score associated with a result at the 16th percentile?
    c) Find a z-score such that only 5% of the Standard Normal Curve is to the right of that z-score.
    d) Find a z-score such that only 35% of the Standard Normal Curve is the left of that z-score.
    e) Find the two z-scores such that the middle 50% of the Standard Normal Curve is between the two z-scores.

2) Doctors often monitor their patients' blood-glucose levels. Suppose that blood-glucose levels are known to be normally distributed with, $\mu = 85$ mg/dL and $\sigma = 25$ mg/dL.
    a) Draw and label sketch of the normal distribution for this situation marking the mean and 1, 2, and 3 standard deviations above and below the mean.
    b) It turns out that doctors consider the blood-glucose level of a patient to be normal if the level is in the middle 94% of all results. What range of blood-glucose levels constitute the middle 94% of all results?
    c) Patients are considered to be at high risk for diabetes if their blood-glucose test comes back in the top 1% of all results. What blood-glucose level marks the start of the top 1% of blood-glucose levels?
    d) Doctors also show concern if there is too little blood-glucose in a patient's system. They will prescribe treatments to patients if their blood-glucose is in the lowest 2% of all patients. What is the blood-glucose level that marks this boundary?

American Diabetes Association®

3) For a given population of high school juniors and seniors, the SAT math scores are normally distributed with a mean of 500 and a standard deviation of 100. For that same population, the ACT math exam has a mean of 18 with a standard deviation of 6.
    a) One school requires that students score in the top 10% on their SAT math exam for admission. What is the minimum score that a student must achieve to be considered for this school?
    b) Another school requires that students score in the top 40% on their ACT math exam for admission. What is the minimum score that a student must achieve to be considered for this school?
    c) One particular school likes to focus on mid-level students and so they only accept students who are in the middle 50% of all ACT math test takers. Between what two scores must a student achieve in order to be considered for acceptance into this school?
    d) One student boasts that they scored at the 85th percentile on their ACT math exam. Another student brags that they scored a 620 on the SAT math exam. Who did better?

4) Many athletes train to try to be selected for the U.S. Olympic team. Suppose for the men's 100 meters, the athletes being considered for the team have a mean time of 10.06 seconds with a standard deviation of 0.07 seconds. In the final qualifying event for the team, only the top 20% of runners will be selected. What time must a runner get to be in the top 20%?

5) A high school basketball coach notices that taller players tend to have more success on his team. As a result, the coach decides that only the tallest 25% of the males in the 11th and 12th grades will be allowed to try out for the team this year. Suppose that the mean height of 11th and 12th grade males is 5 feet 9 inches (69 inches) with a standard deviation of 2.5 inches. How tall must a player be in order to be able to try out for the team?

6) A student comes home to his parents and excitedly claims that he is in the top 90% of his class. Explain why this might not be worth getting excited about.

7) At a certain fast-food restaurant, automatic soft drink filling machines have been installed. For 20-ounce cups, the machine is set to fill up the cups with 19 ounces of soda. Unfortunately, the machine is not perfectly consistent and does not always dispense 19 ounces of soda. Suppose the amount it dispenses produces a normal distribution with a mean of 19 ounces and a standard deviation of 0.6 ounces. It turns out that the 20 ounce cup will actually hold a bit more than 20 ounces. A mathematically inclined worker notices this and starts to record what happens when the machine fills the cups. It turns out that the cups overfill 2% of the time. How much soda will the 20-ounce cup actually hold?

Review Exercises

8) Adult male American bald eagles have a mean wingspan of 79 inches with a standard deviation of 3.5 inches. What percent of these eagles have wingspans longer than 7 feet?

9) Consider the data in the table below where the number of pages is the explanatory variable.

The table lists the weights of ten books and the number of pages in each one.

Number of pages	85	150	100	120	90	140	137	105	115	160
Weight (g)	165	325	200	250	180	285	250	170	230	340

a) Create a scatterplot for the data set. Label your axes.

b) Determine the correlation coefficient, r, for the scatterplot.

c) Give the least-squares linear regression equation. Be sure to define your variables.

d) Using your answer from part c), predict the weight of a book that has 130 pages.

e) Using your answer from part c), predict the number of pages for a book that weighs 295 grams.

10) Consider a standard set of 15 pool balls. Pool balls #1-8 are solid and pool balls #9-15 are striped.

a) If you randomly select one pool ball, what is the probability that it is both solid and odd?

b) If you randomly select one pool ball, what is the probability that it is either solid or odd?

c) If you randomly select two pool balls without replacement, what is the probability that they are either both solid or both striped?

7.4 Chapter 7 Review

In this chapter we have discussed what a density curve is and specifically focused on a special density curve called the normal distribution. The two critical pieces of information that are necessary for analysis of a density curve are the mean and standard deviation. The mean is the center of the distribution while the standard deviation is a measure of spread. We have focused on several key concepts including the 68-95-99.7 Rule and z-scores. We then introduced the Normal Distribution Table and the normalcdf and invNorm commands on our calculators to help us move back and forth between probabilities and percentiles and specific values in our distributions.

