3377 Bayview Avenue
Toronto, ON
M2M 3S4

TEL:
416.226.6620
www.tyndale.ca

Ross, Nancy. *Using Introductory Statistics in the Social Sciences*. Rev. ed. Redding, Calif.: BVT Publishing, 2017.

Using INTRODUCTORY STATISTICS *in the Social Sciences*

1st Edition, Revised

Nancy Ross

BVT Publishing
Academic Publishing Division

Publisher & Director of Business Development: Richard Schofield
Product Development Manager: Shannon Conley
Production and Fulfillment Manager: Janai Escobedo
Textbook Development Coordinator: Tiffany Koopal
Custom Textbook Cover Designer/Typesetter: Suzanne Schmidt
Textbook Specialist: Chad Peterson

Cover photo credit: Shutterstock

eBook[Plus] ISBN: 978-1-5178-0254-7
Loose-Leaf ISBN: 978-1-5178-0255-4
Soft Cover ISBN: 978-1-5178-0253-0

BR032017SS

Contents

3 Measures of Central Tendency 45

4 Measures of Variability 69

5 Z-Scores and the Standard Error 87

CHAPTER 1

Introduction to the Language of Statistics for the Social Sciences

SYNOPSIS

1.1 Introducing the Language of Statistics

- What Are the Terms Used to Describe Different Kinds of Data Sets?
- What Are Variables?
- What Is Random Assignment and How Does It Relate to the Groups in a Study?
- What Does It Mean to Say a Hypothesis Describes Relationships between Variables and/or Constructs?
- What Are the Four Scales of Measurement?
- How Do Survey Questions Sometimes Change the Scale of Measurement for a Variable?
- What Are Discrete and Continuous Scales?

1.2 Taking It Further

- What If a Study Does Not Have a True Independent Variable?
- What If a Study Has More Than One Independent Variable?

There are at least three kinds of social science students who find themselves taking a statistics class. There are those who gladly anticipate the class, because math has always been a strength for them. There are those who enter the course with fear and trepidation, either because math has *not* been a strength for them, or because of the negative stereotypes associated with the word ‘statistics’. And then, of course, there are those who have no idea what to expect from the course and so they have little anticipation, nor worry!

All three of these types of students are likely to find themselves taking a research methods course at some point, in addition to a statistics course. All three types of students will find that a statistics course prepares them well for the quantitative thinking, analysis and writing that will be necessary in order to carry out a research project. For those who are enthusiastic about a math course, you will certainly find some of what you are hoping for. For those who are dreading a math course, you may be pleasantly surprised to find the amount of writing that will be necessary in order to master the course material. Understanding and using statistics is about more than numbers; it is about the appropriate application, interpretation and communication of those numbers in some organized and meaningful way. This chapter provides an introduction to many of the key terms and concepts that must be mastered in order to begin to develop these skills.

1.1 Introducing the Language of Statistics

Statistics are commonly understood to be facts or figures that are presented to represent a **data set**. In baseball, people talk about the frequency of wins or losses, 'at bats' or 'runs batted in' and it is clear that these are statistics. Newscasters report the popularity rating of the current political candidates in the latest poll. The weather forecast provides the temperature but also the record high or record low for this date in the past. In the classroom, a professor might tell you the average exam grade, or the number of extra credits earned by students during the previous section of the course. These numbers have a value in themselves as they describe the group that was measured. However, these numbers can have an even broader value, when they are used them to draw conclusions that go beyond the group at hand.

1.1a What Are the Terms Used to Describe Different Kinds of Data Sets?

Data set
a collection of numbers that represent measured information

Population
the set of individuals of interest for a particular study, about whom you want to draw conclusions

Sample
the set of individuals that is measured, intended to represent the population

Statistic
number that describes a sample

Parameter
number that describes a population

Descriptive statistics
numbers that summarize or organize information about a sample

Inferential statistics
a set of methods and procedures used to make inferences from samples to the populations from which they were drawn

Sampling error
the discrepancy between a sample statistic and the corresponding population parameter

Statistics can more be defined more fully as a set of rules and methods used to organize, summarize or analyze one data set, often for the purposes of drawing conclusions about the larger entity from which the group was drawn. The larger entity is the **population**, and the measured group is called the **sample**.

Suppose, for example, that the Eastern Nazarene College (ENC) baseball team won 10 of their 13 home games in their most recent season. This indicates something about the ENC baseball team in that season, but it also might be used to draw inferences about ENC baseball teams in general. You might like to infer that ENC baseball teams of the future are likely to win more than three quarters of their home games. In this case, last season's baseball team would be your sample, and all ENC baseball teams would make up your population.

