

1

INTRODUCTION TO STATISTICS AND FREQUENCY DISTRIBUTIONS

LEARNING OBJECTIVES

- Explain how you can be successful in this course.
- Explain why many academic majors require a statistics course.
- Describe the four pillars of scientific reasoning.
- Use common statistical terms correctly in a statistical context.
- Identify how a variable is measured.
- Use software to generate frequency distribution tables and graphs.
- Interpret frequency distribution tables, bar graphs, histograms, and line graphs.
- Explain when to use a bar graph, histogram, and line graph.

HOW TO BE SUCCESSFUL IN THIS COURSE

Have you ever read a few pages of a textbook and realized you were not thinking about what you were reading? Your mind wandered to topics completely unrelated to the text, and you could not identify the point of the paragraph (or sentence) you just read. For most of us, this experience is not uncommon even when reading books that we've chosen to read for pleasure. Therefore, it is not surprising that our minds wander while reading textbooks. Although this lack of focus is understandable, it seriously hinders effective

learning. Thus, one goal of this book is to discourage minds from wandering and to encourage reading with purpose. To some extent, you need to force yourself to read with purpose. As you read each paragraph, ask “What is the purpose of this paragraph?” or “What am I supposed to learn from this paragraph?” If you make asking these questions a habit your reading comprehension will improve.

This text is structured to make it easier for you to read with purpose. The short chapters have frequent reading questions embedded in the text that make it easier for you to remember key points. Resist the temptation to go immediately to the reading questions and search for answers in the preceding paragraphs. *Read first, and then answer the questions as you come to them.* Using this approach will increase your understanding of the material in this text.

Reading Question

1. Reading with purpose means
 - a. thinking about other things while you are reading a textbook.
 - b. actively trying to extract information from each paragraph as you read.

Reading Question

2. Is it better to read the paragraph and then answer the reading question or to read the reading question and then search for the answer? It's better to
 - a. read the paragraph, then answer the reading question.
 - b. read the reading question, then search for the question's answer.

The chapter readings introduce the material giving you a foundation in the basics before you attempt the activities at the end of the chapter. The activities give you the opportunity to assess your knowledge of the chapter material by working with research problems. Your emphasis when working on the activities should be on understanding your answers. If you generate a wrong answer, figure out your error. We often think of errors as things that should be avoided at all costs. However, quite the opposite is true. Making mistakes and fixing them is a great way to learn. If you find your errors and correct them, you will probably not repeat them. Resist the temptation to “get the right answer quickly.” It is more important that you understand why every answer is correct.

Reading Question

3. When completing activities, your primary goal should be to get the correct answer quickly.
 - a. True
 - b. False

MATH SKILLS REQUIRED IN THIS COURSE

The math skills required in this course are fairly basic. You need to be able to add, subtract, multiply, divide, square numbers, and take the square root of numbers using a calculator. You also need to be able to do some basic algebra. For example, you should be able to solve the following equation for X : $22 = \frac{X}{3}$. [The correct answer is $X = 66$.]

Reading Question

4. This course requires basic algebra.
- True
 - False

Reading Question

5. Solve the following equation for X : $30 = \frac{X}{3}$.
- 10
 - 90

You will also need to follow the correct order of mathematical operations. As a review, the correct order of operations is (1) the operations in parentheses, (2) exponents, (3) multiplication or division, and (4) addition or subtraction. Some of you may have learned the mnemonic, Please Excuse My Dear Aunt Sally, to help remember the correct order. For example, when solving the equation $(3 + 4)^2$, you would first add $(3 + 4)$ to get 7 and then square the 7 to get 49. Try to solve the next more complicated problem. The answer is 7.125. If you have trouble with this problem, talk with your instructor about how to review the necessary material for this course.

$$X = \frac{(6-1)3^2 + (4-1)2^2}{(6-1) + (4-1)}$$

Reading Question

6. Solve the following equation for X : $X = \frac{(3-1)4^2 + (5-1)3^2}{(3-1) + (5-1)}$
- 11.33
 - 15.25

You will use statistical software or a calculator to perform computations in this course. You should be aware that order of operations is very important when using your calculator. Unless you are very comfortable with the parentheses buttons on your calculator, we recommend that you do one step at a time rather than trying to enter the entire equation into your calculator.

Reading Question

7. Order of operations is only important when doing computations by hand, not when using your calculator.
- True
 - False

Although the math in this course should not be new, you will see new notation throughout the course. When you encounter this new notation, relax and realize that the notation is simply a shorthand way of giving instructions. While you will be learning how to interpret numbers in new ways, the actual mathematical skills in this course are no more complex than the order of operations. The primary goal of this course is teaching you to use numbers to make decisions. Occasionally, we will give you numbers solely to practice computation, but most of the time you will use the numbers you compute to make decisions within a specific, real-world context.

If you ever do research in a subsequent class or for a career after you graduate, you almost certainly will NOT conduct your statistical analyses by hand. Rather, you will use statistical software. It is reasonable to ask, “Why am I crunching numbers by hand if nobody does hand calculations anymore?” Many statistics instructors believe that some amount of number crunching is necessary to understand foundational concepts like variability and expected sampling error. These concepts are so fundamental to this course that computing them by hand helps you understand and subsequently interpret statistical results more accurately than if you don’t understand the math behind these concepts. There is less agreement among statistics educators on how much number crunching is necessary. In this text, we think computing numerous forms of sampling error by hand is helpful, but after you understand these concepts it’s probably more beneficial to focus on learning how to use statistical software.

STATISTICAL SOFTWARE OPTIONS

There are many statistical software packages available, each with advantages and disadvantages. Two widely used packages are IBM SPSS Statistics and R (r-project.org). SPSS, which stands for Statistical Package for the Social Sciences, offers a point-and-click interface that makes running common statistical analyses relatively easy. However, SPSS is expensive and, unless students purchase or rent their own copy of the program, it can be difficult to access off campus. The R statistical package is completely free but is substantially more complicated to use. Rather than having a point-and-click interface, it requires typing code. If you learn how to write the R code (or syntax) you have tremendous flexibility in the statistical analyses you can run; learning the code requires a level of sophistication that presents challenges for introductory statistics students, however. Fortunately, there are other options that offer a point-and-click interface and that are also free. Both JASP (jasp-stats.org) and jamovi (jamovi.org) offer point-and-click interfaces that are more intuitive than SPSS’s interface and both are completely free. Additionally, both JASP and jamovi perform important statistics that SPSS does not perform. For this reason, we recently switched from using SPSS to JASP in our introductory statistics courses. Table 1.1 compares the current capabilities of JASP, jamovi, SPSS, and R. You can find instructions for running common statistical procedures in all three programs online, although you

TABLE 1.1 Comparison of JASP, jamovi, SPSS, and R Statistical Programs

	JASP Version 0.13.1.0	jamovi Version 1.1.9.0	SPSS Version 25	R
Easy point-and-click interface	☑ ⁺	☑ ⁺	☑	
Sample mean, <i>N</i> , <i>SD</i> , and <i>SEM</i>	☑	☑	☑	☑
Mean difference, <i>SD</i> , and <i>SEM</i>	☑	☑	☑	☑
<i>t</i> tests and correlations with <i>p</i> value	☑	☑	☑	☑
Effect size (<i>d</i>)	☑	☑		☑
Confidence interval around mean			☑	☑
Confidence interval around mean difference	☑	☑	☑	☑
Confidence interval around effect sizes	☑	☑		☑
Factorial ANOVA simple effects analysis without requiring syntax	☑	☑		
Price	Free	Free	\$\$\$	Free

Reading Question

8. This textbook can be used with a variety of different statistical programs, some of which are free.
- True
 - False

WHY DO YOU HAVE TO TAKE STATISTICS?