Review Exercises

1) Suppose a teacher gives a test in which the scores on the test are normally distributed with a mean of 10 points and a standard deviation of 2 points.
   a) Draw and label a normal curve to represent this situation. Clearly mark the mean and 1, 2, and 3 standard deviations above and below the mean.
   b) Using the 68-95-99.7 Rule, approximately what percent of students will get a score between 6 and 14?
   c) Using the 68-95-99.7 Rule, approximately what percent of students will get a score between 8 and 16?
   d) Find the percent of students that will get a score between 8 points and 13 points on this test.
   e) What percent of students will score at least 11 points on this test?
   f) What percent of students will score between 5 points and 12 points on this test?
   g) How many points would a student have to score in order to be at the 90th percentile on this test?
   h) What is the z-score associated with a test score of 13 points?
   i) How many points did a student score if their z-score was $z = -1.5$?
2) Which situation below is most likely to be normally distributed?
   a) The heights of all the trees in a forest.
   b) The distances that all the kids at Blaine High School can hit a golf ball.
   c) The number of siblings that each student at Anoka High School has.
   d) The length of time that 6th grade boys at Roosevelt Middle School can hold their breath.

3) The weights of adult male African elephants are normally distributed with a mean weight of 11,000 pounds and standard deviation of 900 pounds.

   a) Between what two weights do the middle 50% of all adult male African elephants weigh?

   b) Suppose one of these elephants weighs 13,400 pounds. At what percentile is this weight?

   c) At what weight would we find the 70th percentile of weights for these elephants?

4) Suppose that IQ test scores are normally distributed with a mean of 100 points and a standard deviation of 15 points.

   a) What z-score is associated with an IQ score of 125 points?

   b) The intelligence organization MENSA requires that members score in the top 2.5% of all IQ test takers to gain membership in the organization. What IQ score must a person score to qualify for MENSA?

   c) What percentage of IQ scores are greater than 125 points?

   d) What percentage of IQ scores are less than 70 points? Use the 68-95-99.7 Rule to approximate your answer.

   e) Who did better, a person with an IQ score of 143 points or someone who was at the 99th percentile on the IQ test? Justify your answer.

5) In a certain city, the number of pounds of newspaper recycled each month by a household produces a normal distribution with a mean of 8.5 pounds and a standard deviation of 2.7 pounds.

   a) Draw and label a sketch for this normal distribution and shade in the region that represents the households that recycle between 6 and 12 pounds of newspaper each month.

   b) What percent of households recycle between 6 and 12 pounds of newspaper each month?

   c) A local newspaper wants to do a story on newspaper recycling in the city. They decide that they would like to base their story on a typical household. After some thought, they decide that 'typical' means that they are in the middle 60% of all households in terms of newspaper recycling. Between what two weights are the 'typical' households?

6) Snowfall each winter in the Twin Cities is normally distributed with a mean of 56 inches and a standard deviation of 11 inches.

a) In what percentage of years does the Twin Cities get less than 3 feet of snow?

b) In what percentage of years does the Twin Cities get more than 6 feet of snow?

c) The winter of 2010-2011 was the fifth snowiest on record for the Twin Cities with a total snowfall of 85 inches. What percentage of years will have snowfalls of more than 85 inches?

d) A winter is considered to be dry if it in the lowest 10% of snowfall totals. What is the maximum amount of snow the Twin Cities could receive to still be called a dry winter?

7) You just got your history test back and found out you scored 37 points. The scores were normally distributed with a mean of 31 points and a standard deviation of 4 points. When you tell your parents how you did, your little brother pipes in that he got a 56 on his math test which was normally distributed with a mean of 40 points and a standard deviation of 11 points. How could you use z-scores to explain to your parents that your score was more impressive than your little brother's score?

8) In 1941, Ted Williams batted 0.406 for the baseball season. He is the last player to hit over 0.400 for an entire major league baseball season. In 2009, Joe Mauer hit 0.365 for the baseball season. In 1941, the batting averages were normally distributed with a mean of 0.260 and a standard deviation of 0.041. In 2009, the batting averages were normally distributed with a mean of 0.262 and a standard deviation of 0.035. Decide which player had a better season compared to the rest of the league during their respective year by comparing z-scores.

9) Suppose that medals will be given out to any student at Champlain Park High School that scores at least 200 points on an aptitude test. The mean score on the aptitude test is 150 points with a standard deviation of 22 points. Approximately how many medals should be ordered if there are 456 students who sign up for the test?