Or suppose that you have just heard that 46% of voters in a poll were confident in the current President's leadership. When this is presented on the news, people are meant to infer that 46% of voters in general are confident in the current President's leadership. The people who were polled make up the sample and all voters make up the population.

Different terms are used to indicate a number that is meant to represent a sample or a number that is meant to represent a population. For example, suppose you took a sample of ENC students in the social sciences and found their average GPA to be roughly 3.2. For comparison purposes, suppose you asked the registrar's office for the GPA of all ENC students, and found it to be 3.15. The number that describes your sample of ENC students from the social sciences (3.2) is referred to as a sample **statistic**. The number that describes the population of all ENC students (3.15) is called a population **parameter**.

Statistics that are used as descriptors of a sample are referred to as **descriptive statistics**. Statistics that are used to make inferences about the entity from which they were drawn, the population, are referred to as **inferential statistics**. Statistical hypothesis testing, which will constitute the majority of the material in this text, is based on a systematic procedure for carrying out inferential statistics in order to test hypotheses about our social world.

It may have occurred to you at this point that making an inference from a sample to a population is not a perfect way to assess the population. The discrepancy between the sample statistic and its corresponding population parameter is referred to as **sampling error**.

Notice that your sample of ENC students' GPAs, which included only social science students, had a slightly different average (3.2) than our population (3.15) which includes students from all majors. If you had used your sample mean as an estimate of the population mean, it would have been an imperfect measure, with a sample error of .05 on the four point GPA scale. It might have been considered a 'decent' representation of the population, but not a perfect one. So why do are samples even used?

It is often unfeasible to measure the actual population of interest in its entirety and thus a sample is the best option if you want to draw conclusions about that population. If careful research methodology is used, there are ways to increase your confidence that a sample truly represents a population. One of the best ways to generate a sample that truly represents a population is to use a random selection procedure (or **random sample**). In this type of sample, every member of the population has an equal chance of being selected to be in your sample.

Contrary to popular belief, for something to truly be chosen at random usually requires a certain amount of effort. People frequently use the word random in everyday conversation, when what they really mean is 'arbitrary'. For example, if you were to say you 'randomly chose' where to sit in the cafeteria last night at supper, this would probably not be entirely true. While you might not have been consciously aware of it, you were probably influenced by things like which seats were open closest to you, where the louder or quieter students were sitting, where you have often sat in the past, or whether students appeared to be clustered in homogenous groups (e.g. athletic teams together, choir members together, commuters together). It was not true that every available seat in the cafeteria had an equal chance of being selected; therefore, your choice was not random. On the other hand, if you were not consciously aware of these things that were influencing your decision making process, you can safely conclude you made your choice without conscious intent - that is, arbitrarily.

So how do you get a truly random sample? The first and crucial step in garnering a random sample is to track down a list of the entire population. This is where the work comes into play. If you want a random sample of ENC students, you need a list of all ENC students. If you want a random sample of all voters in the United States, you need a list of all registered voters in the United States. To truly choose your seat in the cafeteria randomly, you need a list of all available seats to choose from! Otherwise, certain seats are more or less likely to be selected than others.

Once you have your population list, there are a variety of methods by which to choose your random sample. One of those methods might be to use a Bingo-like ball-dispensing machine (which most of us do not happen to own). More commonly, researchers use a random number table that is made up of numbers that have been intentionally, randomly generated by a computer, so that each number and number sequence shows up a similar number of times. The bottom line: random selection requires intentional procedures and a certain degree of effort.

1.1b What Are Variables?

Anything that varies within a study can be considered a **variable**. However, something that is a variable in one study may not be a variable in another study. Suppose one study explored gender and choice of academic major to determine whether males or females are more likely to choose certain majors. In this study both males and females were included so gender is a variable. On the other hand, suppose a study included only females and explored the types of academic majors they selected. Since there were no males included for comparison purposes, gender was not a variable for this second study. Gender was held constant; it did not vary between participants.

All studies include variables, but true experiments include two particular kinds of variables. The variable that describes what has been manipulated is called the **independent variable**. The variable that is measured to assess the expected impact of the independent variable is called the **dependent variable**.

Studies typically have a key variable or set of variables under consideration in the exploration of the topic at hand. However, there are always additional things that vary

Random sample
every member of the population has an equal probability of being selected

Variable
anything that varies between participants or between **conditions**

Condition
one occasion of measurement

Independent variable
a variable that is manipulated or controlled by the researcher; the treatment that differs for each group

Dependent variable
a variable that is measured to determine whether changes have occurred due to the manipulation of the independent variable

within a study, that are not part of what you intended to measure, but may nonetheless impact your results. These variables which go beyond your specific predictions are referred to as **extraneous variables** ("extra" = beyond the prediction). The complications that accompany extraneous variables will be illustrated in the next section.

