Wrapping Up and Looking Forward

Measures of variation, such as the variance, standard deviation, and range, are important descriptive statistics. They provide useful information about the spread of the scores in a distribution, while measures of skew and kurtosis (described in detail in Chapter 4) provide information about the shape of the distribution. Perhaps even more important than their function as descriptors of a single distribution of scores is their role in more advanced statistics such as those coming in later chapters (e.g., ANOVA in Chapters 9, 10, and 11). In the next chapter, we examine the properties of the normal distribution, a distribution with a specific shape and characteristics. Using some of the concepts from Chapter 3, we can see how the normal distribution can be used to make inferences about the population based on sample data.

Work Problems

Suppose that I wanted to know something about the age of customers at my new pizza palace. I take the first six people that walk through the door and ask their age in years. I get the following scores:

6 20 34 42 50 56

1. Calculate and report the range of this distribution.
2. Calculate and report the variance of this distribution.
3. Calculate and report the standard deviation of this distribution.
4. Write a sentence or two in which you explain, using your own words, what a standard deviation is (in general, not this particular standard deviation).
5. Pretend that this distribution of scores represents a population rather than a sample. Calculate and report the standard deviation and explain what it tells you.
6. Explain why the standard deviation you reported for Question 3 differs from the standard deviation you reported for Question 5.
7. Think of an example that illustrates why it is important to know about the standard deviation of a distribution.

For answers to these work problems, and for more work problems, please refer to the website that accompanies this book.

Notes

1 Although the standard deviation is technically not the "average deviation" for a distribution of scores, in practice this is a useful heuristic for gaining a rough conceptual understanding of what this statistic is. The actual formula for the average deviation would be $\Sigma(|X - \text{mean}|)/N$.

2 It is also possible to calculate the variance and standard deviation using the *raw score formula*, which does not require that you calculate the mean. The raw score formula is included in most standard statistics textbooks.

Chapter 4

The Normal Distribution

The **normal distribution** is a concept most people have some familiarity with, even if they have never heard the term. A more familiar name for the normal distribution is the **bell curve**, because a normal distribution forms the shape of a bell. The normal distribution is extremely important in statistics and has some specific characteristics that make it so useful. In this chapter, I briefly describe what a normal distribution is and why it is so important to researchers. Then I discuss some of the features of the normal distribution, and of sampling, in more depth.

Characteristics of the Normal Distribution

In Figure 4.1, I present a simple line graph that depicts a normal distribution. Recall from the discussion of graphs in Chapter 1 that this type of graph shows the frequency, i.e., number of cases, with particular scores on a single variable. So in this graph, the *Y* axis shows the frequency of the cases and the *X* axis shows the scores on the variable of interest. For example, if the variable were scores on an IQ test, the *X* axis would have the scores ranging from smallest to largest. The mean, median, and mode would be 100, and the peak of the line would show that the frequency of cases is highest at 100 (i.e., the mode). As you move away from the mode in either direction, the height of the line goes down, indicating fewer cases (i.e., lower frequencies) at those other scores.

If you take a look at the normal distribution shape presented in Figure 4.1, you may notice that the normal distribution has three fundamental characteristics. First, it is **symmetrical**, meaning that the higher half and the lower half of the distribution are mirror images of each other. Second, the mean, median, and mode are all in the same place, in the center of the distribution (i.e., the peak of the bell curve). Because of this second feature, the normal distribution is highest in the middle, so it is **unimodal**, and it curves downward toward the higher values at the right side of the distribution and toward the lower values on the left side of the distribution. Finally, the normal distribution is **asymptotic**, meaning that the upper and lower tails of the distribution never actually touch the baseline, also known as the *X* axis. This is important because it indicates that the probability of a score in a distribution occurring by chance is never zero. (I present a more detailed discussion of how the normal distribution is used to determine probabilities in Chapter 5.)

Why Is the Normal Distribution So Important?

When researchers collect data from a sample, sometimes all they want to know are the characteristics of the sample. For example, if I wanted to examine the eating habits of 100 first-year college students, I would just select 100 students, ask them what they eat, and summarize my data. These data might give me statistics such as the average number of calories consumed each day by the 100 students in my sample, the most commonly eaten foods, the variety of foods eaten, and

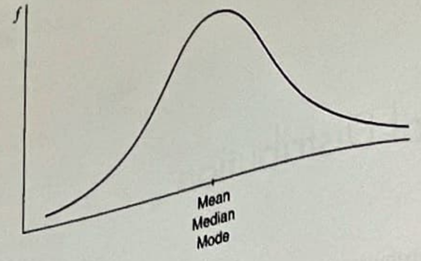

FIGURE 4.1 The normal distribution.

so on. All of these statistics simply *describe* characteristics of my sample, and are therefore called **descriptive statistics**. Descriptive statistics generally are used only to describe a specific sample. When all we care about is describing a specific sample, it does not matter whether the scores from the sample are normally distributed or not.

Many times, however, researchers want to do more than simply describe a sample. Sometimes they want to know what the exact probability is of something occurring in their sample just due to chance. For example, if the average student in my sample consumes 2,000 calories a day, what are the chances, or probability, of having a student in the sample who consumes 5,000 calories or more a day? The three characteristics of the normal distribution are each critical in statistics because they allow us to make good use of **probability** statistics.

In addition, researchers often want to be able to make inferences about the **population** based on the data they collect from their **sample**. To determine whether some phenomenon observed in a sample represents an actual phenomenon in the population from which the sample was drawn, **inferential statistics** are used. For example, suppose I begin with an assumption that in the population of men and women there is no difference in the average number of calories consumed in a day. This assumption of no difference is known as a **null hypothesis** (see Chapter 7 for more information on hypotheses). Now suppose that I select a sample of men and a sample of women, compare their average daily calorie consumption, and find that the men eat an average of 200 calories more per day than do the women. Given my null hypothesis of no difference, what is the probability of finding a difference this large between my samples *by chance?* To calculate these probabilities, I need to rely on the normal distribution, because the characteristics of the normal distribution allow statisticians to generate exact probability statistics. In the next section, I will briefly explain how this works.

The Normal Distribution in Depth

It is important to note that the normal distribution is what is known in statistics as a **theoretical distribution**. That is, one rarely, if ever, gets a distribution of scores from a sample that forms an exact, normal distribution. Rather, what you get when you collect data is a distribution of scores that may or may not approach a normal, bell-shaped curve. Because the theoretical normal distribution is what statisticians use to develop probabilities, a distribution of scores that is not normal may be at odds with these probabilities. Therefore, there are a number of statistics that begin with the assumption that scores are normally distributed. When this assumption is violated (i.e., when the scores in a distribution are not normally distributed), there can be dire consequences.