Image References

Density Curve www.madscientist.blogspot.com

Skewed Distributions http://en.wikipedia.org/wiki/Skewness

68-95-99.7 Normal Curve http://www.rahulgladwin.com/noteblog/biostatistics/descriptive-statistics.php

Earthworms http://www.flowers.vg

Pet Store Window http://perezhilton.com/teddyhilton/

Traffic Jam https://www.rnw.org/

Leafcutter Ant http://www.orkin.com/ants/

Pair of Queens http://www.123rf.com

American Diabetes Association https://americandiabetesassn.wordpress.com/

Track Race http://www.tierraunica.com

Eagle http://www.esa.org

Elephant http://animals.nationalgeographic.com/animals/

Blizzard http://www.csc.cs.colorado.edu

Joe Mauer http://www.mauersquickswing.com

Appendices

Appendix A – Tables

Appendix A, Part 1 - Random Digit Table

Line 101	19223	95034	05756	28713	96409	12531	42544	82853
Line 102	73676	47150	99400	01927	27754	42648	82425	36290
Line 103	45467	71709	77558	00095	32863	29485	82226	90056
Line 104	52711	38889	93074	60227	40011	85848	48767	52573
Line 105	95592	94007	69971	91481	60779	53791	17297	59335
Line 106	68417	35013	15529	72765	85089	57067	50211	47487
Line 107	82739	57890	20807	47511	81676	55300	94383	14893
Line 108	60940	72024	17868	24943	61790	90656	87964	18883
Line 109	36009	19365	15412	39638	85453	46816	83485	41979
Line 110	38448	48789	18338	24697	39364	42006	76688	08708
Line 111	81486	69487	60513	09297	00412	71238	27649	39950
Line 112	59636	88804	04634	71197	19352	73089	84898	45785
Line 113	62568	70206	40325	03699	71080	22553	11486	11776
Line 114	45149	32992	75730	66280	03819	56202	02938	70915
Line 115	61041	77684	94322	24709	73698	14526	31893	32592
Line 116	14459	26056	31424	80371	65103	62253	50490	61181
Line 117	38167	98532	62183	70632	23417	26185	41448	75532
Line 118	73190	32533	04470	29669	84407	90785	65956	86382
Line 119	95857	07118	87664	92099	58806	66979	98624	84826
Line 120	35476	55972	39421	65850	04266	35435	43742	11937

Line 121	71487	09984	29077	14863	61683	47052	62224	51025
Line 122	13873	81598	95052	90908	73592	75186	87136	95761
Line 123	54580	81507	27102	56027	55892	33063	41842	81868
Line 124	71035	09001	43367	49497	72719	96758	27611	91596
Line 125	96746	12149	37823	71868	18442	35119	62103	39244
Line 126	96927	19931	36089	74192	77567	88741	48409	41903
Line 127	43909	99477	25330	64359	40085	16925	85117	36071
Line 128	15689	14227	06565	14374	13352	49367	81982	87209
Line 129	36759	58984	68288	22913	18638	54303	00795	08727
Line 130	69051	64817	87174	09517	84534	06489	87201	97245
Line 131	05007	16632	81194	14873	04197	85576	45195	96565
Line 132	68732	55259	84292	08796	43165	93739	31685	97150
Line 133	45740	41807	65561	33302	07051	93623	18132	09547
Line 134	27816	78416	18329	21337	35213	37741	04312	68508
Line 135	66925	55658	39100	78458	11206	19876	87151	31260
Line 136	08421	44753	77377	28744	75592	08563	79140	92454
Line 137	53645	66812	61421	47836	12609	15373	98481	14592
Line 138	66831	68908	40772	21558	47781	33586	79177	06928
Line 139	55588	99404	70708	41098	43563	56934	48394	51719
Line 140	12975	13258	13048	45144	72321	81940	00360	02428
Line 141	96767	35964	23822	96012	94591	65194	50842	53372
Line 142	72829	50232	97892	63408	77919	44575	24870	04178
Line 143	88565	42628	17797	49376	61762	16953	88604	12724
Line 144	62964	88145	83083	69453	46109	59505	69680	00900
Line 145	19687	12633	57857	95806	09931	02150	43163	58636
Line 146	37609	59057	66967	83401	60705	02384	90597	93600
Line 147	54973	86278	88737	74351	47500	84552	19909	67181
Line 148	00694	05977	19664	65441	20903	62371	22725	53340
Line 149	71546	05233	53946	68743	72460	27601	45403	88692
Line 150	07511	88915	41267	16853	84569	79367	32337	03316

Appendix A, Part 2 – The Normal Distribution Table

For z-scores with z less than or equal to zero

Table 8.1

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3829
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

For z-scores with z greater than or equal to 0

Table 8.2

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	0.5987	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	0.7088	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	0.8289	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	0.8980	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	0.9686	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

Appendix A, Part 3 – A standard deck of 52 cards

Black cards		*Red cards*	
Clubs	**Spades**	**Hearts**	**Diamonds**
A ♣	A ♠	A ♥	A ♦
2 ♣	2 ♠	2 ♥	2 ♦
3 ♣	3 ♠	3 ♥	3 ♦
4 ♣	4 ♠	4 ♥	4 ♦
5 ♣	5 ♠	5 ♥	5 ♦
6 ♣	6 ♠	6 ♥	6 ♦
7 ♣	7 ♠	7 ♥	7 ♦
8 ♣	8 ♠	8 ♥	8 ♦
9 ♣	9 ♠	9 ♥	9 ♦
10 ♣	10 ♠	10 ♥	10 ♦
Jack ♣	Jack ♠	Jack♥	Jack ♦
Queen ♣	Queen ♠	Queen ♥	Queen ♦
King ♣	King ♠	King ♥	King ♦

Appendix A, Part 4 - Results for the total of two 6-sided dice

+	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Appendix B – Glossary and Index

95% confidence statement – Page 120, Section 4.4

A confidence statement is a summary statement of the findings of a study. All confidence statements have the form 'We are 95% confident that the true proportion of (parameter of interest) will be between (low value of confidence interval) and (high value of confidence interval).'