You are probably reading this book because you are required to take a statistics course to complete your degree. Students majoring in business, economics, nursing, political science, pre-medicine, psychology, social work, and sociology are often required to take at least one statistics course. There are many reasons why statistics is a mandatory course for students in these varied disciplines. But the primary reason is that in every one of these disciplines people make decisions that have the potential to improve people's lives, and these decisions should be informed by data. For example, a psychologist may conduct a study to determine if a new treatment is effective for depression. Based on this

study, the psychologist will need to decide if the treatment is or is not effective. If the wrong decision is made, an opportunity to help people with depression may be missed. Even more troubling, a wrong decision might harm people. While statistical methods will not eliminate wrong decisions, understanding statistical methods will allow you to reduce the number of wrong decisions you make. You are taking this course because the professionals in your discipline recognize that statistical methods can improve decision making. Statistics make us more professional.

Reading Question

9. Why do many disciplines require students to take a statistics course? Taking a statistics course
 - a. is a way to employ statistics instructors, which is good for the economy.
 - b. can help people make better decisions in their chosen professions.

THE FOUR PILLARS OF SCIENTIFIC REASONING

People in many professions rely on statistical evidence to make scientifically supported decisions. For example, medical, sociological, economic, and psychological researchers all rely on statistical evidence when they recommend one course of action (e.g., a drug treatment, a social policy, a tax policy, or a learning strategy) rather than another. When the statistical evidence is evaluated thoughtfully, there is little doubt that these statistically justified decisions benefit our society with healthier people, better schools, and better government. However, all too frequently statistical evidence is not evaluated thoughtfully. Unfortunately, statistical evidence can be intentionally misrepresented or unintentionally misinterpreted. Either leads to less-than-optimal, even detrimental, outcomes (e.g., higher death rates, lower graduation rates, or less trust in government). Without being overly dramatic, thoughtfully interpreted statistical evidence can be a positive force for improvement, but misinterpreted statistical evidence hinders, even obstructs, progress. Clearly, our society would be better off if more of us could thoughtfully interpret statistical evidence.

Constructing accurate scientific conclusions based on statistical evidence can be tricky, but two simple recommendations can help. First, when constructing a scientific conclusion, use multiple types of evidence. You will learn that different statistical procedures have different advantages and potential problems. Therefore, by using multiple statistical procedures to inform the same scientific decision you gain the advantages of each; if used properly, these differing advantages can compensate for the potential weakness of any single statistical procedure. If constructing a sound scientific conclusion were like constructing a house, using multiple statistical procedures would be like using multiple different types of materials (e.g.,

wood, glass, and aluminum) to make a house rather than just one material. A house incorporating the advantages of wood, glass, and aluminum is generally more desirable than one constructed from just one type of material. Similarly, a scientific conclusion constructed from multiple statistical procedures possesses more desirable characteristics. Second, when constructing a scientific conclusion, evaluate the strength of your evidence and adjust your conclusion accordingly. Grand scientific conclusions, those made with great certainty, require extraordinary evidence. When your evidence is weak, your conclusions should be tentative. The best scientific conclusions are not overly certain but rather appropriately certain; they convey uncertainty when the evidence is uncertain.

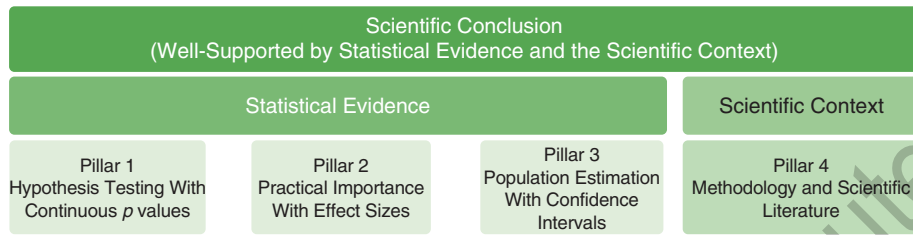
Within any scientific context, using multiple types of evidence and adjusting the certainty of one's conclusions based on the evidence's strength are sound recommendations. Within the medical and social sciences, following this advice frequently means using hypothesis testing, effect sizes, and confidence intervals and then interpreting these statistics based on the methodology used to generate them and their consistency with the scientific literature. The first three are different types of statistical evidence, each offering its own strengths and weaknesses. We will discuss each of these in depth in later chapters. Consistent with using multiple kinds of evidence, we advocate using these three statistical procedures jointly when answering scientific questions and then interpreting (i.e., contextualizing) them based on methodology and scientific literature. Collectively, as Figure 1.1 suggests, these four types of evidence form the foundation of a sound, appropriately certain, scientific conclusion. In fact, we view these four sources of evidence as so foundational to scientific decision-making that from this point forward we refer to them as the four pillars of scientific reasoning. For clarity, the four pillars of scientific reasoning are

1. hypothesis testing with a continuous p value,
2. practical importance with effect size,
3. population estimation with confidence intervals, and
4. research methodology and scientific literature.

Using these four pillars when making scientific decisions will likely lead to well-supported, scientific conclusions. We will be discussing each of these pillars in greater detail in subsequent chapters, but for now we offer a brief description of each.

The first pillar is hypothesis testing with a continuous p value, *a formal multiple-step procedure for evaluating a null hypothesis*. This statistical procedure is also called **significance testing** or **null hypothesis testing**. In later chapters, you will learn a variety of statistics that test different hypotheses. All of the hypothesis-testing procedures that you will learn are needed because of one fundamental problem that

FIGURE 1.1 Concept Map Depicting How the Four Pillars of Scientific Reasoning Support a Sound Scientific Conclusion



plagues all researchers—the problem of sampling error. For example, researchers evaluating a new depression treatment want to know if it effectively lowers depression in all people with depression, called the population of people with depression. However, researchers cannot possibly study every depressed person in the world. Instead, researchers have to study a subset of this population, perhaps a sample of 100 people with depression. *The purpose of any sample is to represent the population from which it came.* In other words, if the 100 people with depression are a good sample, they will be similar to the population of people with depression. Thus, if the average score on a clinical assessment of depression in the population is 50, the average score of a good sample will also be 50. Likewise, if the ratio of women with depression to men with depression is 2:1 in the population, it will also be 2:1 in a good sample. Of course, you do not really expect a sample to be exactly like the population. *The differences between a sample and the population create **sampling error**.*

Reading Question

10. All hypothesis-testing procedures were created so that researchers could
 - a. study entire populations rather than samples.
 - b. deal with sampling error.

Reading Question

11. If a sample represents a population well, it will
 - a. respond in a way that is similar to how the entire population would respond.
 - b. generate a large amount of sampling error.

While hypothesis testing is extremely useful, it has limitations. Therefore, another primary purpose of this book is to describe these limitations and how researchers address them by using two additional statistical procedures. **Effect sizes** describe the magnitude of a study's results, helping researchers determine if a research result is large enough to be useful or if it is too small to be meaningful in real-world situations. In other words, effect sizes help researchers determine the *practical importance* of a study's results. Another valuable statistical procedure is a confidence interval. **Confidence intervals** estimate the plausible values that might occur in the population based on the likely sampling error.

Copyright ©2022 by SAGE Publications, Inc.

in the study. Samples rarely represent the population perfectly; consequently, confidence intervals help researchers estimate how a study's results might generalize to the entire population. Each of these statistical procedures help researchers give meaning to the results of a hypothesis test. In fact, the American Psychological Association (APA) Publication Manual recommends that researchers use effect sizes and confidence intervals whenever hypothesis tests are used (American Psychological Association, 2020). These three statistical procedures are most beneficial when they are used side by side.

Reading Question

12. Effect sizes and confidence intervals help researchers
 - a. interpret (i.e., give meaning to) the results of hypothesis tests.
 - b. address the limitations of hypothesis tests.
 - c. Do both of the above.

After computing the hypothesis test, effect size, and confidence intervals, you need to **contextualize these statistical results** by evaluating the methodological rigor of the study and comparing the results to the existing scientific literature. If a study's methodology had limitations, your confidence in the statistical results should be limited as well. A major methodological flaw might even make the statistical results meaningless. Accurately interpreting statistical results requires careful consideration of the methodology used to generate those results. Contextualizing results also requires comparing your results to those in the scientific literature. You should have more confidence in results that have been replicated many times than you have in those that contradict the existing scientific literature. While contradictory results are not necessarily wrong, it is wise to interpret them cautiously when they contradict well-established theory or a large body of empirical findings.