The most obvious consequence of violating the assumption of a normal distribution is that the probabilities associated with a normal distribution are not valid. For example, if you have a

normal distribution of scores on some variable (e.g., IQ test scores of adults in the United States), you can use the probabilities based on the normal distribution to determine exactly what percentage of the scores in the distribution will be 120 or higher on the IQ test (see Chapter 5 for a description of how to do this). But suppose the scores in our distribution do not form a normal distribution. Suppose, for some reason, we have an unusually large number of high scores (e.g., over 120) and an unusually small number of low scores (e.g., below 90) in our distribution. If this were the case, when we use probability estimates based on the normal distribution, we would underestimate the actual number of high scores in our distribution and overestimate the actual number of low scores in our distribution.

The Relationship Between the Sampling Method and the Normal Distribution

As I discussed in Chapter 1, researchers use a variety of different methods for selecting samples. Sometimes samples are selected so that they represent the population in specific ways, such as the percentage of men or the proportion of wealthy individuals (**representative sampling**). Other times, samples are selected randomly with the hope that any differences between the sample and the population are also random, rather than systematic (**random sampling**). Often, however, samples are selected for their convenience rather than for how they represent the larger population (**convenience sampling**). The problem of violating the assumption of normality becomes most problematic when our sample is not an adequate representation of our population.

The relationship between the normal distribution and the sampling methods is as follows. The probabilities generated from the normal distribution depend on (a) the shape of the distribution and (b) the idea that the sample is not somehow systematically different from the population. If I select a sample randomly from a population, I know that this sample may not look the same as another sample of equal size selected randomly from the same population. But any differences between my sample and other random samples of the same size selected from the same population would differ from each other randomly, not systematically. In other words, my sampling method was not **biased** such that I would continually select a sample from one end of my population (e.g., the more wealthy, the better educated, the higher achieving) if I continued using the same method for selecting my sample. Contrast this with a convenience sampling method. If I only select schools that are near my home or work, I will continually select schools with similar characteristics. For example, if I live in the "Bible Belt" (an area of the southern United States that is more religious, on average, than other parts of the country), my sample will probably be biased in that my sample will likely hold more fundamentalist religious beliefs than the larger population of schoolchildren. Now, if this characteristic is not related to the variable I am studying (e.g., achievement), then it may not matter that my sample is biased in this way. But if this bias is related to my variable of interest (e.g., "How strongly do American schoolchildren believe in God?"), then I may have a problem.

Suppose that I live and work in Cambridge, Massachusetts. Cambridge is in a section of the country with an inordinate number of highly educated people because there are a number of high-quality universities in the immediate area (Harvard, MIT, Boston College, Boston University, etc.). If I conduct a study of student achievement using a convenience sample from this area, and try to argue that my sample represents the larger population of students in the United States, probabilities that are based on the normal distribution may not apply. That is because my sample will be more likely than the national average to score at the high end of the distribution. If, based on my sample, I try to predict the average achievement level of students in the United States, or the percentage that score in the bottom quartile, or the score that marks the 75th percentile, all of these predictions will be off, because the probabilities that are generated by the normal distribution assume that the sample is not biased. If this assumption is violated, we cannot trust our results.

Skew and Kurtosis

Two characteristics used to describe a distribution of scores are **skew** and **kurtosis.** When a sample of scores is not normally distributed (i.e., not the bell shape), there are a variety of shapes it can assume. One way a distribution can deviate from the bell shape is if there is a bunching of scores at one end and a few scores pulling a tail of the distribution out toward the other end. If there are a few scores creating an elongated tail at the higher end of the distribution, it is said to be **positively skewed** (see Figure 4.2). If the tail is pulled out toward the lower end of the distribution, the shape is called **negatively skewed** (See Figure 4.3). As you can see, and as discussed in Chapter 2, the mean in a skewed distribution is pulled in the direction of the tail. Skew does not affect the median as much as it affects the mean, however. So a positively skewed distribution will have a higher mean than median, and a negatively skewed distribution will have a lower mean than median. If you recall that the mean and the median are the same in a normal distribution, you can see how the skew affects the mean relative to the median.

As you might have guessed, skewed distributions can distort the accuracy of the probabilities based on the normal distribution. For example, if most of the scores in a distribution occur at the lower end with a few scores at the higher end (positively skewed distribution), the probabilities that are based on the normal distribution will underestimate the actual number of scores at the lower end of this skewed distribution and overestimate the number of scores at the higher end of the distribution. In a negatively skewed distribution, the opposite pattern of errors in prediction will occur.

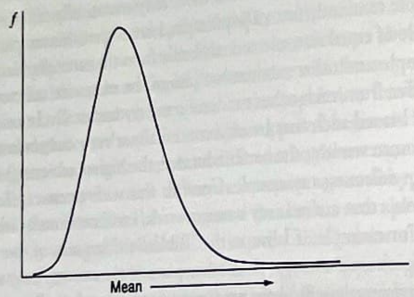

FIGURE 4.2 A positively skewed distribution.

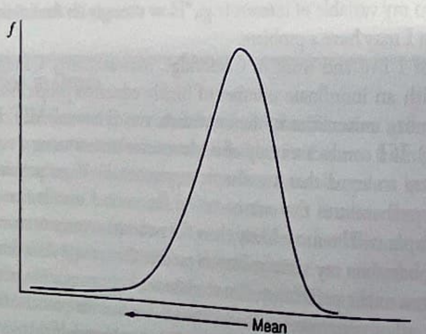

FIGURE 4.3 A negatively skewed distribution.

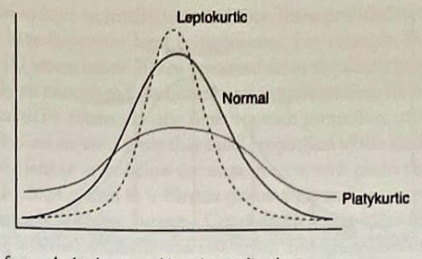

FIGURE 4.4 A comparison of normal, platykurtic, and leptokurtic distributions.

Kurtosis refers to the shape of the distribution in terms of height, or flatness. When a distribution is symmetrical but has a peak that is higher than that found in a normal distribution, it is called **leptokurtic.** When a distribution is flatter than a normal distribution, it is called **platykurtic.** Because the normal distribution contains a certain percentage of scores in the middle area (i.e., about 68 percent of the scores fall between one standard deviation above and one standard deviation below the mean), a distribution that is either platykurtic or leptokurtic will likely have a different percentage of scores near the mean than will a normal distribution. Specifically, a leptokurtic distribution will probably have a greater percentage of scores closer to the mean and fewer in the upper and lower tails of the distribution, whereas a platykurtic distribution will have more scores at the ends and fewer in the middle than will a normal distribution. These different shapes of distributions are presented in Figure 4.4.

Example 1: Applying Normal Distribution Probabilities to a Normal Distribution

In Chapter 5, I go into much more detail explaining how the normal distribution is used to calculate probabilities and percentile scores. In this chapter, I offer a more general example to provide you with an idea of the utility of the normal distribution.