Back to Back Stem Plots – Page 184, Section 5.6

A stem plot in which two sets of numerical data share the stems in the middle, with one set having its leaves going to the right and the other set having its leaves going to the left.

Bar Graph – Page 137, Section 5.1

A graph in which each bar shows how frequently a given category occurs. The bars can go either horizontally or vertically. Bars should be of consistent width and need to be equally spaced apart. The categories may be placed in any order along the axis.

Bias - Page 95, 103, Section 4.1, 4.2

Bias occurs when a measurement repeatedly reports values that are either too high or too low.

Bin Width

See Class Size

Bivariate Data - Page 200, Section 6.1

Numerical data that measures two variables.

Blind Study - Page 126, Section 4.5

A study in which the subject does not know exactly what treatment they are getting.

Block Design - Page 128, Section 4.5

A study in which subjects are divided into distinct categories with certain characteristics (for example, males and females) before being randomly assigned treatments in an experiment.

Box Plot (Box and Whisker Plot) - Page 171, Section 5.5

A display in which a numerical data set is divided into quarters. The 'box' marks the middle 50% of the data and the 'whiskers' mark the upper 25% and lower 25% of the data.

Categorical Variable - Page 93, 136, Section 4.1, 5.1

Variables that can be put into categories, like favorite color, type of car you own, your sports jersey number, etc...

Census - Page 97, 101, Section 4.1, 4.2

A special type of study in which data is gathered from every single member of the population.

Center - Page 147, 156, Section 5.2, 5.3

Typically, it is the mean, median, or the mode of a data set. In a normal distribution curve the mean, median, and mode all mark the center. If a data set is skewed or has outliers, it is standard practice to use the median as the center.

Chance Behavior - Page 26, Section 2.1

Events whose outcomes are not predictable in the short term, but have long term predictability.

Class Size (Bin Width) - Page 164, Section 5.4

A consistent width that all bars on a histogram have. A quick estimation of a reasonable class size is to roughly divide the range by a value from about 7 to 10.

Coincidence - Page 215, Section 6.2

A relationship between two variables that simply occurs by chance.

Combination - Page 15, Section 1.4

An arrangement of a set of objects in which the order does not matter.

$$_{n}C_{r} = \frac{n!}{r!(n-r)!}$$

Common Response - Page 214, Section 6.2

A situation in which two variables have a strong correlation but are actually responding to an additional lurking variable.

Complement of an Event - Page 26, Section 2.1

The probability of an event, 'A', NOT occurring. It can be thought of the opposite of an event and can be notated as Ac or A'. P(A') = 1 – P(A)

Compound Event - Page 33, Section 2.2

An event with two or more steps such as drawing a card and then rolling a die.

Conditional Probability - Page 54, Section 2.5

The probability of a particular outcome happening assuming a certain prerequisite condition has already been met. A clue that a conditional probability is being considered is the word 'given' or the vertical bar symbol, |.

Confidence Interval - Page 119, Section 4.4

The range of answers included within the margin of error. Typically, we use a 95% confidence interval meaning it is very likely (95% chance) that the parameter lies within this range.

Confounding - Page 215, Section 6.2

Occurs when two variables are related, but it is not a clear cause/effect relationship because there may be other variables that are influencing the observed effect.

Context - Page 156, 204 Section 5.3

The specific realities of the situation we are considering. We often consider the labels and units when defining the context.

Contingency Table

See Two-Way Table

Control - Page 125, Section 4.5, 6.1

A researcher in an experiment establishes control when one of the treatment groups receives either a placebo or the currently accepted treatment.

Control Group - Page 125, Section 4.5

A group in an experiment that does not receive the actual treatment, but rather receives a placebo or a known treatment.

Convenience Sample - Page 106, Section 4.2

A biased sampling method in which data is only gathered from those individuals who are easy to access or are conveniently located.

Correlation (r) - Page 210-213, Section 6.2

A statistic that is used to measure the strength and direction of a linear correlation whose values range from -1 to 1. The sign of the correlation (+/-) matches the sign of the slope of the regression equation. A correlation value of 0 indicates no linear relationship whatsoever.

Data - Page 93, Section 4.1

A collection of facts, measurements, or observations about a set of individuals.