1.1c What Is Random Assignment and How Does It Relate to the Groups in a Study?

In a true experiment, participants are randomly assigned to two or more groups, and then something is done differently to each group (i.e. the independent variable is manipulated), and participants are measured to assess their response (the dependent variable). **Random assignment** involves putting people into groups through some intentionally unbiased means, so that every participant has an equal chance of ending up in any of the groups. When used properly, random assignment distributes the impact of any extraneous variables equally between your groups, so that any measured differences found between the groups can only be due to the manipulated variable.

You might flip a coin for each person in order to determine whether they should go in the first or second group. You might use a random number table to determine which participants go into each of your groups. The key here is that you do not use naturally occurring groups, because this method prevents you from knowing the two groups are essentially equal before you begin.

For example, suppose you wanted to know whether students do better on exams if they have taken short answer quizzes before the exam. You could take a sample of 80 students, and randomly assign them to either the group of people who take quizzes, or the group of people who do not take quizzes before the exam. In this case you are manipulating whether or not they take quizzes (the independent variable). Then, you would examine all of the exam grades (the dependent variable) to determine which group produced the higher average. Without random assignment you might instead simply compared two separate sections of the same course where one section had quizzes and the other did not. However, there are many things that vary between two separate sections of a course: the professor, the time of day it is taught, the classroom, the semester in which the course occurs, etc. If differences are found between exam grades in the two sections of the course, it will not be clear whether the result is due to the quizzes or the other extraneous variables.

Suppose you wanted to know whether multiple choice quizzes led to higher exam grades than short answer quizzes. You could randomly assign half of your 80 students to take multiple choice quizzes, and the other half to take short answer quizzes. In this case, the 'type of quiz' is the independent variable, and the subsequent 'exam grade' is the dependent variable.

In the first example, notice that one of the groups received something (quizzes) while the other group did not (no quizzes). In such a design, the group that receives something is referred to as the **experimental group**, while the group that receives nothing is referred to as the **control group**. In the second example, notice that both groups received something (quizzes), but they did not receive the exact same thing. In this design there are two experimental groups. The group that received multiple choice quizzes is one experimental group, and the group that received short answer quizzes is the second experimental group.

The exposure participants have to the independent variable is commonly referred to as a **treatment**. This is not necessarily some kind of therapeutic treatment; the term is used generically to refer to all types of exposure to an independent variable.

Extraneous variables
variables beyond the study's goal, which may interact with or complicate the measurement of variables of interest

Random assignment
assigning people to groups/conditions so that every participant has an equal chance of ending up in any of the groups

Experimental group
a group of participants in an experiment that receives a treatment (some exposure to the independent variable)

Control group
a group of participants in an experiment that do not experience any treatment (have no exposure to the independent variable)

Treatment
exposure to an independent variable

1.1d What Does It Mean to Say a Hypothesis Describes Relationships between Variables and/or Constructs?

Every research project is carried out to test a hypothesis. A **hypothesis** is a statement of what you think is true about the relationship between your variables, and therefore it articulates what you expect to find in your study. In an experiment with an independent and dependent variable, your hypothesis is a statement that expresses that the independent variable will influence the dependent variable, and often describes how it will influence that variable. For example, in the first scenario above you might have predicted: Taking short answer quizzes leads to better exam grades. In the second scenario, you might have predicted: The type of quizzes taken before an exam influences exam grades. Sometimes stating the hypothesis helps you identify your independent and dependent variables.

Be sure to note that a hypothesis is always a statement, not a question. Because the outcome of the study is unknown when the hypothesis is written, it can be tempting to format it as a question. However, that would be called a 'research question' rather than an actual hypothesis.

In the social sciences, many of the variables measured in order to test hypotheses are considered hypothetical **constructs**. They are different from physically measurable quantities, such as height, weight, temperature or age, because these are characteristics of a person or object that are externally visible, or that exist outside the mind. On the other hand, constructs, such as variables like happiness, intelligence, criminal intent, or leadership potential are all characteristics that are internal, or mental. These constructs are called 'hypothetical' because they cannot be measured physically. They are not externally observable. Nonetheless, these construct labels help describe an internal basis for observable characteristics or behaviors that typically occur together. One of the great challenges in the development of social science has been to determine how to externally measure something that is inherently internal. Those measurements are typically assessments of the observable characteristics or behaviors that are associated with the construct.

Construct: hypothetical (or intangible) concepts used to organize observations in terms of underlying (internal) mechanisms.