Reading Question

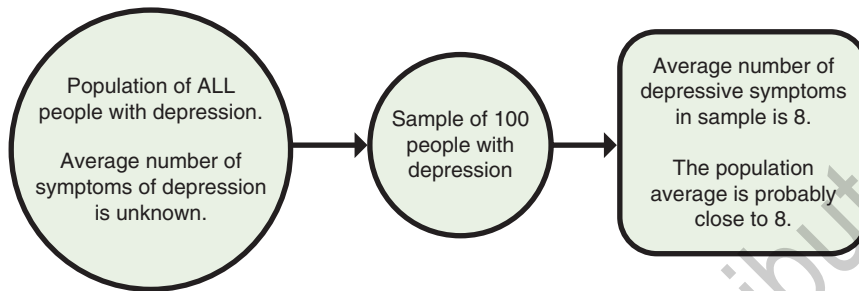
13. Before constructing a final scientific conclusion, you should consider
 - a. the statistical evidence provided by hypothesis tests, effect sizes, and confidence intervals.
 - b. the methodological strengths and weaknesses of the study that generated the data.
 - c. the existing scientific literature.
 - d. All of the above.

POPULATIONS AND SAMPLES

Suppose that a researcher studying depression gave a new treatment to a sample of 100 people with depression. Figure 1.2 is a pictorial representation of this research scenario. The large circle on the left represents a **population**, *a group of all things that share a set of characteristics*. In this case, the “things” are people, and the characteristic they

Copyright ©2022 by SAGE Publications, Inc.

FIGURE 1.2 A Pictorial Representation of Using a Sample to Estimate a Population Parameter (i.e., Inferential Statistics)



all share is depression. Researchers want to know what the mean depression score for the population would be if all people with depression were treated with the new depression treatment. In other words, researchers want to know the **population parameter**, *the value that would be obtained if the entire population were actually studied*. Of course, the researchers don't have the resources to study every person with depression in the world, so they must instead study a **sample**, *a subset of the population that is intended to represent the population*. In most cases, the best way to get a sample that accurately represents the population is by taking a **random sample** from the population. When taking a **random sample**, *each individual in the population has the same chance of being selected for the sample*. In other words, while researchers want to know a population parameter, their investigations usually produce a **sample statistic**, *the value obtained from the sample*. The researchers then use the sample statistic value as an estimate of the population parameter value. The researchers are making an *inference* that the sample statistic is a value similar to the population parameter value based on the premise that the characteristics of those in the sample are similar to the characteristics of those in the entire population. *When researchers use a sample statistic to infer the value of a population parameter* it is called **inferential statistics**. For example, a researcher studying depression wants to know how many depressive symptoms are exhibited by people in the general population. He can't survey everyone in the population and so he selects a random sample of people from the population and finds that the average number of symptoms in the sample is eight (see Figure 1.2). If he then inferred that the entire population of people would have an average of eight depressive symptoms, he would be basing his conclusion on inferential statistics. It should be clear to you that if the sample did not represent the population well (i.e., if there was a lot sampling error), the sample statistic would NOT be similar to the population parameter. In fact, **sampling error** is defined as *the difference between a sample statistic value and an actual population parameter value*.

Reading Question

14. The value obtained from a population is called a
- statistic.
 - parameter.

Reading Question

15. Parameters are
- always exactly equal to sample statistics.
 - often estimated or inferred from sample statistics.

Reading Question

16. When a statistic and parameter differ,
- it is called an inferential statistic.
 - there is sampling error.

The researchers studying depression were using inferential statistics because they were using data from a sample to infer the value of a population parameter. The component of the process that makes it inferential is that researchers are using data they actually have to estimate (or infer) the value of data they don't actually have. In contrast, researchers use **descriptive statistics** *when their intent is to describe the data that they actually collected*. For example, if a clinical psychologist conducted a study in which she gave some of her clients a new depression treatment and she wanted to describe the average depression score of only those clients who got the treatment, she would be using descriptive statistics. Her intent is only to describe the results she observed in the clients who actually got the treatment. However, if she then wanted to estimate what the results would be if she were to give the same treatment to additional clients, she would then be performing inferential statistics.

Reading Question

17. Researchers are using descriptive statistics if they are using their results to
- estimate a population parameter.
 - describe the data they actually collected.

INDEPENDENT AND DEPENDENT VARIABLES

Researchers design experiments to test if one or more variables cause changes to another variable. For example, if a researcher thinks a new treatment reduces depressive symptoms, he could design an experiment to test this prediction. He might give a sample of people with depression the new treatment and a placebo treatment to another sample of people with depression. Later, if those who received the new treatment had lower levels of depression, he would have evidence that the new treatment reduces depression. In this experiment, the type of treatment each person received (i.e., new treatment vs. no treatment) is the **independent variable (IV)**. This experiment's IV has two **IV levels**: (1) the

new treatment and (2) the placebo treatment. The main point of the study is to determine if the two different IV levels were differentially effective at reducing depressive symptoms. More generally, *the IV is a variable with two or more levels that are expected to have different impacts on another variable*. In this study, after both samples of people with depression were given their respective treatment levels, the amount of depression in each sample was compared by counting the number of depressive symptoms in each person. In this experiment, the number of depressive symptoms observed in each person is the **dependent variable (DV)**. Given that the researcher expects the new treatment to work and the placebo treatment not to work, he expects the new treatment DV scores to be lower than the placebo treatment DV scores. More generally, *the DV is the outcome variable that is used to compare the effects of the different IV levels*.

Reading Question

18. The IV (independent variable) in a study is the
 - a. variable expected to change the outcome variable.
 - b. outcome variable.

In true experiments, those in which researchers manipulate a variable so that some participants have one value and others have a different value, the manipulated variable is referred to as the IV. For example, if a researcher gives some participants a drug (Treatment A) and others a placebo (Treatment B), this manipulation defines the IV of Treatment as having two levels, namely, drug and placebo. However, in this text, we also use IV in a more general way. The IV is any variable predicted to influence another variable even when the IV was not manipulated. For example, if a researcher predicted that women would be more depressed than men, we will refer to gender as the IV because it is the variable that is expected to influence the DV (i.e., depression score). If you take a research methods course, you will learn an important distinction between *manipulated* IVs (e.g., Type of Treatment: Drug vs. Placebo) and *measured* IVs (e.g., Gender: Male vs. Female). Very briefly, the ultimate goal of science is to discover causal relationships, and manipulated IVs allow researchers to draw causal conclusions while measured IVs do not. You can learn more about this important distinction and its implications for drawing causal conclusions in a research methods course. It is important to keep in mind that statistics alone do not allow you to determine if an IV causes changes in a DV. An understanding of research methodology is just as important as an understanding of statistics in order to critically evaluate research claims.

Reading Question

19. All research studies allow you to determine if the IV causes changes in the DV.
 - a. True
 - b. False

IDENTIFY HOW A VARIABLE IS MEASURED

Scales of Measurement

All research is based on measurement. For example, if researchers are studying depression, they will need to devise a way to measure depression accurately and reliably. The way a variable is measured has a direct impact on the types of statistical procedures that can be used to analyze that variable. Generally speaking, researchers want to devise measurement procedures that are as precise as possible because more-precise measurements enable more-sophisticated statistical procedures. Researchers recognize four different **scales of measurement** that vary in their degree of measurement precision: (1) nominal, (2) ordinal, (3) interval, and (4) ratio (Stevens, 1946). Each of these scales of measurement is increasingly more precise than its predecessor, and, therefore, each succeeding scale of measurement allows more-sophisticated statistical analyses than its predecessor.

Reading Question

20. The way a variable is measured
 - a. determines the kinds of statistical procedures that can be used on that variable.
 - b. has very little impact on how researchers conduct their statistical analyses.