Density Curve - Page 236, Section 7.1

A curve that gives a rough description of a distribution. The curve is smooth and always has an area equal to 1 or 100%.

Direct Cause and Effect - Page 214, Section 6.2

A situation in which one variable causes a specific effect to occur with no lurking variables.

Direction - Page 210, Section 6.2

One of three general results reported for a linear regression. It will be reported as either being positive, negative, or 0.

Disjoint

See Mutually Exclusive Events

Dot Plot - Page 154, Section 5.3

A simple display that places a dot above each marked value on the x-axis. There is a dot for each result, so results that occur more than once will be shown by stacked dots.

Double Blind - Page 126, Section 4.5

A study in which neither the person administering the treatments nor the subject knows which treatment is being given.

Empirical Rule (68-95-99.7 Rule) - Page 238, Section 7.1

A rule stating that in a normal distribution, 68% of the data is located within one standard deviation of the mean, 95% of the data is located within two standard deviations of the mean, and 99.7% of the data is located within three standard deviations of the mean.

Event - Page 1, Section 1.1

Any action from which a result will be recorded or measured.

Expected Value - Page 67, Section 3.1

The average result over the long run for an event if repeated a large number of times.

Experiment - Page 97, 124, Section 4.1, 4.5

A study in which the researchers impose a treatment on the subjects.

Explanatory Variable - Page 125, 200, Section 4.5, 6.1

The x-axis variable. It can often be viewed as the 'cause' variable or the independent variable.

Factorial - Page 7, Section 1.2

A number followed by an exclamation point indicated repeated multiplication down to 1. For example, $4! = 4 \times 3 \times 2 \times 1$.

Fair Game - Page 76, Section 3.2

A game in which neither the player nor the house has an advantage. An average player over the long run will neither gain nor lose money. In other words, the expected value of the game is the same as the cost to play the game.

Five-Number Summary - Page 171, Section 5.5

A description of data that includes the minimum, first quartile, median, third quartile, and maximum numbers which can be used to create a box plot.

Form - Page 204, Section 6.1

A general description of the pattern in a scatterplot. Typical descriptions include linear, curved, or random (no specific form).

Frequency Table - Page 137, Section 5.1

A table that shows the number of occurrences in each category.

Fundamental Counting Principle - Page 4, Section 1.2

A rule that states that in order to find the number of outcomes for a multi-step event, simply multiply the number of possibilities from each step of the event.

Histogram - Page 164, Section 5.4

A special bar graph for a numerical data set. In a histogram, each bar has the same bin width and there is no space between consecutive bars. Each bar tracks the number or frequency of results in its given range.

Independent Events - Page 33, Section 2.2

Two events are independent if the outcome of one event does not change the probability for the outcome for the other event.

Individual - Page 93, Section 4.1

This is the person, animal, or object being studied.

Interquartile Range (IQR) - Page 174, Section 5.5

The distance between the lower and upper quartiles. $IQR = Q_3 - Q_1$

Instrument of Measurement - Page 94, Section 4.1

This is the tool used to make measurements. Some examples of instruments include rulers, scales, thermometers, or speedometers.

Intersection of Events - Page 42, Section 2.3

In a Venn Diagram, it includes the results that are members of more than one group simultaneously. We use the symbol, ∩, to indicate the intersection and think of the intersection as those parts of the diagram that include both A and B.

Law of Large Numbers - Page 26, 84, Section 2.1, 3.3

A rule that states that we will eventually get closer to the theoretical probability as we greatly increase the number of times an event is repeated.

Line Graph

See Time Plot

Lurking Variable - Page 124, 214, Section 4.5, 6.2

An additional variable that was not taken into account in a particular situation.

Margin of Error - Page 119, Section 4.4

It is the distance we move above and below the mean to help establish a 95% confidence interval in which we believe the true parameter is located. An approximation for the margin of error for a 95% confidence interval is M.O.E $= \pm \frac{1}{\sqrt{n}}$ where n represents the sample size.

Mean (Average) - Page 147, 237, Section 5.2, 7.1

The sum of all the numbers divided by the number of values in a data set. It is also located at the center of a normal distribution and is a good measure of center for symmetric data sets.

Median - Page 147, Section 5.2

The data result in the middle of a data list that has been organized from smallest to largest. If there are two middle data values, then the median is located halfway between those two values. Visually, it marks the spot where half of the area of a graph is below the median and half of the area is above the median. It is common to use the median as your measure of center for skewed data sets or data sets that contain outliers.

Mode - Page 147, Section 5.2

The result that appears most frequently in a data set. It also occurs at the highest point of a density curve.

Multistage Random Sample - Page 104, Section 4.2

A sampling technique that uses randomly selected sub-groups of a population before random selection of individuals occurs.

Mutually Exclusive Events (Disjoint) - Page 41, Section 2.3

Outcomes that cannot occur at the same time. For example, if a single card is drawn from a standard deck, the outcomes of a diamond and a black card are mutually exclusive.