There is no blood test for happiness, for example. Nonetheless, people do tend to recognize it when they see it. The construct of happiness draws together and explains a collection of behaviors and experiences. It is something humans have 'constructed' (thus the name construct) in order to help understand and make sense of the world around us. Options can be generated for assessing those constructs by measuring some set of the behaviors or activities that are commonly associated with the construct. People who are happy, for example, often smile or laugh more frequently. They may be inclined to make more optimistic statements than an unhappy person. They may be more playful, or more energetic than others. These behaviors are measurable. Note that these behaviors are not happiness itself, but rather they are typical external expressions of happiness. A person could choose to engage in all of these behaviors, and yet not be happy. Imagine a comedian, for example, who finds it a part of his job to behave in ways that appear happy, whether or not his heart is in it at the time. Therefore, a measurement of happiness in terms of a specific set of observable behaviors is not a perfect measurement, but it can be a decent approximation most of the time. The quality of such a measurement is also dependent on whether you measure *enough of* the *right* indicators of the construct. Is it enough, for example, to measure happiness by the frequency with which someone laughs? If you add the frequency with which people make optimistic statements to your picture of happiness, are those two indicators enough to distinguish between happy and unhappy people? Should you also measure a person's self-reported happiness level, perhaps, on a scale from 1 to 10? What about having a friend also rate

Hypothesis
statement that predicts the outcome of a study. For an experiment, it is a statement that indicates that the independent variable will influence the dependent variable.

Construct
hypothetical (or intangible) concepts used to organize observations in terms of underlying (internal) mechanisms

that person on that same scale so that you have a peer report as well as self-reported happiness? Any combination of the above could be used, and you can likely think of additional alternatives as well.

When you decide exactly how you will measure a construct for a particular study, you have what is called an **operational definition** for that construct. Various operational definitions might work for different research purposes. So individual researchers must operationally define a construct, such as happiness, in order for others to know how they have *operationalized* the construct in their context.

Another example that illustrates the value of using different operational definitions is the construct of intelligence. Many different Intelligence Quotient (IQ) tests exist today. Why are there so many choices? This is because intelligence is one of those underlying mechanisms which can be expressed in behavior, but which is not physically measurable itself. Different IQ tests vary in their length, content, and purposes. When someone criticizes an IQ test, they are questioning whether the items on the test accurately assess intelligence, that is, whether that particular IQ test is a good operational definition of intelligence.

Scores on an IQ test are not the only option available as an indicator of intelligence. Perhaps a more entertaining indicator might be scores on a Jeopardy-like game, where participants demonstrate their knowledge by using certain clues to draw out that information. Or, in the college setting, sometimes GPA might be used as an operational definition for intelligence. Cumulative GPA, semester GPA, or major GPA could be chosen. These are all different ways you might operationalize intelligence.

1.1e What Are the Four Scales of Measurement?

Once you have decided the variable or variables you will use to operationalize your construct, you can determine the scale or level of measurement involved for each. Every variable also corresponds to a specific scale of measurement. These levels range from the most basic to the most sophisticated kind of categorization. Mastering these scales of measurement is very important for a statistics course, since it impacts our choice of statistical analysis once you get to the hypothesis testing process.

The simplest type of variable is on the **nominal scale** of measurement. For this scale, categories are distinguishable by their names, but there is no specific order in which those categories can be arranged from lowest to highest. There is no numerical value associated with those categories. For example, a person's gender can be categorized as 'male' or 'female' but there is no quantitative basis for ranking one gender over the other. Eye color is another example; people can be grouped according to whether they might have blue, brown, green, or hazel colored eyes. There is no reason to say that 'brown' is higher or lower than 'hazel', or 'green' is higher or lower than 'blue'. One's choice of academic major is another nominal variable. Biology is not higher or lower than Movement Arts. Psychology is not higher or lower than Literature.

The next level of measurement adds the element of order to the scale. **Ordinal scales** have categories that can be put into ascending order. However, these categories are not necessarily all equal distances apart as you move along the scale. For example, consider an employee's position within a company. Positions might include trainee, entry level, supervisor, manager, Vice President, CEO. These positions can be put in order, but the conceptual jump from trainee to entry level employee might not be equal to the jump from manager to Vice-President. This might be evidenced by the fact that people typically stay in these positions for different lengths of time, and raises can be very different when a person is promoted from one position to the next.