For example, researchers could describe depression using a nominal scale by categorizing people with different kinds of major depressive disorders into groups, including those with melancholic depression, atypical depression, catatonic depression, seasonal affective disorder, or postpartum depression. **Nominal scales** of measurement *categorize things into groups that are qualitatively different from other groups*. Because nominal scales of measurement involve categorizing individuals into qualitatively distinct categories, they yield **qualitative** data. In this case, clinical researchers would interview each person and then decide which type of major depressive disorder each person has. With nominal scales of measurement, it is important to note that the categories are not in any particular order. A diagnosis of melancholic depression is not considered to be “more depressed” than a diagnosis of atypical depression. With all other scales of measurement, the categories are ordered. For example, researchers could also measure depression on an ordinal scale by ranking individual people in terms of the severity of their depression. **Ordinal scales** of measurement also categorize people into different groups but on ordinal scales these groups are rank ordered. In this case, researchers might interview people and diagnose them with a “mild depressive disorder,” “moderate depressive disorder,” or “severe depressive disorder.” An ordinal scale clearly indicates that people *differ in the amount of something they possess*. Thus, someone who was diagnosed with mild depressive disorder would be less depressed than someone diagnosed with moderate depressive disorder. Although ordinal

scales rank diagnoses by severity, they do not quantify how much more depressed a moderately depressed person is relative to a mildly depressed person. To make statements about how much more depressed one person is than another, an interval or ratio measurement scale is required. Researchers could measure depression on an interval scale by having people complete a multiple-choice questionnaire that is designed to yield a score reflecting the amount of depression each person has. **Interval scales** of measurement *quantify how much of something people have*. While the ordinal scale indicates that some people have more or less of something than other people, the interval scale is more precise, and indicates *how much* of something someone has. Another way to think about this is that for interval scales the intervals between categories are equivalent whereas for ordinal scales the intervals are not equivalent. For example, on an ordinal scale, the interval (or distance) between a mild depressive disorder and a moderate depressive disorder may not be the same as the interval between a moderate depressive disorder and a severe depressive disorder. However, on an interval scale, the distances between values are equivalent. For example, if people completed a well-designed survey instrument that yielded a score between 1 and 50, the difference in the amount of depression between scores 21 and 22 would be the same as the difference in the amount of depression between scores 41 and 42. Most questionnaires used for research purposes yield scores that are measured on an interval scale of measurement. **Ratio scales** of measurement also *involve quantifying how much of something people have but a score of zero on a ratio scale indicates that the person has none of the thing being measured*. For example, if people are asked how much money they earned last year, the income variable would be measured on a ratio scale: not only are the intervals between values equivalent, but also there is an absolute zero point. A value of zero means the complete absence of income last year. Because they involve quantifying how much of something an individual has, interval and ratio scales yield **quantitative** data. Interval and ratio scales are similar in that they both determine how much of something someone has but some interval scales can yield a negative number while the lowest score possible on a ratio scale is zero. Within the behavioral sciences, the distinction between interval and ratio scales of measurement is not usually very important. Researchers typically use the same statistical procedures to analyze variables measured on interval and ratio scales of measurement.

Although most variables can be easily classified as nominal, ordinal, or interval/ratio, some data are more difficult to classify. Researchers often obtain data by asking participants to answer questions on a survey. These survey responses are then combined into a single measure of the construct.

For example, participants may answer a series of questions related to depression using a Likert scale with response items 1 = *strongly disagree*, 2 = *somewhat disagree*, 3 = *neutral*, 4 = *somewhat agree*, 5 = *strongly agree*. A participant's response to one item on this type of scale is thought of as ordinal but when researchers combine a participant's responses to all of the questions into a single depression score it is typically thought of as an interval/ratio

scale of measurement. Another example may help. Suppose you took a test and each item was either right or wrong. Each question would be on an ordinal scale of measurement. If, however, you computed the percent of items you got correct, that total score would be on an interval/ratio scale. The same basic idea applies to Likert scales. The entire scale is treated differently than any single item. Although there is not complete agreement among statisticians on this issue, most researchers classify questionnaire and survey scores as interval (e.g., Carifio & Perla, 2007). Thus, in this book scores on surveys will be considered interval/ratio data.

Reading Question

21. Researchers typically treat questionnaire/survey scores as which scale of measurement?
- Nominal scale of measurement.
 - Ordinal scale of measurement.
 - Interval scale of measurement.

When trying to identify the scale of measurement of a variable, it can also be helpful to think about what each scale of measurement allows you to do. For example, if you can only count the number of things in a given category, you know that you have a nominal scale. Table 1.2 summarizes what you can do with each type of scale and provides examples of each scale of measurement.

TABLE 1.2 ● The Four Scales of Measurement, What They Allow, and Examples

Scale of Measurement	What the Scale Allows You to Do	Examples
Nominal	COUNT the number of things within different categories.	<u>Favorite pet</u> : 5 dogs, 12 cats, 7 fish, 2 hamsters.
		<u>Marital status</u> : 12 married, 10 divorced, 2 separated.
Ordinal	COUNT AND RANK some things as having more of something than others (but DO NOT QUANTIFY how much of it they have).	<u>Annual income</u> : above average, average, or below average.
		<u>Speed (measured by place of finish in a race)</u> : 1st, 2nd, 3rd, etc.
Interval	COUNT, RANK, AND QUANTIFY how much of something there is but a score of zero does not mean the absence of the thing being measured.	<u>Temperature</u> : -2° F, 98° F, 57° F; 0° F is not the absence of heat.
Ratio	COUNT, RANK, AND QUANTIFY how much of something there is and a score of zero means the absence of the thing being measured.	<u>Annual income</u> : \$25,048, \$48,802, \$157,435, etc.
		<u>Number of text messages sent in a day</u> : 0, 3, 351, 15, etc.

Reading Question

22. The scale of measurement that quantifies the thing being measured (i.e., indicates *how much* of it there is) is _____ scale(s) of measurement.
- the nominal
 - the ordinal
 - both the interval and ratio

Reading Question

23. The scale of measurement that categorizes objects into different kinds of things is _____ scale(s) of measurement.
- the nominal
 - the ordinal
 - both the interval and ratio

Reading Question

24. The scale of measurement that indicates that some objects have more of something than other objects but not how much more is _____ scale(s) of measurement.
- the nominal
 - the ordinal
 - both the interval and ratio

Discrete vs. Continuous Variables

Variables can also be categorized as discrete or continuous. A **discrete variable** is measured in whole units rather than fractions of units. For example, the variable “number of siblings” is a discrete variable because someone can only have a whole number of siblings (e.g., no one can have 2.7 siblings). A **continuous variable** is measured in fractions of units. For example, the variable “time to complete a test” is a continuous variable because someone can take a fraction of minutes to complete a test (e.g., 27.39 minutes). Nominal and ordinal variables are always discrete variables. Interval and ratio variables can be either discrete or continuous.

Reading Question

25. If a variable can be measured in fractions of units, it is a _____ variable.
- discrete
 - continuous

GRAPHING DATA

The first step in all statistical analyses is graphing because doing so helps you understand your data. For example, if you were looking at the number of siblings that college students have, you could begin by looking at a graph to determine how many siblings most students have. Inspection of the graph also allows you to find out if there is anything odd in the data file that requires further examination. For example, if you graphed the data and found that most people reported having between 0 and 4 siblings but one person reported having 20

Three basic types of graphs are (1) **bar graphs**, (2) **histograms**, and (3) **line graphs**. The names of the first two are a bit misleading because both are created using bars. The only difference between a bar graph and a histogram is that in a bar graph the bars do not touch while in a histogram they do touch. In general, use bar graphs when the data are discrete or qualitative. The spaces between the bars of a bar graph emphasize that there are no possible values between any two categories (i.e., bars). For example, when graphing the number of children in a family, a bar graph is appropriate because there is no possible value between any two categories (e.g., you cannot have 1.5 children). When the data are continuous, use a histogram. For example, if you are graphing the variable “time to complete a test,” and you were creating a bar for each minute category, the bars would touch to indicate that the variable we are graphing is continuous (i.e., 27.46 minutes is possible).