In Figure 4.5, I present a graphic of the normal distribution divided into standard deviation intervals. In a normal distribution, we know that a certain proportion of the population is contained between the mean and one standard deviation above the mean. Because the normal distribution is symmetrical, we also know that the same proportion of the population is contained between the mean and one standard deviation *below* the mean. Because the normal distribution has the mode in the middle, where the mean of the population is, we also know that the further we get from the mean, the smaller the proportion of the population we will find. Therefore, there is a smaller proportion of the population contained between one and two standard deviations above the mean than there is between the mean and one standard

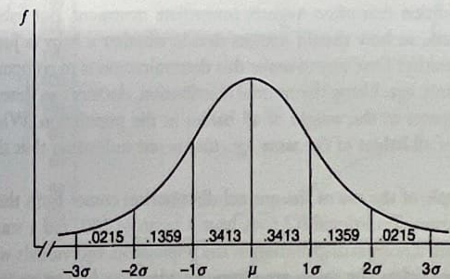

FIGURE 4.5 The normal distribution divided into standard deviation intervals.

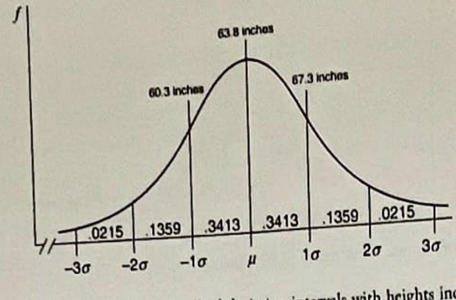

FIGURE 4.6 The normal distribution divided into standard deviation intervals with heights indicated for the mean and one standard deviation above and below the mean.

deviation above the mean. For example, .3413 (or 34.13 percent) of the population will have scores between the mean and one standard deviation above the mean in a normal distribution, but only .1359 (or 13.59 percent) will have scores between one standard deviation and two standard deviations above the mean.

We can use this information to understand where an individual score in a distribution falls in relation to the rest of the scores in the population. For example, if I know that the average height for an adult woman is 63.8 in., with a standard deviation of 3.5 in., I can use the normal distribution to determine some height characteristics of the population. For example, as seen in Figure 4.6, I know that 34.13 percent of women will be between 63.8 and 67.3 in. tall, because 63.8 in. is the mean and 67.3 in. is the mean plus one standard deviation (i.e., 63.8 + 3.5 = 67.3). Similarly, I know that 68.26 percent of women will be between and 60.3 and 67.3 in. tall, because that is the proportion of the population that is contained between one standard deviation below and one standard deviation above the mean. As we will discuss in Chapter 5, the normal distribution allows us to determine, very specifically, the proportion of scores in a population (or normally distributed sample) that falls below or above an individual score, and the probability of randomly selecting an individual with a score that is above or below that score. This is the basis of many statistics that you may have encountered (e.g., what your percentile score was on a standardized test, or what percentage of the population you are taller than), as well as the inferential statistics that we will discuss in the latter chapters of this book.

The normal distribution, and the ability to use the normal distribution to calculate probabilities, is used in the real world all the time. For example, the weight of infants and toddlers at various ages has a normal distribution. Pediatricians use these normal distributions to determine whether young children are growing normally. When toddlers or infants are severely underweight for their age, they may be diagnosed with a condition called failure to thrive. This is a serious medical condition that often requires immediate treatment. But babies and toddlers come in a variety of sizes, so how should doctors decide whether a baby is just small as opposed to dangerously unhealthy? One way to make this determination is to compare the child with other children of the same age. Using the normal distribution, doctors can determine how the weight of one baby compares to the weight of all babies in the population. When a baby weighs less than 95 percent of all babies of the same age, that is one indication that the baby may be failing to thrive.

Another example of the use of the normal distribution comes from the world of intelligence testing and IQ scores. Traditional IQ tests have a mean of 100, and a standard deviation of 15, and the scores form a normal distribution in the population. Individuals who score two standard deviations or more below the mean are often considered as having an intellectual disability. If you look at Figure 4.5, you will notice that only 2.15 percent of the normal distribution scores at least two standard deviations below the mean, so this is the proportion of the population that

would be expected to have an intellectual disability. These probabilities associated with the normal distribution have important legal consequences. For example, in some states in the U.S., individuals with IQ scores below 70 are protected from the death penalty because they are not considered intelligent enough to fully understand the crimes they have committed or the nature of the death penalty. To determine who deserves such protection, probabilities associated with the normal distribution are used. Only that small proportion of the distribution with IQ scores of two standard deviations or more below the mean receive such protection from the death penalty in many states. (In 2014, a man in a Florida prison with an IQ score of around 70 was spared from the death penalty when the Supreme Court decided that IQ scores alone were not a sound basis for deciding whether someone was intelligent enough to understand the death penalty.) These are just two of many examples of the use of the normal distribution, and the probabilities that are associated with it, in real life.

Example 2: Applying Normal Distribution Probabilities to a Nonnormal Distribution

To illustrate some of the difficulties that can arise when we try to apply the probabilities that are generated from using the normal distribution to a distribution of scores that is skewed, I present a distribution of sixth-grade students' scores on a measure of self-esteem. In these data, 677 students completed a questionnaire that included four items designed to measure students' overall sense of self-esteem. Examples of these items included "On the whole, I am satisfied with myself" and "I feel I have a number of good qualities." Students responded to each item of the questionnaire using a 5-point rating scale with 1 = *Not at all true* and 5 = *Very true*. Students' responses on these four items were then averaged, creating a single self-esteem score that ranged from a possible low of 1 to a possible high of 5. The frequency distribution for this self-esteem variable is presented in Figure 4.7.

As you can see, the distribution of scores presented in Figure 4.7 does not form a nice, normal, bell-shaped distribution. Rather, most of the students in this sample scored at the high end of the distribution, and a long tail extends out toward the lower end of the scale. This is a negatively skewed distribution of scores. The happy part of this story is that most of the students in this sample appear to feel quite good about themselves. The sad part of the story is that some of the assumptions of the normal distribution are violated by this skewed distribution. Let's take a look at some specifics.

One of the qualities of a normal distribution is that it is symmetrical, with an equal percentage of scores between the mean and one standard deviation below the mean as there are between the mean and one standard deviation above the mean. In other words, in a normal distribution,

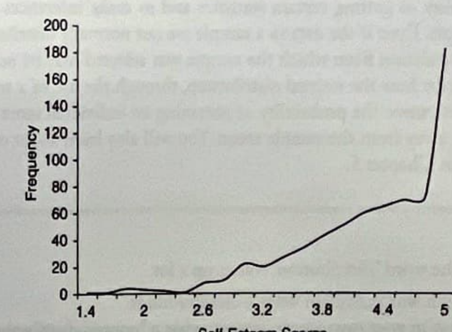

FIGURE 4.7 Frequency distribution for self-esteem scores.

there should be about 34 percent of the scores within one standard deviation above the mean and another 34 percent within one standard deviation below the mean. In our distribution of self-esteem scores presented earlier, the mean is 4.28 and the standard deviation is .72. A full 50 percent of the distribution falls between the mean and one standard deviation above the mean in this group of scores (see Figure 4.7). So, although I might predict that about 16 percent of my distribution will have scores more than one standard deviation above the mean in a normal distribution, in my skewed distribution of self-esteem scores, I can see that there are no students with scores more than one standard deviation above the mean. In Chapter 5, I present a more thorough discussion of how to use the normal distribution to calculate standard deviation units, probabilities, and percentile scores in a normal distribution.