A scale that does have equal intervals between its categories is called (big surprise) an **interval scale**. The Fahrenheit temperature scale is one example. The jump from 45 degrees to 46 degrees is the exact same change in temperature as the jump between 85 degrees and 86 degrees. However, interval scales are missing something called absolute zero. For interval scales, there may be no zero at all, or if it does exist on the scale

Operational definition
how a construct is measured *in one particular study*

Nominal scales
categories are distinguishable by name only, and cannot be put in an ascending or descending order

Ordinal scales
categories can be put in order from lowest to highest, but are not equal distances apart

Interval scales
categories can be put in ascending order, and have equal intervals between the categories, but the scale does not include absolute zero

it is really just a label for a particular category and does not represent the absence of the quantity at hand. Contrast the temperature scale with the income scale. On the temperature scale, zero does not mean 'no heat at all' (although sometimes it may feel like it). However, on the income scale, zero income does in fact mean 'no income at all' (an unfortunate reality for many).

Many measurement inventories created to measure hypothetical constructs are created on an interval scale. Consider a personality inventory designed to tell a person where they fall on an extraversion-introversion continuum. Scores might range from as high as +10 (very extraverted) to as low as –10 (very introverted). Here, zero would not mean 'no personality', but would indicate that a person is inclined toward a balance of extraverted and introverted tendencies.

The last and most sophisticated scale of measurement is the **ratio scale**. Ratio scales have equal ascending intervals, and also include absolute zero. Income, mentioned above, would be one example of a ratio scale. Physical quantities are typically ratio variables, such as height, weight, or age. Examples of these might include tree height, weight of a frozen yogurt order, or age of a building). Drug dosage in milligrams of a drug would be another example of a ratio scale variable. This scale is called a ratio scale because absolute zero makes it meaningful to talk about these variables in terms of ratios (e.g. that tree is twice as tall as the other, or that building is twice as old as this one, or that patient is taking twice as much as the other).

1.1f How Do Survey Questions Sometimes Change the Scale of Measurement for a Variable?

Special note should be taken when determining the scale of measurement for variables measured in survey form. Often, survey questions can turn interval or ratio variables into nominal or ordinal variables by nature of the response alternatives offered. For example, if a survey asked 'What is your age?' and then provided a space in which respondents can enter their actual age, this would be a ratio scale variable in units of 'years'.

What is your age? __________

One the other hand, if a survey asked 'What is your age?' and then gave the following categories for respondents to choose from, the variable age would have become an ordinal variable with categories which include varying ranges of ages:

What is your age?
- younger than 18 years
- 18–20 years
- 21–23 years
- older than 23 years

This variable might be more aptly named 'age category' instead of 'age', since it does not allow respondents to indicate their exact age.

Often variables like age, income, number of hours at a task, etc., are broken into categories such as these. Remember that the scale of measurement is the scale on which people respond, not necessarily the scale of the underlying variable.

1.1g What Are Discrete and Continuous Scales?

Have you ever been amused by statistics which claim that the average household has 2.5 children, or 1.5 cars, or 2.75 cell phones? These statistics are amusing because you cannot truly have half a child, car or cell phone. Some variables can meaningfully be broken into smaller and smaller units, while other variables have indivisible categories that cannot be broken into smaller parts. Consider a student's major in college. Perhaps

Ratio scales
categories can be put in ascending order, have equal intervals between categories, and includes absolute zero on the scale

she started as a Psychology major, and then changed her mind to become a Social Work major. At any given point in time, she is either (as far as the registrar's office is concerned), one or the other. She cannot officially be halfway between the two. Another example might be class rank. If Susan is first in her class, and Phillip is second in that same class, then there can be no person who is halfway between the two. Those categories are indivisible. Scales that cannot be divided into smaller units are called **discrete scales**.

Other types of variables have categories that can be divided into smaller units. For example, distance can be talked about in terms of kilometres, metres, centimeters, milimetres, etc. Theoretically, no matter how many times you divide a distance, you still have a meaningful quantity to talk about that could exist in reality. These scales are referred to as **continuous scales**.

Unlike the four levels of measurement described above, the distinction between discrete and continuous scales can sometimes break down when specific examples are considered. Consider measures of the human population, for example. This can be measured in trillions of people, or billions, millions, hundreds of thousands, and so forth. Number of people would seem to be a continuous scale, and is typically treated as such for statistical or mathematical purposes. Technically it is not true that you can divide people into an infinite number of smaller units, because you cannot have half, or part of a person. Similarly, a family cannot actually have 2.75 cell phones in reality. They might have 2, or 3, but not something part way between those two categories. Therefore, a variable is considered to be discrete or continuous based on whether it can be broken into meaningful units *to a point* rather than being *infinitely divisible* into smaller units.

1.2 Taking It Further

1.2a What If a Study Does Not Have a True Independent Variable?