Reading Question

26. What type of graph is used for discrete data or qualitative data?
- A bar graph.
 - A histogram.

Reading Question

27. What type of graph is used for continuous data?
- A bar graph.
 - A histogram.

Reading Question

28. In bar graphs, the bars
- touch.
 - don't touch.

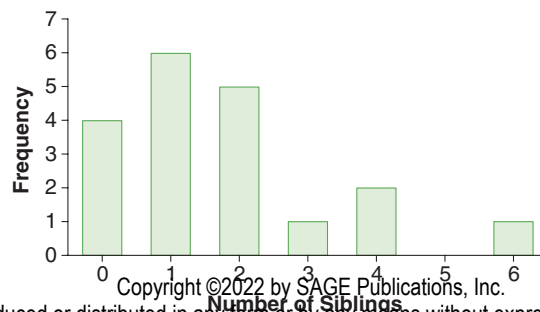
Reading Question

29. In histograms, the bars
- touch.
 - don't touch.

To create either a bar graph or a histogram, you should put categories on the x -axis and the number of scores in a particular category (i.e., the frequency) on the y -axis. For example, suppose we asked 19 students how many siblings they have and obtained the following responses:

0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 6

FIGURE 1.3 Bar Graph of Variable, Number of Siblings, Collected From a Sample of 19 Students



To graph these responses, you would list the range of responses to the question “How many siblings do you have?” on the x -axis (i.e., in this case 0 through 6). The y -axis is the frequency within each category. For each response category, you will draw a bar with a height equal to the number of times that response was given. For example, in the bar graph (Figure 1.3), 4 people said they had 0 siblings and so the bar above the 0 has a height of 4.

Reading Question

30. Use the graph to determine how many people said they had 1 sibling.
- 4
 - 5
 - 6

FIGURE 1.4 Frequency Histogram of Variable, Height in Inches, Collected From a Sample of 13 Volleyball Players

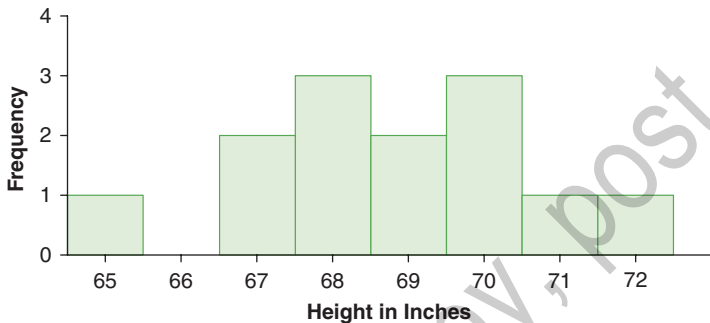
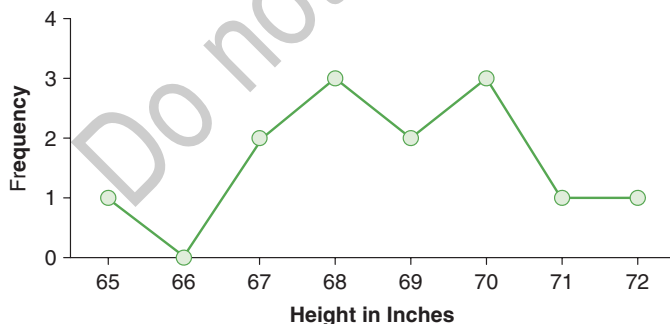


FIGURE 1.5 Frequency Line Graph of Variable, Height in Inches, Collected From a Sample of 13 Volleyball Players



The procedure for creating a histogram is similar to that for creating a bar graph. The only difference is that the bars should touch. For example, suppose that you recorded the height of players on a volleyball team and obtained the following heights rounded to the nearest inch:

65, 67, 67, 68, 68, 68, 69, 69,
70, 70, 70, 71, 72

Height in inches is continuous because there are an infinite number of possible values between any two categories (e.g., between 68 and 69 inches). The data are continuous so we create a histogram (i.e., the bars touch) as shown in Figure 1.4.

Whenever a histogram is appropriate, you may also use a **line graph** in its place. To create a line graph, you use dots to indicate frequencies and connect adjacent dots with lines (Figure 1.5).

Whether the data are discrete or continuous should determine how the data are graphed. You should use a bar graph for discrete data and a histogram or a line graph for continuous data. Nominal data should be graphed with a bar graph.

graph. Throughout the text we will use these guidelines, but you should be aware of the fact that histograms and bar graphs are often used interchangeably outside of statistics classes.

Reading Question

31. Line graphs can be used whenever a _____ is appropriate.
- histogram
 - bar graph

Reading Question

32. What type of graph should be used if the data are measured on a nominal scale?
- A histogram.
 - A bar graph.

SHAPES OF DISTRIBUTIONS

A **distribution** is a group of scores. If a distribution is graphed, the resulting bar graph or histogram can have any “shape,” but certain shapes occur so frequently that they have specific names. The most common shape you will see is a bell curve (see Figure 1.6). These bell-shaped distributions are also called *normal distributions* or *Gaussian distributions*. One important characteristic of normal distributions is that the most frequent scores pile up in the middle and as you move farther from the middle, the frequency of the scores gets less. Additionally, normal distributions are symmetrical in that the right and left sides of the graph are identical.

For the purposes of this class, you do not need to know the exact mathematical properties that define the normal curve. However, you should know that a normal curve looks bell shaped and symmetrical. You will use the normal curve frequently in this class.

The normal curve is important because many variables, when graphed, have a normal shape; this fact will be very important in later chapters. While normal curves are common, there are specific ways for graphs to deviate from a bell shape.

Some of these deviations have specific names.

FIGURE 1.6 Frequency Histogram of Test Scores That Form a Normal Curve

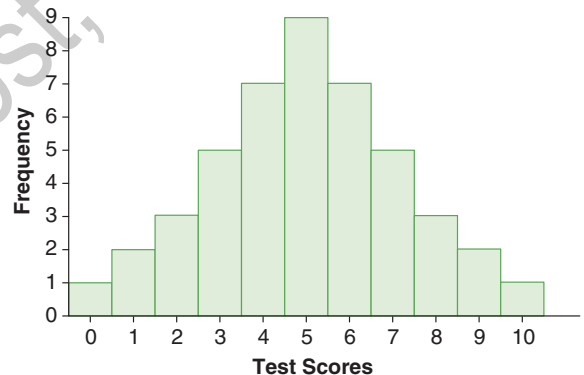


FIGURE 1.7 Positively Skewed Distribution of Scores

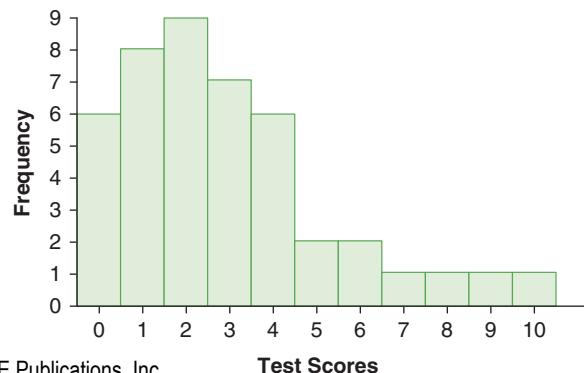
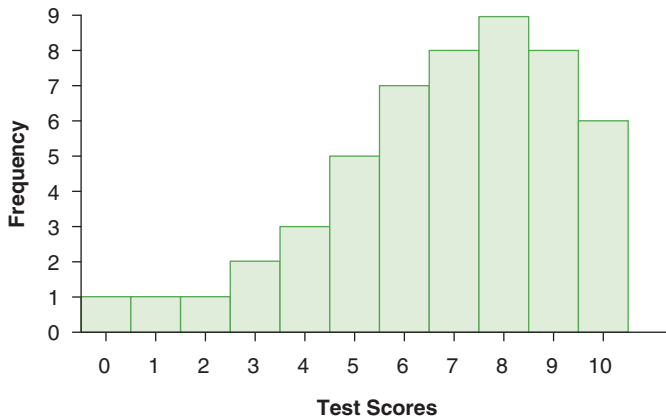


FIGURE 1.8  **Negatively Skewed Distribution of Scores**

For example, graphs can deviate from bell shaped because of **skew**. A skewed distribution is asymmetrical, meaning that the right and left sides are not identical. Instead, the scores are shifted such that most of them occur on one side of the scale with fewer scores on the other side of the scale. For example, the distributions in Figures 1.7 and 1.8 are both skewed, but in different ways. The positively skewed distribution (Figure 1.7) has the majority of the scores on the low end of the distribution with just a few scores on the higher end. The negatively skewed

distribution is the opposite. Distinguishing between positive and negative skew is as easy as noticing which side of the distribution has the longer “tail” (i.e., which side takes longer to descend from the peak to zero frequency). In positively skewed distributions, the longer tail points toward the right, or the positive side of the x -axis. In negatively skewed distributions (Figure 1.8), the longer tail points toward the left, or the negative side of the x -axis. There are statistics that you can compute to quantify exactly how skewed a distribution is (see Field, 2017, for an excellent discussion), but we will just eyeball the graphs to determine if they deviate from normal.