As this example demonstrates, the probabilities that statisticians have generated using the normal distribution may not apply well to skewed or otherwise nonnormal distributions of data. This should not lead you to believe, however, that nonnormal distributions of scores are worthless. In fact, even if you have a nonnormal distribution of scores in your sample, these scores can create normal sampling distributions for use in inferential statistics (a sampling distribution is a theoretical distribution of sample statistics drawn from many samples taken from a single population – see Chapter 6). What is perhaps most important to keep in mind is that a nonnormal distribution of scores may be an indication that your sample differs in important and systematic ways from the population that it is supposed to represent. When making inferences about a population based on a sample, be very careful to define the population precisely and to be aware of any biases you may have introduced by your method of selecting your sample. It is also important to note, however, that not all variables are normally distributed in the population. Therefore, nonnormal sample data may be an accurate representation of nonnormal population data, as well as an indication that the sample does not accurately represent the population. The normal distribution can be used to generate probabilities about the likelihood of selecting an individual or another sample with certain characteristics (e.g., distance from the mean) from a population. If your sample is not normal and your method of selecting the sample may be systematically biased to include those with certain characteristics (e.g., higher than average achievers, lower than average income, etc.), then the probabilities of the normal distribution may not apply well to your sample.

Wrapping Up and Looking Forward

The theoretical normal distribution is a critical element of statistics, primarily because many of the probabilities that are used in inferential statistics are based on the assumption of normal distributions. As you will see in the coming chapters, statisticians use these probabilities to determine the probability of getting certain statistics and to make inferences about the population based on the sample. Even if the data in a sample are not normally distributed, it is possible that the data in the population from which the sample was selected may be normally distributed. In Chapter 5, I describe how the normal distribution, through the use of *z* scores and standardization, is used to determine the probability of obtaining an individual score from a sample that is a certain distance away from the sample mean. You will also learn about other fun statistics like percentile scores in Chapter 5.

Work Problems

1. In statistics, the word "distribution" comes up a lot.
   a. In your own words, explain what a distribution is.
   b. Then, again in your own words, explain what a "normal distribution" is.
2. What does "asymptotic" mean and why is it important?

3. In statistics, we use the normal distribution a lot.
   a. What is so great about it? In other words, what does it do for us?
   b. Describe something that the normal distribution lets us, as statisticians, do that we cannot do without it.
4. Many believe that most human traits form a normal distribution. Height, weight, intelligence, musical ability, friendliness, attractiveness, etc. are all examples of things that might form a normal distribution.
   a. First, explain whether you agree with this assumption, and why or why not.
   b. Second, think of an example of a trait that does NOT form a normal distribution in the population.
5. If you know that in the population of adults, the average number of hours slept per night is 7, with a standard deviation of 2, what proportion of the population would you expect to sleep between 7 and 9 hours per night?

For answers to these work problems, and for additional work problems, please refer to the website that accompanies this book.

CHAPTER 5

Standardization and z Scores

If you know the mean and standard deviation of a distribution of scores, you have enough information to develop a picture of the distribution. Sometimes researchers are interested in describing individual scores within a distribution. Using the mean and the standard deviation, researchers are able to generate a **standard score**, also called a **z score**, to help them understand where an individual score falls in relation to other scores in the distribution. Through a process of **standardization**, researchers are also better able to compare individual scores in the distributions of two separate variables. Standardization is simply a process of converting each score in a distribution to a z score. A z score is a number that indicates how far above or below the mean a given score in the distribution is in standard deviation units. So standardization is simply the process of converting individual **raw scores** in the distribution into standard deviation units.

Suppose that you are a college student taking final exams. In your biology class, you take your final exam and get a score of 65 out of a possible 100. In your statistics final, you get a score of 42 out of 200. On which exam did you get a "better" score? The answer to this question may be more complicated than it appears. First, we must determine what we mean by "better." If better means the percentage of correct answers on the exam, clearly you did better on the biology exam. But if your statistics exam was much more difficult than your biology exam, is it fair to judge your performance solely on the basis of the percentage of correct responses? A fairer alternative may be to see how well you did compared to other students in your classes. To make such a comparison, we need to know the mean and standard deviation of each distribution. With these statistics, we can generate z scores (a.k.a. standard scores).

Suppose the mean on the biology exam was 60 with a standard deviation of 10. That means you scored 5 points above the mean, which is half of a standard deviation above the mean (higher than the average for the class). Suppose further that the average on the statistics test was 37 with a standard deviation of 5. Again, you scored 5 points above the mean, but this represents a full standard deviation over the average. Using these statistics, on which test would you say you performed better? To fully understand the answer to this question, let's examine standardization and z scores in more depth.

Standardization and z Scores in Depth

As you can see in the previous example, it is often difficult to compare two scores for two variables when the variables are measured using different scales. The biology test in the example was measured on a scale from 1 to 100, whereas the statistics exam used a scale from 1 to 200. When variables have such different scales of measurement, it is almost meaningless to compare the raw scores (i.e., 65 and 42 on these exams). Instead, we need some way to put these two exams on the same scale, or to *standardize* them. One of the most common methods of standardization used in statistics is to convert raw scores into standard deviation units, or z scores. The formula for doing this is very simple and is presented in Table 5.1.

TABLE 5.1 Formula for calculating a z score

$$z = \frac{\text{raw score} - \text{mean}}{\text{standard deviation}}$$

or

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

or

$$z = \frac{X - \bar{X}}{s}$$

where X is the raw score,
μ is the population mean,
σ is the population standard deviation,
$\bar{X}$ is the sample mean,
s is the sample standard deviation.

As you can see from the formulas in Table 5.1, to standardize a score (i.e., to create a *z* score), you simply subtract the mean from an individual raw score and divide this by the standard deviation. So if the raw score is above the mean, the *z* score will be positive, whereas a raw score below the mean will produce a negative *z* score. When an entire distribution of scores is standardized, the average (i.e., mean) *z* score for the standardized distribution will always be zero, and the standard deviation of this distribution will always be 1.0.

Notice that there are actually two formulas presented in Table 5.1. The first one includes the population mean (μ, ***mu***) and the population standard deviation (σ, ***sigma***). This is the formula that you use when you are working with a distribution of scores from a population, and you know the population mean and standard deviation. When you are working with a distribution of scores from a sample, the correct formula to use for calculating standard scores (i.e., *z* scores) is the bottom formula in Table 5.1. This uses the sample mean ($\bar{X}$, ***x-bar***) and the sample standard deviation (s).