Earlier a distinction was made between independent (manipulated) and dependent (measured) variables. These two terms can only be applied to studies where a true experiment has been carried out. A manipulated variable is one that is controlled by the researcher, which means that the researcher applies different treatments, or different levels of treatment to different groups or conditions. In addition to having an independent variable, true experiments involve random assignment to conditions or groups. For example, suppose a researcher wanted to compare the effectiveness of types of therapies for anxiety: cognitive behavioral therapy (CBT) or exposure therapy. In a true experiment, she might randomly assign half of the participants to the CBT condition, and half to the exposure therapy condition. In this case 'type of therapy' is a true independent variable.

However, some studies sound very much like experiments, but lack the random assignment to conditions and/or manipulation of an independent variable. These studies are called quasi-experimental. For example, suppose the researcher examining the impacts of these two types of therapy found it impossible to randomly assign people to conditions. However, she was able to instruct one therapist at her agency to use CBT, and another to use exposure therapy, but she could not control which clients worked with each therapist. Since she is manipulating the kind of therapy being employed, it can be said that she manipulated an independent variable. However, clients were not randomly assigned to work with each therapist.

A more likely scenario is that she would not be allowed to tell her colleagues which type of therapy to use. In that case, she could compare two groups of clients already receiving either CBT or exposure therapy. Thus, neither manipulation of a variable, nor random assignment would be employed.

Discrete scales

onsists of separate, indivisible categories. No values can exist between neighboring categories

Continuous scales

categories are divisible into an infinite number of smaller units

Both of these variations on an experiment are referred to as **quasi-experimental designs**.

1.2b What If a Study Has More Than One Independent Variable?

Suppose you wanted to explore the impact of journaling on the stress levels of college students. Previous research suggests that there is a connection between writing about stress and academic performance in college students. (e.g. Lumley & Provenzano, 2003; Klein & Boals, 2001). To further explore this topic, imagine that you randomly assigned participants to one of four different conditions. One group wrote about the stress of adjusting to college for three 20 minute sessions once a week for three weeks. The second group also wrote about the stress of adjusting to college, but they carried out the activity for 20 minutes every day for 21 days. The third group wrote about general life stresses instead of adjusting to college, and they wrote for three 20 minute sessions once a week for three weeks. The fourth group wrote about general life stresses for 20 minutes every day for 21 days. At the end of the semester, the average GPAs of each group were compared.

		Topic of writing	
		Adjusting to College	**General Life Stress**
Quantity of writing	Twenty minutes *once a week* for three weeks	Group One	Group Three
	Twenty minutes *every day* for three weeks	Group Two	Group Four

This study involves two separate manipulated (independent) variables. The topic of the writing was manipulated (adjustment to college vs. general life stress) and the quantity of writing was manipulated (once a week vs. every day for three weeks).

An experiment with one independent variable allows the researcher to test one hypothesis. However, an experiment with two independent variables allows the researcher to test three hypotheses. This is because it is possible that each independent variable has an effect (these are called main effects), but it also possible that they interact in their impact. Using the example above, this means that the following hypotheses could be tested:

1. The impact of writing about stressful events on students' GPAs is influenced by the topic about which they write (adjustment to college vs. general life stress).
2. The impact of writing about stressful events on students' GPAs is influenced by the quantity of writing they complete (once a week for three weeks, or every day for three weeks).
3. There is an interaction between the impact of the topic and quantity of writing in their impact on students' GPAs.

If you found that the average GPA of those who wrote about adjustment to college (Groups One and Two) was higher than the average GPA of those who wrote about general life stress (Groups Three and Four), it would appear that the topic influenced GPAs. This would mean there is a main effect of the topic of writing.

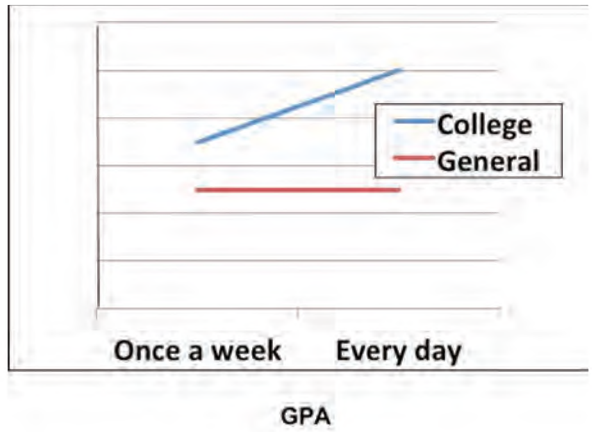
If you found that the average GPA of those who wrote once a week for three weeks (Groups One and Three) was lower than the average GPA of those who wrote every day for three weeks (Groups Two and Four), it would appear that the quantity of writing also mattered. That is, there is a **main effect** of quantity of writing.