Reading Question

33. The scores on an exam are distributed such that most scores are low (between 30% and 50%), but a couple of people had very high scores (i.e., above 95%). How is this distribution skewed?
- Positively skewed.
 - Negatively skewed.

Distributions also vary in **kurtosis**, which is the extent to which they have an exaggerated peak versus a flatter appearance. Distributions that have a higher, more-exaggerated peak than a normal curve are called leptokurtic while those that have a flatter peak are called platykurtic. Figures 1.9 and 1.10 display a leptokurtic and platykurtic distribution, respectively. As with skew, there are ways to quantify kurtosis in a distribution (again, see Field, 2017), but we will just eyeball it in this book.

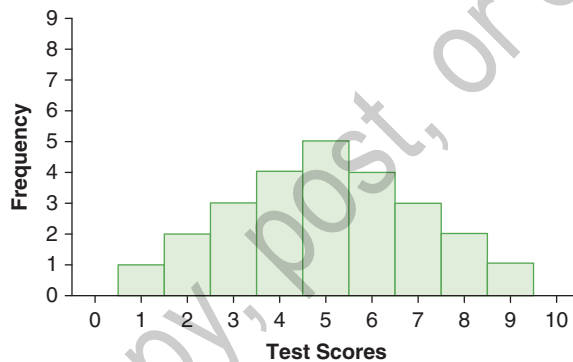
Reading Question

34. Distributions that are flatter than a normal distribution are called
- platykurtic.
 - leptokurtic.

FIGURE 1.9 Example of a Leptokurtic Distribution



FIGURE 1.10 Example of a Platykurtic Distribution



FREQUENCY DISTRIBUTION TABLES

Graphing data is typically the best way to see patterns in the data (e.g., normal, leptokurtic, or platykurtic). However, some precision is often lost with graphs. Therefore, it is sometimes useful to look at the raw data in a **frequency distribution table**. To create a frequency distribution table, you need to know the measurement categories as well as the number of responses within a given measurement category. For example, suppose that a market researcher asked cell phone users to respond to the following statement: “I am very happy with my cell phone service provider.” People were asked to respond with 1 = *strongly agree*, 2 = *agree*, 3 = *neither agree nor disagree*, 4 = *disagree*, 5 = *strongly disagree*. The responses are listed below:

1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5

It is probably obvious that a string of numbers like that one is not a particularly useful way to present data. A frequency distribution table organizes the data, so it is easier to interpret; one is shown in Table 1.3.

Copyright ©2022 by SAGE Publications, Inc.

This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

TABLE 1.3 Frequency Distribution Table of the Variable “I Am Very Happy With My Cell Phone Service Provider”

	X	f	Percent	Cumulative Percent
Strongly agree	1	2	8.70	8.70
Agree	2	4	17.39	26.09
Neither agree nor disagree	3	7	30.43	56.52
Disagree	4	6	26.09	82.61
Strongly disagree	5	4	17.39	100.00

The first column (X) represents the possible response categories. People *could* respond with any number between 1 and 5, therefore the X column (i.e., the measurement categories) must include all of the *possible* response values, namely 1 through 5. In this case, we chose to put the categories in ascending order from 1 to 5, but they could also be listed in descending order from 5 to 1.

The next column (f) is where you record the frequency of each response. For example, 4 people gave responses of 5 (*strongly disagree*) and so a 4 is written in the “ f ” column across from the response category of 5 (*strongly disagree*).

The next column (percent) is simply the number of responses in each category (i.e., the frequency f) divided by the total number of responses (i.e., $N = 23$). For example, 4 people responded with “strongly agree” so $4/23 = 17.39\%$. The percent column is a useful way to compare the relative number of responses in each category.

The final column is the cumulative percent or percentile. A percentile score is the percent of respondents with a score equal to or less than a given score. For example, in Table 1.3 26.09% of the scores are equal to or lower than the score of 2 “Agree.” So, the score of 2 is at the 26.09 percentile. In this example, this means that 26.09% of the sample agree or strongly agree with the statement, “I am happy with my cell phone provider.”

Reading Question

35. The value for “ f ” represents the
- number of measurement categories.
 - number of responses within a given measurement category.

Reading Question

36. In Table 1.3 how many people responded with an answer of 3?
- 2
 - 4
 - 7

**Reading
Question**

37. What score is at the 56.52 percentile?
- 2
 - 3
 - 4

**Reading
Question**

38. What percent of the sample disagreed or strongly disagreed with the statement “I am happy with my cell phone provider”?
- 82.61%
 - 26.09%
 - 43.48%

Activity 1.1

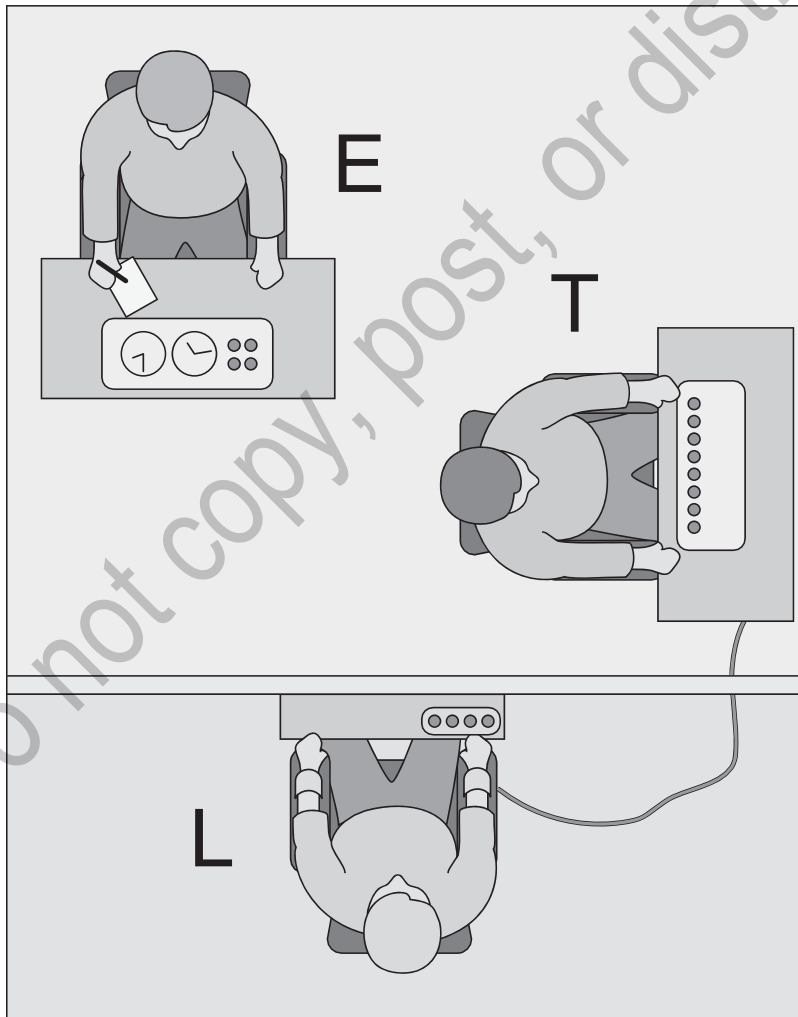
How Obedient Were Milgram’s Participants?