Interpreting z Scores

z scores tell researchers instantly how large or small an individual score is relative to other scores in the distribution. For example, if I know that one of my students got a *z* score of −1.5 on an exam, I would know that student scored 1.5 standard deviations below the mean on that exam. If another student had a *z* score of .29, I would know the student scored .29 standard deviation units above the mean on the exam.

Let's pause here and think for a moment about what *z* scores do *not* tell us. If I told you that I had a *z* score of 1.0 on my last spelling test, what would you think of my performance? What you would know for sure is that (a) I did better than the average person taking the test, (b) my score was one standard deviation above the mean, and (c) if the scores in the distribution were normally distributed (Chapter 4), my score was better than about 84 percent of the scores in the distribution. But what you would *not* know would be (a) how many words I spelled correctly, (b) if I am a good speller, (c) how difficult the test was, (d) if the other people that took the test are good spellers, (e) how many other people took the test, and so on. As you can see, a *z* score alone does not provide as much information as we might want. To further demonstrate this point, suppose that after I told you I had a *z* score of 1.0 on the spelling test, I went on to tell you that the average score on the test was 12 out of 50 and that everyone else who took the test was seven years old. Not very impressive in that context, is it?

Now, with the appropriate cautions in mind, let's consider a couple more uses of *z* scores and standardization. One of the handiest features of *z* scores is that, when used with a normally

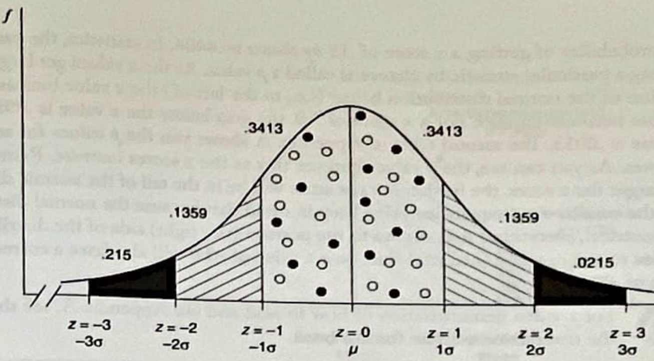

FIGURE 5.1 The standard normal distribution. This is the normal distribution divided into standard deviation units, and it shows the proportion of the normal distribution that falls between each *z* score.

distributed set of scores, they can be used to determine **percentile scores**. Percentile scores are scores that indicate the percentage of the distribution that falls below a given score. For example, the score that marks the 75th percentile in a distribution is the score at which 75 percent of the distribution falls below, and 25 percent falls above. If you have a normal distribution of scores, you can use *z* scores to discover which score marks the 90th percentile of a distribution (i.e., that raw score at which 90 percent of the distribution scored below and 10 percent scored above). This is because statisticians have demonstrated that in a normal distribution, a precise percentage of scores will fall between the mean and one standard deviation above the mean. Because normal distributions are perfectly symmetrical, we know that the exact same percentage of scores that falls between the mean and one standard deviation *above* the mean will also fall between the mean and one standard deviation *below* the mean. In fact, statisticians have determined the precise percentage of scores that will fall between the mean and any *z* score (i.e., number of standard deviation units above or below the mean). A table of these values is provided in Appendix A. When you also consider that in a normal distribution the mean always marks the exact center of the distribution, you know that the mean is the spot in the distribution in which 50 percent of the cases fall below and 50 percent fall above. With this in mind, it is easy to find the score in a distribution that marks the 90th percentile, or any percentile, for that matter. In Figure 5.1, we can see the percentages of scores in a normal distribution that fall between different *z* score values. This figure contains the *standard normal distribution.*

TIME OUT FOR TECHNICALITY: INTERPRETING APPENDIX A

Using the values in Appendix A is simple once you get the hang of it. The left column shows the *z* score value to the nearest tenth. If you need to get more precise than that, you can use the values in the top row. For example, if you have a *z* score of .15, then you find the intersection of the .1 row with the .05 column to create your *z* value of .15. If you go to that intersection, you will see that you get a value of .5596. This number indicates the proportion of the normal distribution that falls *below* this *z* value. So using Appendix A, we can conclude that .5596, or 55.96 percent, of the distribution has a *z* score of .15 *or less.* To find the proportion of the normal distribution that would be above a *z* score of .15, you simply subtract .5596 from the total of 1.0: 1.0 − .5596 = .4404. This value tells us that the probability of getting a z score of .15 *by chance* is .4404. In statistics, the probability of getting a particular statistic by chance is called a *p* value. As the z values get larger, the proportion of the normal distribution below (i.e., to the left of) the z value increases, and the *p* value becomes smaller. For a z score of 3.0, the area below the z value is .9987, and the *p* value is .0013. The second table in Appendix A shows you the *p* values for several large z scores. As you can see, the *p* values become tiny as the z scores increase. Remember that the larger the z score, the further out the score will be in the tail of the normal distribution, and the smaller the frequencies. Also, keep in mind that because the normal distribution is symmetrical, everything that applies to the positive (i.e., right) side of the distribution also applies to the negative (i.e., left) side. So a z value of −3.0 will also have a corresponding *p* value of .0013.

For a video demonstration of how to read and use Appendix A, see the video on the companion website for this book.

Let us consider an example. The Scholastic Aptitude Test (SAT) is an exam many high school students in the United States take when they are applying to college. One portion of the SAT covers mathematics and has a range of 200–800 points. Suppose I know that the average SAT math score for males is 517, with a standard deviation of 100, which forms a normal distribution. In this distribution, I already know that the score that marks the 50th percentile is 517. Suppose I want to know the score that marks the 90th percentile. To find this number, I have to follow a series of simple steps.

Step 1: Using the z score table in Appendix A, find the z score that marks the 90th percentile. To do this, we need to remember that the 90th percentile is the score at which 90 percent of the distribution falls below and 10 percent of the distribution falls above. In Appendix A, the values represent the proportion of the normal distribution that falls to the left of (i.e., *below*) the given z values. So in this example we are looking for the value in Appendix A that is closest to .90, because that is the value that marks the 90th percentile. In Appendix A, the closest value we can find to .90 is .8997, and that corresponds to a z value of 1.28 (i.e., where the $z = 1.2$ row intersects the .08 column). So $z = 1.28$ in this example.