Negative Linear Association - Page 205, Section 6.1

A situation such that as one numerical variable increases, another numerical variable decreases.

Non-Response - Page 108, Section 4.2

A non-sampling error in which individuals selected for a study do not participate or do not answer questions in a survey.

Normal Distribution Curve - Page 237, Section 7.1

A bell-shaped curve that describes a symmetrical data set such that the most frequent results occur near the mean and results become less frequent as you move further from the mean.

Numerical Variable - Page 93, Section 4.1

A variable that can be assigned a numerical value, such as height, distance, or temperature.

Observational Study - Page 97, 124, Section 4.1, 4.5

A study in which researchers do not impose a treatment on the individuals being studied. Data is collected by observing the individuals, surveying the individuals, or collecting data from the individuals from information that is already available. (Observe but do not disturb)

Outcome - Page 1, Section 1.1

A possible result of an event.

Outlier - Page 155, 178, 204, Section 5.3, 5.5, 6.1

A value that is unusual when compared to the rest of a data set. High outliers will be greater than $Q_3 + 1.5$ IQR. Low outliers will be below $Q_1 - 1.5$ IQR.

Parallel Box Plots - Page 183, Section 5.6

Multiple box plots graphed on the same axes to compare multiple data sets.

Parameter - Page 111, Section 4.2

A value that describes the truth about a population. The value is frequently unknown so a parameter is often given as a description of truth.

Permutation - Page 10, Section 1.3

A specific order or arrangement of a set of objects or items. In a permutation, the order in which the items are selected matters.

Pictograph - Page 141, Section 5.1

A bar graph that uses pictures instead of bars. These graphs can be misleading because pictures measure height and width, where bar graphs measure only height. To be effective, all the pictures used must be the same size.

Pie chart - Page 139, Section 5.1

A graph which shows each category as a part of the whole in a circle graph. Pie charts can be used if exactly 100% of the results from a particular situation are known.

Placebo - Page 126, Section 4.5

A fake treatment that is similar in appearance to the real treatment.

Placebo Effect - Page 126, Section 4.5

The placebo effect occurs when a subject starts to experience changes simply because they believe they are receiving a treatment.

Population - Page 101, Section 4.2

The entire group of individuals we are interested in. A population is often described using the word 'all'.

Positive Linear Association - Page 205, Section 6.1

A situation in which as one numerical variable increases, the other numerical variable also increases.

Prime Number - Page 42, Section 2.3

A number that has exactly 2 factors. Remember, 1 is not a prime number!

Probability - Page 26, Section 2.1

The likelihood of a particular outcome occurring.

Probability Model - Page 49, Section 2.4

A table that lists all the values for the outcomes of an event and their respective probabilities. The sum of all the probabilities in a probability model must equal 1.

Processing Errors - Page 109, Section 4.2

An error commonly made due to issues like poor calculations or inaccurate recording of results.

Prospective Studies - Page 124, Section 4.5

A study which follows up with study subjects in the future in an effort to see if there were any long-term effects.

Quartile 1 - Page 172, Section 5.5

The median of all the values to the left of the median. Do not include the median itself in this calculation if the median is one of the data points.

Quartile 3 - Page 172, Section 5.5

The median of all the values to the right of the median. Do not include the median itself in this calculation if the median is one of the data points.

Random Digit Table - Pages 82, 114, Section 3.3, 4.3, Appendix A

A long list of randomly chosen digits from 0 to 9, usually generated by computer software or calculators. A table of random digits can be found in Appendix A, Part 1.

Random Event - Page 26, Section 2.1

An event is random if it does not have short-term predictability but it has long-term predictability. For example, a coin flip is a random event because we do not know what will happen on the next flip, but we can be reasonably sure that about 50% of a long series of flips will land on heads.

Random Sampling Error - Page 107, Section 4.2

Even though a sample is randomly selected, it is entirely possible that a particular result within the population will be over-represented causing us to be significantly different from the parameter. Larger sample sizes reduce random sampling error. The margin of error is stated with most studies to account for random sampling error.

Range - Page 148, 174, Section 5.2, 5.5

A basic description of how spread out a data set is. It is calculated by subtracting the smallest number from the largest number in a data set.

Reliability - Page 95, Section 4.1

How consistently a particular measurement technique gives the same, or nearly the same measurement.

Response Bias - Page 109, Section 4.2

Occurs when an individual responds to a survey with an incorrect or untruthful answer. This type of bias can frequently happen when questions are potentially sensitive or embarrassing.

Response Variable - Page 125, 200, Section 4.5, 6.1

This is the y-axis variable. It can often be thought of as the 'effect' variable or dependent variable.

Retrospective Study - Page 124, Section 4.5

A study in which information about a subject's past is used in the study.

Sample - Page 102, Section 4.2

A representative subset of a population.

Sample Space - Page 1, Section 1.1

A list of all the possible outcomes that may occur.

Sample Survey - Page 97, Section 4.1

A survey that uses a subset of the population in order to try to make predictions about the entire population.