Quasi-experimental design
a study that seems like an experiment but lacks either random assignment to conditions, or manipulation of an independent variable

Quasi-experimental variable
a variable that differs for the conditions of a study but was not manipulated by the researcher or for which participants were not randomly assigned to conditions

Main effect
significant impact of one independent variable

However, suppose the quantity of writing made a greater difference when the topic was adjustment to college than it did when the topic was general life stresses? (This hypothetical result is illustrated in the graph below). This means that there is an **interaction** between the two variables in their impact on GPAs.



Interaction
when the impact of one independent variable is moderated by the impact of another independent variable

The example above illustrates what is typically referred to as a two-way interaction, because two independent variables are involved. However, it is possible to add additional independent variables to experiments, which means it is also possible to have three-way interactions, four-way interactions, and so forth. If this seems a bit complicated, remember that research ultimately attempts to understand the causes and effects of everyday life which is also rather complicated!

CHAPTER 1 Review

Key Terms

Data set: a collection of numbers that represent measured information

Population: the set of individuals of interest for a particular study, about whom you want to draw conclusions.

Sample: the set of individuals that is measured, intended to represent the population.

Statistic: number that describes a sample

Parameter: number that describes a population

Descriptive Statistics: numbers that summarize or organize information about a sample.

Inferential Statistics: a set of methods and procedures used to make inferences from samples to the populations from which they were drawn.

Sampling error: the discrepancy between a sample statistic and the corresponding population parameter.

Random Sample: every member of the population has an equal probability of being selected.

Variable: anything that varies between participants or between conditions.

Condition: one occasion of measurement.

Independent variable: a variable that is manipulated or controlled by the researcher; the treatment that differs for each group.

Dependent variable: a variable that is measured to determine whether changes have occurred due to the manipulation of the independent variable.

Extraneous variables: variables beyond the study's goal, which may interact with or complicate the measurement of variables of interest.

Random assignment: assigning people to groups/conditions so that every participant has an equal chance of ending up in any of the groups.

Experimental Group: a group of participants in an experiment that receives a treatment (some exposure to the independent variable).

Control Group: a group of participants in an experiment that do not experience any treatment (have no exposure to the independent variable).

Treatment: exposure to an independent variable.

Hypothesis: statement that predicts the outcome of a study. For an experiment, it is a statement that indicates that the independent variable will influence the dependent variable.

Construct: hypothetical concepts used to organize observations in terms of underlying (internal) mechanisms.

Operational Definition: how a construct is measured *in one particular study*.

Nominal Scales: categories are distinguishable by name only, and cannot be put in an ascending or descending order.

Ordinal Scales: categories can be put in order from lowest to highest, but are not equal distances apart.

Interval scales: categories can be put in ascending order, and have equal intervals between the categories, but the scale does not include absolute zero.

Ratio scales: categories can be put in ascending order, have equal intervals between categories, and includes absolute zero on the scale.

Discrete scales: Consists of separate, indivisible categories. No values can exist between neighboring categories.

Continuous scales: categories are divisible into an infinite number of smaller units.

Quasi-experimental design: a study that seems like an experiment but lacks either random assignment to conditions, or manipulation of an independent variable.

Quasi-experimental variable: a variable that differs for the conditions of a study but was not manipulated by the researcher or for which participants were not randomly assigned to conditions.

Main effect: significant impact of one independent variable

Interaction: when the impact of one independent variable is moderated by the impact of another independent variable.

Summary

There are two kinds of *data sets* commonly described with numbers. One is the *population*—the entire group about which you want to draw conclusions. The other is a *sample*—a subset of the population used as a representative of the population. Numbers that describe a population are referred to as *parameters*. Numbers that describe a sample are referred to as *statistics*.

When numbers are used simply to summarize or organize a data set, they are referred to as *descriptive statistics*. When sample statistics are used to draw conclusions about a population, this process is referred to as *inferential statistics*. The degree to which our sample statistics vary from their corresponding population parameters is referred to as *sampling error*. The best way to minimize sampling error is to use a *random sample*.

Anything that varies between participants or conditions in a study is called a *variable*. Each occasion of measurement in a study is called a *condition*. A true experiment involves manipulating at least one variable, and a manipulated variable is called an *independent variable*. The variable measured for change caused by the independent variable is referred to as the *dependent variable*. Experiments also involve randomly assigning participants to conditions. If you have enough participants, *random assignment* should cause extraneous variables to be evenly distributed between groups, avoiding any bias they might otherwise induce. In addition to the variables of interest in a study, *extraneous variables* are things that vary between individuals or conditions, and could bias or distort results if they are not controlled.