You are probably somewhat familiar with Stanley Milgram’s (1974) famous studies on obedience to authority. If you have never seen a video of one of his studies, it would be helpful to watch one that demonstrates the procedures. There is a link to a video on the textbook website, but you can find videos online by searching for “Milgram Experiment.” What you may not know is that Milgram conducted many studies using the same basic procedures but varying different elements of the study. In the first study, men between the ages of 20 and 50 were recruited to participate in a one-hour study of memory at Yale. When each participant arrived at the laboratory, he met an Experimenter as well as another participant. In fact, both the Experimenter and the other participant were confederates hired by Milgram to play specific roles in the study. The Experimenter said the study’s purpose was testing a theory of learning suggesting that people learn information more effectively when they receive punishment for their memory error. To lend credence to this cover story, the Experimenter showed the participants a book that presumably contained this theory and he stated that the theory suggested that parents could help children learn to behave by spanking them. He continued the cover story by “explaining” that it is unclear based on current scientific studies how much punishment is necessary for the most effective learning. The Experimenter said one participant would serve as the Teacher in the study and one participant would serve as the Learner. He asked the participants if they preferred being the Learner or the Teacher. They were both allowed to express their preferences, but ultimately the Experimenter decided that the only fair way was to draw from a hat. Of course, the draw was rigged so that the real participant was always the Teacher and the confederate was always the Learner. Once the roles were assigned, the Learner/confederate was strapped into a chair and an electrode was placed on his wrist. The participant was then led to another room and the learning task was explained to him. In this task, the Teacher read a set of word pairs. Later there was a testing session where the Learner indicated which words were paired together by

Copyright ©2022 by SAGE Publications, Inc.

ACTIVITY FIGURE 1.1 Shock Labels in Milgram's Studies

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30			
15	30	45	60	75	90	105	120	135	150	165	180	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405	420	435	450			
VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS	VOLTS			
SLIGHT SHOCK	---	---	---	MODERATE SHOCK	---	---	---	STRONG SHOCK	---	---	---	VERY STRONG SHOCK	---	---	---	INTENSE SHOCK	---	---	---	EXTREME INTENSITY SHOCK	---	---	---	DANGER SEVERE SHOCK	---	---	---	X	X	X		
																														W	W	W



pressing a switch. Whenever the Learner got an answer wrong, the Teacher was instructed to punish the Learner by giving him increasingly severe shocks that were labeled from 1 (15 Volts, Slight Shock) to 30 (450 Volts, XXX).

Before the study began, participants were given a sample shock at 45 volts (Level 3) so that they would believe that they were shocking the Learner. The Experimenter stayed in the room with the Teacher during the study. If the Teacher expressed any reluctance about shocking the Learner, the Experimenter prodded him with the following increasingly directive phrases: “Please continue,” “The experiment requires that you continue,” “It is absolutely essential that you continue,” and “You have no other choice, you must go on.”

In the first study, the Teacher was in another room and could not speak to the Learner/confederate. The only communication occurred when the Learner pressed a switch to indicate his answer. However, after being “shocked” with 300 volts, the Learner pounded on the wall. After 315 volts the pounding ceased, and the Learner did not press the switch again. Of course, the real purpose of the study was to determine how much “shock” the participant would administer to the Learner/confederate when asked to do so by an authority figure.

Measurement

1. The dependent variable in this study is how far each Teacher went on the shock scale. A score of 0 would mean that they were unwilling to administer any shocks while a score of 30 would mean that they went to the very end of the scale (450 volts labeled XXX). The fact that each numerical label on the scale increased by exactly 15 volts suggests that this “shock administered” variable is measured on a _____ scale of measurement. (nominal, ordinal, or interval/ratio)
2. As shown in Activity Figure 1.1, in addition to the numerical labels (e.g., 15 volts, 30 volts) the Teacher saw verbal labels that grouped the specific voltages into increasingly severe punishments (e.g., slight shock, moderate shock, strong shock, etc.). These verbal labels represent the _____ scale of measurement. (nominal, ordinal, or interval/ratio)
3. Teachers thought they were giving shocks by pressing levers that increased in steps of 15 volts. There were no possible responses between any two levers. Does this suggest that the dependent variable is discrete or continuous?

Statistics, Parameters, and Sampling Error

4. In Experiment 1, Milgram recruited a sample of 40 men between the ages of 20 and 50 from Connecticut. If we use the results from this sample to infer what the results would be if he had collected data from the entire population of 20- to 50-year-old men, would the purpose of the study be descriptive or inferential?

5. If the _____ of participants represented the _____ of men in Connecticut between the ages of 20 and 50 well, the study would have a relatively small amount of sampling error.
 - a. population; sample
 - b. sample; population
6. If the study has a relatively small amount of sampling error, the sample _____ will be similar to the population _____.
 - a. statistics; parameters
 - b. parameters; statistics

Interpreting Frequency Distribution Tables and Graphs

7. Before beginning this study, Milgram asked a group of psychiatrists to predict the maximum number of volts people would administer. Download the Milgram data file on the textbook website and create a frequency distribution table and graph of the variable *PsychiatristPredictions*. How many psychiatrists made predictions?
8. What was the highest level of “administered shock” that a psychiatrist predicted?
9. The predictions from the psychiatrists spanned a wide range of values. What was the most common response?
10. What was the second-most-common response from psychiatrists?
11. How many psychiatrists thought that people would go higher than 13 (very strong shock; 195 volts) on the scale?
12. What percent of psychiatrists thought that people would go higher than 13 (very strong shock; 195 volts) on the scale?
13. What is the psychiatrist prediction at the 87.2th percentile? In other words, what is the predicted “administered shock” value that has 87.2% of the distribution equal to or lower than it?
14. Milgram also asked college students to predict the maximum number of volts people would administer. Create a frequency distribution table and graph for the variable *CollegeStudentPredictions*. How many college students thought that people would go higher than 13 (very strong shock; 195 volts) on the scale?
15. What percent of college students thought that people would go higher than 13 (very strong shock; 195 volts) on the scale?

16. The actual data from Milgram's first study are in the Milgram data file. Create a frequency distribution table and graph for the *ExperimentOne* variable. How many people participated in the study?
17. Look at the graph you created. Many statistical analysis programs exclude categories with no responses from being listed on the x -axis. In this study, the possible categories ranged from 0 to 30. In your program, are all categories from 0 to 30 listed on the x -axis or are some response categories omitted? This is important because graphs look very different if response categories are omitted from the x -axis.
18. The numbers in the data file refer to the maximum shock level each person gave. Milgram defined "obedient" as going all the way to Level 30 (450 volts, XXX). What percent of people went all the way to Level 30 in Experiment 1?
19. The highest level that one psychiatrist predicted that people would go to was Level 20. How many people went higher than Level 20 in Experiment 1?
20. What percent of people went higher than Level 20 in Experiment 1?
21. The highest level that one college student predicted that people would go to was Level 14. What percent of people went higher than Level 14 in Experiment 1?
22. In Experiment 1 there was no verbal communication from the Learner after the testing session began. In Experiment 2 (done with different participants), the Teacher could hear the Learner protest through the walls. The protests became increasingly vehement as the shocks increased until all sound ceased at 330 volts and the Learner stopped responding/pressing buttons during the memory test. Create a frequency distribution table and graph for Experiment 2. What percent of people went all the way to Level 30 in Experiment 2?
23. Experiments 1 and 2 yielded similar numbers of people who were classified as "obedient" in that they delivered the maximum level of shocks. However, the two distributions of scores are not identical. For which experiment did some people stop shocking the Learner sooner?
24. In Experiments 1 and 2, the Learner and Teacher were in different rooms. In Experiment 3, the Learner and Teacher were in the same room. Create a frequency distribution table and graph for Experiment 3. What percent of people went all the way to Level 30 in Experiment 3?
25. Choose the best summary of the results from Experiments 1–3.
 - a. Participants were less obedient to authority than was expected. Most administered fairly low levels of shock. A small minority administered severe shocks. Neither communication nor physical proximity impacted the levels of shock administered to the Learner.