Sampling Frame - Page 103, Section 4.2

A list of all members of a population.

Scatterplot - Page 200, Section 6.1

Graphs that represent a relationship between two numerical variables where each data point is shown as a coordinate point on a scaled grid.

SCOFD - Page 203-206, Section 6.1

This is an acronym used for the description of a scatterplot and stands for Strength, Context, Outliers, Form, and Direction.

Simple Random Sample (SRS) - Page 103, Section 4.2

A sample where all possible groups of a particular size are equally possible. It can be thought of as putting names of all members of a population in a hat and randomly drawing until the desired sample size is reached.

Simulation - Page 82, Section 3.3

A model of a real situation that can be used to make predictions about what might really happen. Often, tables of random digits are used to carry out simulations.

Skewed Distribution - Page 155, 236, Section 5.3, 7.1

A distribution in which the majority of the data is concentrated on one end of the distribution. Visually, there is a 'tail' on the side with less data and this is the direction of the skew.

SOCCS - Page 154-156, Section 5.3

An acronym used to remember the key information to discuss for a distribution: Shape, Outliers, Center, Context, and Spread.

Spread - Page 156, Section 5.3

A way to measure variability of a data set. Common measures of spread are the range, standard deviation, and IQR.

Standard Deviation - Page 174, 237, Section 5.5, 7.1

A measure of spread relative to the mean of a data set. Use this measurement for any data set which is approximately normally distributed.

Statistic - Page 111, Section 4.2

A number that describes results from sample. This number is often a percentage and is used to make an approximation of the parameter.

Stem Plot - Page 157, Section 5.3

A method of organizing data that sorts the data in a visual fashion. The stem is made up of all the leading digits of a piece of data and the leaf is the final digit. No commas or decimal points should be used in a stem plot.

Stratified Random Sample - Page 104, Section 4.2

A sample in which the population is divided into distinct groups called strata before a random sample is chosen from each strata.

Strength - Page 203, 210, Section 6.1, 6.2

One of three measurements reported for a best-fit line that describes how close the data is to being perfectly linear.

Subjects - Page 125, Section 4.5

The individuals that are being studied in an experiment.

Symmetrical Distribution - Page 155, Section 5.3

A distribution in which the left side of the distribution looks like a mirror image of the right side of the distribution.

Systematic Random Sample - Page 104, Section 4.2

A sampling method in which the first selection is made randomly and then a 'system' is used to make the remaining selections. For example, randomly select one person from a list and then select every 14th person after that.

Theoretical Model - Page 26, 82 Section 2.1, 3.3

A model that gives a picture of exactly the frequencies of what should happen in a situation involving probability.

Theoretical Probability - Page 26, Section 2.1

A mathematical calculation of the likelihood that a given outcome will occur.

Time Plot (Line Graph) - Page 145, Section 5.2

A graph that shows how a numerical variable changes over time.

Tree Diagram - Page 2, 4, 48 Section 1.1, 1.2, 2.4

A visual representation of a multi-step event where each successive step branches off from the previous step.

Two-Way Table (Contingency Table) - Page 55, Section 2.5

A table which tracks two characteristics from a set of individuals. For example, we might track gender and grade of all the students in your high school.

Undercoverage - Page 107, Section 4.2

A sampling error in which an entire group or groups of subjects are left out or underrepresented in a study.

Union of Events - Page 41, Section 2.3

A union includes all results that are in either one category, another category, or both categories in a Venn diagram. We use the symbol $\cup$ and can think of a union as anything belonging to either A, B, or both A and B.

Validity - Page 95, Section 4.1

A measurement technique is valid if it is a reasonable way to collect data.

Variables - Page 93, Section 4.1

Characteristics about the individuals in a study in which researchers might have interest.

Venn Diagrams - Page 29, 42, Section 2.1, 2.3

Diagrams that represent outcomes or categories using intersecting circles.

Voluntary Response Survey - Page 105, Section 4.2

A biased sampling method in which participants get to choose whether or not to participate in the survey. The bias occurs because those who are most passionate about an issue will be more likely to respond.

Wording of a Question - Page 108, Section 4.2

The wording of a question can be used to manipulate individuals in a survey such that they are more likely to respond a certain way in the survey which causes bias.

Z-Score - Page 245, Section 7.2

A measure of the number of standard deviations a particular data point is away from the mean in a normal distribution. If a z-score is positive, the value is larger than the mean and if it is negative, it is less than the mean.

Appendix C – Calculator Help

This appendix is not meant to be a full guide for calculators common to students who take this course. Rather, it is intended to highlight some of the locations to access a variety of commands commonly used on a TI-30XS Multiview Scientific Calculator and a TI-84 Plus Graphing Calculator. One online source that can be helpful for those of you with graphing calculator issues can be found on the Prentice Hall website at http://www.prenhall.com/divisions/esm/app/calc_v2/.

Topic 1 - Combinations, Permutations, and Factorials

TI-30 XS Multiview

Access located in the **prb** menu. Enter the value for n, select nCr or nPr, and then enter the value for r.