Step 2: Convert this z score back into the original unit of measurement. Remember that the SAT math test is measured on a scale from 200 to 800. We now know that the mean for males who took the test is 517, and that the 90th percentile score of this distribution is 1.28 standard deviations above the mean (because $z = 1.28$). So what is the actual SAT math score that marks the 90th percentile? To answer this, we have to convert our z score from standard deviation units into raw score units and add them to the mean. The formula for doing this is

$$X = \mu + (z)(\sigma)$$

In this equation, X is the raw score we are trying to discover, μ is the average score in the distribution, z is the z score we found, and σ is the standard deviation for the distribution. Plugging our numbers into the formula, we find that

$$X = 517 + (1.28)(100)$$
$$X = 517 + 128$$
$$X = 645$$

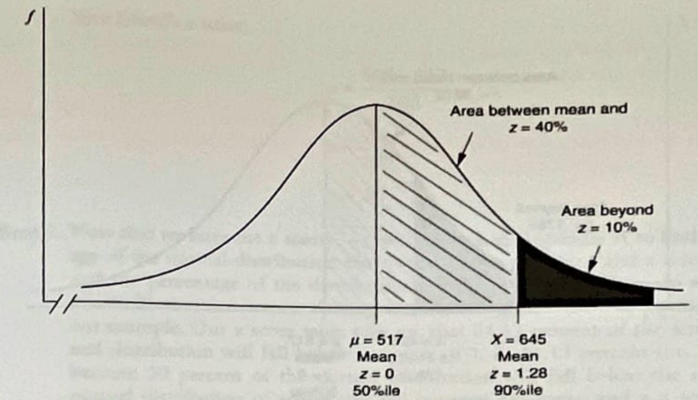

FIGURE 5.2 The score that marks the 90th percentile of this distribution.

Step 3: Now we can wrap words around our result and answer our original question. When doing this, it is often helpful to use the original question when stating our finding, as follows.

Question: *What is the score that marks the 90th percentile of the distribution of male students' SAT math scores?*

Answer: *The score of 645 marks the 90th percentile of the distribution of male students' SAT math scores.* This z score, the percentile score, and the corresponding raw score are depicted in Figure 5.2.

Just as we can use z scores to find the raw score that marks a certain percentile in a distribution, we can also use z scores to help us convert a known raw score into a percentile score. For example, if I know that a student in my distribution has a score of 425 on the SAT math test, I might want to know the percentage of the distribution that scored above and below 425. This is the type of conversion that has happened when students' standardized test scores are published in the local newspaper using percentiles under headlines such as "California Students Score in 45th Percentile on National Test!" Similarly, when a proud parent exclaims "My Johnny is in the top 10 percent in height for his age group!" a conversion from a raw score to a percentile score has taken place, with the help of a z score. Here's how it's done:

Step 1: We must begin by converting the raw score into a z score. In our example, the raw score is 425 ($X = 425$). To convert this into a z score, we simply recall our mean ($\mu = 517$) and our standard deviation ($\sigma = 100$) and then plug these numbers into the z score formula:

$$z = \frac{425 - 517}{100}$$
$$z = \frac{-92}{100}$$
$$z = -.92$$

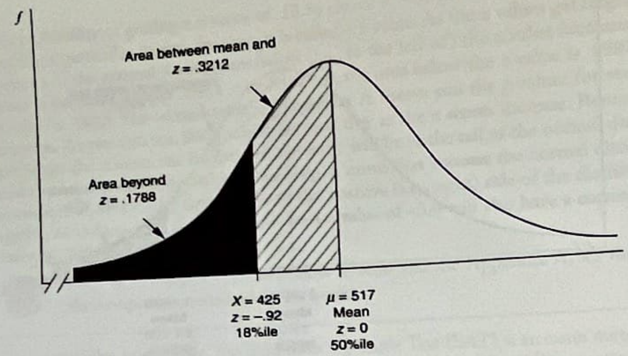

FIGURE 5.3 The percentage of the distribution scoring above and below 425.

Step 2: Now that we have a z score, we need to look in Appendix A to find the percentage of the normal distribution that falls below a z score of −.92. Notice that we are dealing with a negative z score in our example. Most z score tables only report positive z scores, but because normal distributions are symmetrical, the percentage of the distribution that falls below a negative z score is the same as the percentage that falls above the same positive z score value. Similarly, the percentage that falls above a positive z score is identical to the percentage that falls below a negative z score. My z score table in Appendix A tells me that 82.12 percent of the normal distribution will fall below a z value of .92 and 17.88 percent will fall above it (because 1 − .8212 = .1788, or 17.88 percent). The fact that the normal distribution is symmetrical tells us that 17.88 percent of the normal distribution will fall *below* a z score of −.92. So 17.88 percent of the population would be expected to score 425 or lower on the math portion of the SAT and 82.12 percent would score above 425. Figure 5.3 shows the raw score of 425, the corresponding z score, and the proportion of the normal distribution that falls below and above this z score and raw score.

z scores used with a normal distribution can also be used to figure out the proportion of scores that fall between two raw scores. For example, suppose that you got a score of 417 on the SAT math test and your friend got a score of 567. "Wow!" your friend says. "I blew you away! There must be about 50 percent of the population that scored between you and me on this test." Your ego bruised, you decide to see if your friend is right in his assessment. Here's what you need to do.

Step 1: First, convert each raw score into z scores. Recall the mean (μ = 517) and standard deviation (σ = 100) and then plug these numbers into the z score formula:Your z score:

$$z = \frac{417 - 517}{100}$$

$$z = \frac{-100}{100}$$

$$z = -1.00$$

Your friend's z score:

$$z = \frac{567 - 517}{100}$$

$$z = \frac{50}{100}$$

$$z = .50$$

Step 2: Now that we have the z scores, we need to look in Appendix A to find the percentage of the normal distribution that falls between the mean and a z score of −1.00, and the percentage of the distribution that falls between the mean and a z score of .50. Notice that we are dealing with one negative and one positive z score in our example. Our z score table tells me that 84.13 percent of the scores in a normal distribution will fall below a z score of 1, so 34.13 percent (i.e., .8413 − .50, because 50 percent of the normal distribution will fall below the mean) of the normal distribution of scores will fall between the mean and a z score of −1.00. Similarly, 69.15 percent of the normal distribution will fall below a z score of .50, so 19.15 percent will fall between the mean and a z score of .50. To determine the total proportion of scores that fall between these two z scores, we need to add the two proportions together: .3413 + .1915 = .5328. Note that we are *adding* the two proportions together because one of the raw scores (417) was *below* the mean and the other (567) was *above* the mean.

Step 3: Admit defeat in a bitter and defensive way. "Ha ha," you say to your friend. "It is not 50 percent of the population that scored between you and me on the SAT math test. It was 53.28 percent!" (See Figure 5.4 for a depiction of the proportion of the normal distribution that is contained between these two z scores.)