- b. Participants were less obedient to authority than was expected. Very few people administered severe shocks. Most people gave fairly moderate levels of shock. When participants were able to communicate with the Learner or were in close proximity to the Learner, most participants administered very mild shocks.
- c. Participants were more obedient to authority than was expected. Many administered shocks at the highest levels. Surprisingly, when participants were able to communicate with the Learner or were in close proximity to the Learner, they administered more-severe shocks.
- d. Participants were more obedient to authority than was expected. Many administered shocks at the highest levels. When participants were able to communicate with the Learner or were in close proximity to the Learner, they administered less-severe shocks.

More About Milgram's Studies and Data Collection Ideas

Criticism of Milgram's work often focuses on the many ethical issues associated with these studies (e.g., Baumrind, 1964; Perry, 2013). In addition to these legitimate ethical concerns, modern researchers also question the rigor of the research methods used (Griggs, 2017). For example, although there was a script that the Experimenter was supposed to follow, he often went off script and improvised (Perry, 2013). Milgram was also selective in choosing which results to publish. For example, in one study the Learner and Teacher knew each other (family, friends, and neighbors). In those cases, disobedience rose to 85% (Perry, 2013). Although Milgram's studies were purportedly about obedience to authority, some modern researchers have argued that Milgram wasn't studying blind obedience to authority, but rather an "engaged followership" where the participants identify with the experimenter and believe they are helping the scientific community (Haslam et al., 2016). Clearly, Milgram's work has had a dramatic impact on psychology and popular culture, but we should approach these studies with the same caution as we do any studies.

- 26. In addition to learning from the legitimate ethical concerns raised by Milgram (1974), we should also learn the importance of
 - a. recognizing how changing the experimental procedure (e.g., having the Teacher know the Learner personally) can dramatically change the results and the subsequent interpretation. Researchers should be careful not to construct overly simplistic or overly broad conclusions (e.g., people tend to blindly follow authority).

- b. considering the methodological rigor of data collection procedures used in a study when interpreting its results. If a study has poor procedural control (e.g., an Experimenter going off script) its results will be less reliable.
- c. Both answers are important considerations when interpreting scientific results.

Milgram conducted additional studies that were not included in this activity. The data file includes data from Experiment 4 where Milgram brought the Teacher and the Learner even closer together. In this study, the Learner was in the same room as the Teacher AND the Teacher had to physically touch the Learner and place his hand on an electric shock plate. Milgram also collected data with women as the Teachers, and these data are included in the data file as well. Try to guess how you think people would respond in these studies and then look at the data to see how people actually responded. Alternatively, you could ask people to guess how far people would go on the shock scale and see how close their guesses are to the actual data.

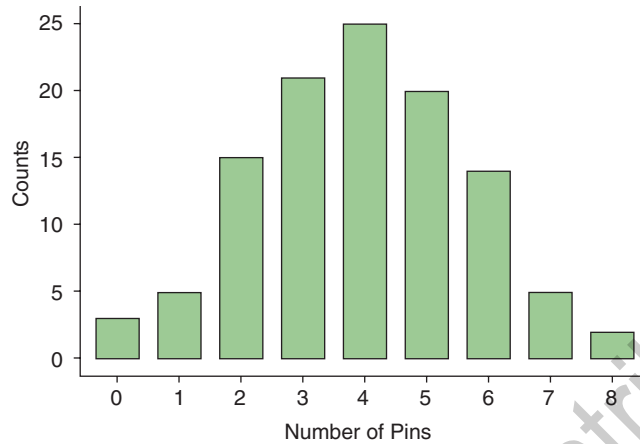
27. Optional question (ask your instructor if this is required): Investigate the data from the two additional experiments included in the provided data file. How would you summarize the results from these studies? Try to compose a brief paragraph of these results. What important conclusions (if any) can you draw from these data?

Activity 1.2

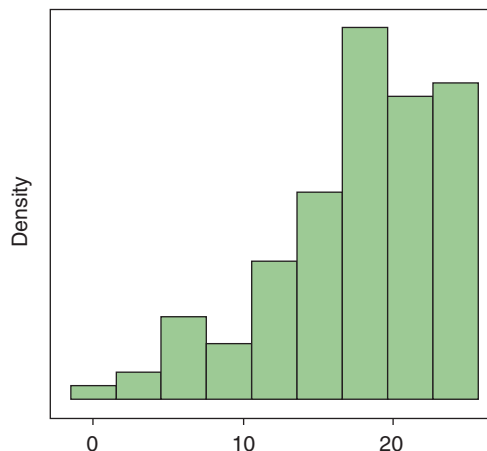
How Do Psychologists Measure Aggression?

Milgram used the level of “shock administered” as a measure of obedience to authority. Other researchers, using different research paradigms, also used “shock administered” as a measure of aggression (e.g., Parrott & Zeichner, 2002). Below are examples of different ways that researchers have measured aggression.

1. DeWall et al. (2013) measured aggression with the voodoo doll task. In this task, participants get a voodoo doll and are told they can “release their negative energy” by stabbing the doll with pins. The researchers count the number of pins stuck in the doll as a measure of aggression. Is number of pins discrete or continuous?
2. Is number of pins used nominal, ordinal, or interval/ratio?
3. A researcher obtains the following data on the number of pins stuck in a voodoo doll from a group of participants. Are these data best described as normal, platykurtic, leptokurtic, positively skewed, or negatively skewed?



4. Suppose that a researcher obtained data on the voodoo doll task from 100 participants and found that the number of pins used was negatively skewed. Does this suggest that most people used more pins or fewer pins? It might help to draw a negatively skewed distribution to visualize the results.
5. Lieberman et al. (1999) measured aggression with the Hot Sauce Allocation task. Participants get hot sauce and are told to decide how much hot sauce someone else should eat. The amount of hot sauce is measured in grams. Is the amount of hot sauce discrete or continuous?
6. Is the scale of measurement for the amount of hot sauce nominal, ordinal, or interval/ratio?
7. A researcher obtains the following data on the amount of hot sauce given to a group of participants. Are these data best described as normal, platykurtic, leptokurtic, positively skewed, or negatively skewed?



8. In addition to these behavioral measures (number of pins and grams of hot sauce), aggression is also measured using paper and pencil measures. One such measure is the Aggression Questionnaire (Buss & Perry, 1992). One section of this questionnaire assesses Physical Aggression with nine different questions (e.g., If somebody hits me, I hit back; I get into fights a little more than the average person). Participants respond to each of the nine questions on a 5-point scale where 1 = *extremely uncharacteristic of me* and 5 = *extremely characteristic of me*. This 5-point scale is an example of a _____ scale of measurement. (nominal, ordinal, or interval/ratio)
9. Researchers do not typically analyze just one question. Instead, they sum a participant's responses to all nine questions; therefore, an individual's score on the entire nine-question aggression scale could range between a low of 9 to a high of 45. Although each item on the questionnaire is measured on an ordinal scale, when researchers sum individual items intended to measure the same thing (e.g., physical aggression) the resulting value is commonly treated as a more precise scale of measurement (e.g., Carifio & Perla, 2007). This would be a _____ scale of measurement. (nominal, ordinal, or interval/ratio)

Researchers who study aggression also look at other variables that are associated with aggression. All the variables below are associated with aggression. Identify each as nominal, ordinal, or interval/ratio.

10. Gender of the participants.
11. Age of the participants.
12. Outside temperature (Fahrenheit).
13. Education level (less than high school, completed high school, some college, bachelor's degree, advanced degree).
14. Amount of alcohol consumed (in ounces) per week.
15. Testosterone levels measured as below average, average, and above average.

Do not copy, post, or distribute