TI-84 Plus

Access located in the **Math**, **PRB** menu. Enter the value for n, select nCr or nPr, and then enter the value for r

Topic 2 – Random Number Generators

TI-30 XS Multiview

Access located in the **prb** menu. Select rand, enter lowest value, enter highest value.

TI-84 Plus

Access located in the **Math**, **PRB** menu. Select RandInt, enter lowest value, enter highest value, enter number of random values desired.

Topic 3 – Means and Standard Deviations

TI-30 XS Multiview

Enter data into L_1 in the **data** menu. Press **2nd data (stat)** and select 1-Var Stats. Arrow down to find the mean, $\bar{x}$, and the standard deviation, s_x.

TI-84 Plus

Enter data in L1 by selecting **STAT** and **EDIT**. Press STAT and CALC and then select 1-Var Stats. Arrow down to find the mean, $\bar{x}$, and the standard deviation, Sx.

Topic 4 – Correlations, Slopes, and Y-Intercepts

TI-30 XS Multiview

Enter data into L_1 and L_2 in the **data** menu. Press **2nd data (stat)** and select 2-Var Stats for L_1 and L_2. Arrow down to find the slope (a), the y-intercept (b) and the correlation coefficient (r).

TI-84 Plus

Enter data in L1 and L2 by selecting **STAT** and **EDIT**. Press **STAT** and **CALC** and then select LinReg(ax+b). Be sure the Xlist and Ylist are L_1 and L_2. If you wish to store you equation into the Y= menu, press VARS, Y-VARS, Function, and Y_1. If the correlation (r) does not show up, go to **2nd CATALOG** and select DiagnosticOn.

Topic 5 – Normal Distributions

TI-30 XS Multiview

This calculator cannot perform normal distribution calculations.

TI-84 Plus

To find the percent of area in a normal curve, select **2nd DISTR** and select normalcdf(. Enter the lower bound, upper bound, mean, and standard deviation. To find a value from a percentile in a normal distribution, select 2nd DISTR and select invNorm(. Enter the %tile, mean, and standard deviation.

Image References

Random Digit Table http://uwsp.edu/math

Normal Distribution Table http://www.regentsprep.org

TI-30XS Multiview Calculator http://education.ti.com

TI-84 Plus Graphing Calculator http://education.ti.com

Appendix D – Selected Answers

Problem Set 7.1

1a) Sketch
1b) 1
1c) Skewed Right
2a) Sketch
2b) 50%
2c) 68%
2d) 34%
2e) 16%
2f) 2.5%
3) Sketch
4a) 11.85, 12.15
4b) 11.70, 12.30
4c) 11.55, 12.45
5) $\mu = 15, \sigma \approx 3$
6a) Sketch
6b) 2.5%
6c) 9,680 & 13,320
6d) 81.5%
7) c)
8a) Sketch
8b) 50%
8c) 16%, 2.5%
8d) 4,200
9a) 171 seconds
9b) 132 & 158 seconds
10a) 336
10b) 10
11a) Sketch
11b) 13 minutes
11c) 64
12) 70
13) $\bar{x} \approx 5.45, S_x \approx 1.51$
14) 7,920
15) 3, 8, 12.5, 15, 20
16) Voluntary Response

Problem Set 7.2

1) Sketch
2) 1
3) -1
4) 1.15
5) 60th
6) 8th
7) 50th
8) 31%
9) 47%
10) 65th
11) 7th
12) 36%
13) 19%
14) 35%
15) 98th
16) 117
17) -0.36
18) 96%
19) 16%
20) 79%
21) Ricky
22) The male
23) Yes
24) 1
25) 0.005
26) Stem Plot
27) No, Common Response
28) 1.31

Problem Set 7.3

1a) 0.99
1b) -0.99
1c) 1.65
1d) -0.39
1e) -0.67 & 0.67
2a) Sketch
2b) 38 & 132
2c) 143
2d) 34
3a) 628
3b) 20
3c) 14 & 22
3d) The SAT student
4) 10 seconds or less
5) 5ft, 10.5 in.
6) Answers will vary.
7) 20.23 oz.
8) 8%
9a) Scatterplot
9b) 0.96
9c) $\hat{y}$ = 2.34x - 41.59
9d) 263 grams
9e) 144 pages
10a) 0.27
10b) 0.8
10c) 0.47

Chapter 7 Review

1a) Sketch
1b) 95%
1c) 83.85%
1d) 77.5%
1e) 31%
1f) 83.5%
1g) 12.6
1h) 1.5
1i) 7
2) d)
3a) 10,393 & 11,607
3b) 99th
3c) 11,472
4a) 1.67
4b) 130
4c) 5%
4d) 2.5%
4e) The 143 is better.
5a) Sketch
5b) 73%
5c) 6.2 & 10.8 pounds
6a) 3.5%
6b) 7.3%
6c) 0.4%
6d) 41.9 inches
7) Answers will vary.
8) Williams did better.
9) 6