Finally, we can use z scores and percentile scores to determine the proportion of scores in a normal distribution that fall between two raw scores on the same side of the mean. For example,

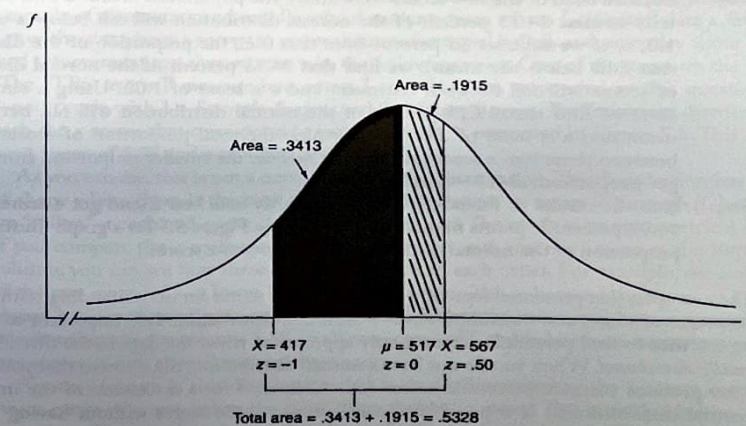

FIGURE 5.4 Proportion of scores in a distribution between two raw scores.

suppose you have another friend who got a raw score of 617 on the SAT math test. Now you want to determine the proportion of the population that scored between 567 and 617 on the test. Here is what you do.

Step 1: First, convert the raw scores into z scores. Recall the mean ($\mu = 517$) and the standard deviation ($\sigma = 100$) and then plug these numbers into the z score formula.

Friend 1's z score:

$$z = \frac{567 - 517}{100}$$

$$z = \frac{50}{100}$$

$$z = .50$$

Friend 2's z score:

$$z = \frac{617 - 517}{100}$$

$$z = \frac{100}{100}$$

$$z = 1.00$$

Step 2: Now that we have the z scores, we need to look in Appendix A to find the percentage of the normal distribution that falls between the mean and a z score of 1.00, and the percentage of the distribution that falls between the mean and a z score of .50. Notice that now we are dealing with two positive z scores in our example because both of the raw scores were above the population mean. Our z score table tells us that 84.13 percent of the normal distribution will fall below a z score of 1.0, so if we subtract 50 percent from that (i.e., the proportion of the distribution that falls below the mean), we find that 34.13 percent of the normal distribution of scores will fall between the mean and a z score of 1.00. Using a similar process, we find that 19.15 percent of the normal distribution will fall between the mean and a z score of .50. To determine the total proportion of scores that fall between these two z scores, we need to *subtract* the smaller proportion from the larger proportion: .3413 − .1915 = .1498.

Step 3: Rub the results in Friend 1's face. "Ha ha! My *new* best friend got a score that was 14.98 percentile points higher than yours!" (See Figure 5.5 for a graph illustrating the proportion of the normal distribution between two z scores.)

The examples just presented represent handy uses of z scores for understanding both an entire distribution of scores and individual scores within that distribution. It is important to note that using z scores to find percentile scores is only appropriate when the data in the distribution are *normally distributed*. When you do not have a normal distribution, the z scores that you calculate will *not* produce accurate percentile scores. (See Chapter 4 for a discussion of the importance of normal distributions.) It *is* possible to calculate percentile scores without having a normal distribution. To do this, you do not convert z scores to percentile scores. Rather, you rank order your data and find the score at which a certain percentage of the scores fall above and a certain

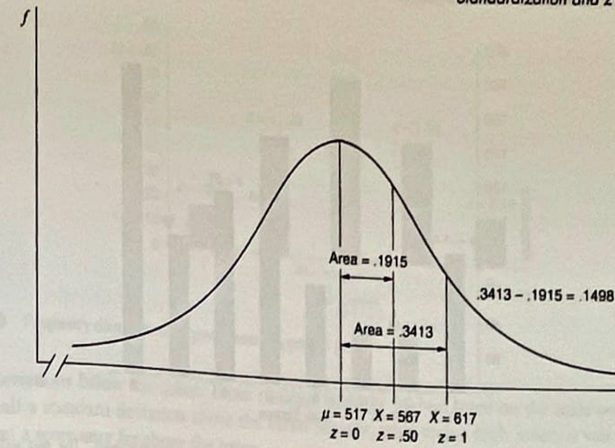

FIGURE 5.5 Proportion of scores in a distribution between two raw scores (both above the mean).

percentage fall below. This is exactly the procedure you used to find the median of a simple frequency distribution in Chapter 3. The median is, after all, simply the score that marks the 50th percentile in a distribution.

Standardized scores are used in a variety of statistics and are perhaps most helpful for comparing scores that are measured using different scales of measurement. As discussed earlier in this chapter, it is difficult to compare two scores that are measured on different scales (e.g., height and weight) without first converting them into a common unit of measurement. Standardizing scores is simply this process of conversion. In the final section of this chapter, I present and briefly describe two distributions of scores described by both raw scores and z scores.

Examples: Comparing Raw Scores and z Scores

To illustrate the overlap between raw scores and standardized z scores, I first present data from a sample of elementary and middle school students from whom I collected data a few years ago. I gave these students a survey to assess their motivational beliefs and attitudes about a standardized achievement test they were to take the following week. One of the items on the survey read, "The ITBS test will measure how smart I am." Students responded to this question using an 8-point scale with 1 = *Strongly disagree* and 8 = *Strongly agree*. The frequency distribution, along with the z scores that correspond to each raw score, is presented in Figure 5.6. This distribution has a mean of 5.38 and a standard deviation of 2.35.

As you can see, this is not a normal, bell-shaped distribution. This distribution has an odd sort of shape where there is the hint of a normal distribution in scores 2 through 7, but then there are "spikes" at each end, particularly at the higher end. The result is an asymmetrical distribution. If you compare the z scores on top of each column with the raw scores at the bottom of each column, you can see how these scores are related to each other. For example, we can see that all of the raw scores of 5 or lower have negative z scores. This is because the mean of a distribution always has a z score of zero, and any raw scores below the mean will have negative z scores. In this distribution, the mean is 5.38, so all raw scores of 5 and below have negative z scores and all raw scores of 6 or above have positive z scores.

Another feature of this distribution that is clearly illustrated by the z scores is that there is a larger range of scores below the mean than above the mean. This is fairly obvious, because the mean is well above the midpoint on this scale. The highest scores in this distribution are just a little more than one standard deviation above the mean ($z = 1.12$), whereas the lowest scores are

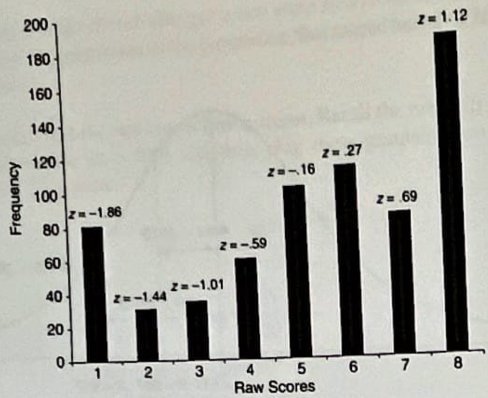

FIGURE 5.6 Frequency distribution, with *z* scores associated with each raw score, for the item "The test will show how smart I am."

nearly two standard deviations below the mean ($z = -1.86$). Finally, the inclusion of standard deviation scores with each raw score allows us to immediately determine how many standard deviations away from the mean a particular raw score falls. For example, we can see that a student who had a raw score of 3 on this variable scored just about exactly one standard deviation below the mean ($z = -1.01$).

For our second example, I have chosen a variable with a much smaller standard deviation. Using the same 8-point scale described earlier, students were asked to respond to the item "I think it is important to do well on the ITBS test." Students overwhelmingly agreed with this statement, as the mean (7.28) and relatively small standard deviation (1.33) revealed. The frequency distribution for the scores on this item is presented in Figure 5.7.

In this graph, we can see that the distribution is highly skewed, with most students circling the number 8 on the scale. Because so many students answered similarly, the standard deviation is quite small, with only relatively few scores at the lower end of the distribution. The small standard deviation coupled with the high mean create a situation where very low scores on the scale have extremely small *z* scores. For example, the few students with a raw score of 1 on the scale ($n = 7$) had *z* scores of −4.72, indicating that these students were more than 4⅔ standard

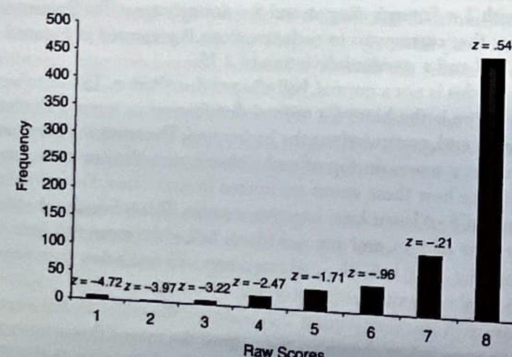

FIGURE 5.7 Frequency distribution for the item "Important to do well."

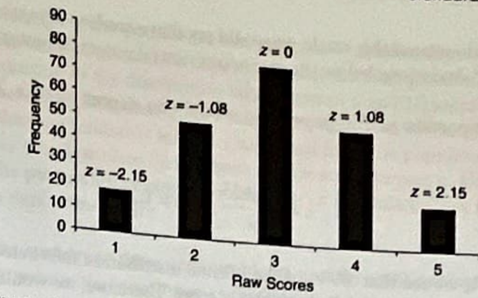

FIGURE 5.8 Frequency distribution for performance-approach goals.

deviations below the mean. Those students with the highest score on the scale were only about half a standard deviation above the mean because, with such a high mean, it was impossible to get a score very far above the mean.

The two examples provided above both illustrate the relation between the *z* scores and the raw scores for distributions that are skewed. Please note that because these data were not normally distributed, it would be inappropriate to calculate percentile scores from the *z* scores derived from these data. If you did need to calculate percentile scores from a skewed or otherwise non-normal distribution, you could use the ranking method described earlier in the chapter. In both of the distributions presented in Figures 5.6 and 5.7, the means were above the midpoint on the scale, and subsequently there was a greater range of *z* scores below than above the mean. This is not the case when the scores are normally distributed. To illustrate this, I use data from a different dataset. I used surveys to measure a sample of high school students' motivational goals in school. One goal that I measured is known as a performance-approach goal. This goal reflects a concern, or a desire, to outperform classmates and peers for the sake of demonstrating superior ability. The items on the survey were measured using a scale from 1 to 5 (1 = *Not at all true* and 5 = *Very true*). The frequency distribution is presented in Figure 5.8.

This distribution of scores had a mean of 3.00 and a standard deviation of .92. As you can see, the data are quite normally distributed. When the data are normally distributed, we would expect most of our cases to have *z* scores at or near zero because in a normal distribution, most of the cases are near the mean. Also notice that as we move farther away from the mean (i.e., *z* scores over 2.0 or less than −2.0), there are fewer cases. In a normal distribution, then, the probability of finding a particular *z* score becomes smaller as the value of the *z* score moves further away from zero. As Figures 5.6 and 5.7 illustrate, this is not always the case in skewed distributions.

Worked Examples

In this section I will work through three examples to demonstrate how to calculate and wrap words around a standard score (i.e., a *z* score), how to find the proportion of the normal distribution that would fall beyond a certain raw score, and how to find the proportion of the normal distribution that would fall between two raw scores.

Suppose that in the population of college students in the U.S. the average number of units taken per semester is 15 with a standard deviation of 4. What is the standard score for a person in this population that takes 12 units per semester?

$$z = \frac{12-15}{4}, \text{ so } z = -.75$$

Wrapping words around this result, we would say that a student who takes 12 units per semester is .75 standard deviations below the mean for units taken per semester by college students in the U.S.

Now, what proportion of the population would you expect to take 20 units or more in a semester?

$$z = \frac{20-15}{4}, \text{ so } z = 1.25$$

In Appendix A, we see that .8944 of the normal distribution falls below a z score of 1.25, so .1056, or 10.56 percent, would fall above this z score. Therefore, we would expect 10.56 percent of this population to take 20 units or more per semester.

Now, what proportion (and percentage) of the population would we expect to take between 17 and 20 units per semester? We already know that .1056 of the distribution would take 20 units or more. So now we need to calculate the z score for the raw score of 17:

$$z = \frac{17-15}{4}, \text{ so } z = .50$$

In Appendix A, we see that .6915 of the normal distribution falls below a z score of .50, and .3085 falls above. To find the proportion of the distribution that falls *between* the z scores of 1.25 and .50, we subtract one proportion from the other:

$$.3085 - .1056 = .2029$$

Now we can wrap words around our results: .2029, or 20.29 percent, of this population would be expected to take between 17 and 20 units per semester.

For brief videos demonstrating how to calculate and interpret z scores, probabilities, and percentile scores, please refer to the website that accompanies this book.

Wrapping Up and Looking Forward

z scores provide a handy way of interpreting where a raw score is in relation to the mean. We can use z scores to quickly and easily determine where an individual score in a distribution falls relative to other scores in the distribution, either by interpreting the z score in standard deviation units or by calculating percentile scores. Using the table of probabilities based on the normal distribution presented in Appendix A, we can also use z scores to determine how unusual a given score in a distribution is (i.e., the probability of obtaining an individual score of that size when selecting the individual at random). In the next chapter, I will use information about the mean, standard deviation, normal distributions, z scores, and probability to explain one of the most important concepts in statistics: the standard error.

Work Problems

Suppose you know that in the population of full-time employees in the United States, the average number of vacation days taken off work per year is 10 with a standard deviation of 4. Please answer the following questions using this information, assuming that the number of vacation days taken forms a normal distribution.

1. What percentage of the population takes at least 7 vacation days off per year?
2. What is the number of vacation days that marks the 30th percentile of this distribution?
3. What proportion of the distribution takes between 6 and 10 vacation days off per year?
4. What percentage of the distribution takes between 6 and 15 vacation days off per year?
5. Suppose that you randomly select an individual from the population who takes 20 vacation days off per year. Calculate that person's z score and interpret it. What does it tell you?
6. What is the probability of randomly selecting an individual from the population who takes 8 vacation days or more off?

For the answers to these work problems, and for additional work problems, please refer to the website that accompanies this book.