Experiments always contain at least two groups or conditions. The simplest experiment involves one *experimental group*, that receives the experimental treatment, and one *control group* that does not receive treatment. Experiments can also involve multiple experimental groups. The experimental *treatment* is any exposure to the independent variable and does not need to be a treatment in the therapeutic sense.

A *hypothesis* is a statement that predicts the outcome of a study, and in an experiment it predicts how the independent variable will influence the dependent variable. Many hypotheses involve constructs which are hypothetical concepts that cannot be physically measured, but can be inferred by observable behaviors. An *operational definition* is a description of how a construct will be measured or operationalized in any particular study.

Every variable falls on one of four scales of measurement. Those four scales are *nominal* (categories are distinguishable by name only), *ordinal* (categories can be put in ascending order, but are not equal distances apart), *interval* (categories can be ordered and are equal distances apart, but the scale does not include absolute zero), and *ratio* (categories can be put in order, have equal intervals between them, and the scale includes absolute zero).

Variables can also be described as *discrete or continuous*. Discrete scales consist of indivisible categories, while categories on continuous scales can be divided into an infinite number of smaller units.

Studies with variables compare but do not involve manipulating an independent variable and/or randomly assigned people to conditions are called *quasi-experiments*. These grouping variables are referred to as *quasi-independent variables*.

Experiments can also have more than one independent variable. When this is the case it is possible to check for an effect of each variable (these are referred to as *main effects*) as well as for an *interaction* between those variables.

Multiple Choice Review

1. Suppose in your work as a criminal psychologist, you gave a survey to inmates in a medium security correctional institution in order to better understand their perspective on life. Out of the 678 inmates, you have 100 of them complete your survey. Among other things, you learned that the average age of the group who completed your survey was 33 years. The 100 people who completed your survey would be called your:
   a. Sample
   b. Population
   c. Experimental group
   d. Control group

2. You found that the average age of inmates who completed your survey was 33 years of age. If you used that to infer that the average age of inmates in that prison was 33 years, you would be doing what?
   a. Creating a random sample
   b. Inferential statistics
   c. Descriptive statistics
   d. Hypothesis testing

3. Suppose you used your results to infer that the average age of inmates in correctional institutions across the United States was 33 years of age. Now your population would be:
   a. The 100 inmates who completed your survey
   b. The 678 inmates at the correctional institution where you did your survey.
   c. All inmates in correctional institutions across the United States.
   d. All inmates in correctional institutions anywhere.

4. The average age of those who completed the survey was 33 years. This would be referred to as a ____________________. Inferring that the average age of inmates at the correctional institution in question is 33 years would be making a claim about a ____________________.
   a. Sample parameter; population statistic
   b. Population parameter; Sample statistics
   c. Population statistic; Sample parameter
   d. Sample statistic; population parameter

5. Imagine that the director of a network of social service agencies wanted to determine whether a new continuing education program improved the ability of social workers at these agencies to serve their clients. He took a sample of 100 social workers in the network, and randomly assigned half of them to participate in the program, while not including the other half. As a measure of the social workers' ability to serve their clients, he used evaluations completed by the supervisors of each social worker in the study. Six months after the program had taken place, he had the supervisors complete evaluations on each of the 100 social workers who were in the study. Those evaluations were significantly better for those social workers who had experienced the program than for those who had not. In this study, the social workers who were part of the sample, but did not participate in the program would be considered his
   a. Independent variable
   b. Dependent variable
   c. Experimental group
   d. Control group

6. Suppose that a researcher critiqued this director's study by saying that the 100 social workers chosen might have been younger than the average social worker and therefore more motivated to partake in continuing education programs. This researcher is claiming that the sample may not be
   a. Biased
   b. Representative
   c. Extraneous
   d. Experimental

7. If the director responds to the researcher by pointing out that the 100 participants selected from the network were chosen by a computer generated number table and therefore should not be different in average age from the rest of the population, he is pointing out that his sample was
   a. Deliberate
   b. Independent
   c. Random
   d. Arbitrary

8. If the director further explains that the 100 participants were assigned to their conditions by flipping a coin (those who got 'heads' were in the program, and those who got 'tails' were not), he can then claim that the ages of the two conditions were likely equal before they began the program. This is because
   a. Random sampling balances extraneous variables between conditions
   b. Random assignment balances extraneous variables between conditions
   c. Random assignment balances independent variables between conditions
   d. Random sampling balances extraneous variables between conditions

9. Imagine that a graduate student in experimental psychology wanted to know whether monkeys trained in sign language have more developed motor cortex than monkeys who are not trained in sign language. She randomly assigns 30 monkeys to be taught sign language, and the rest serve as her control group. Then monkeys in the first group are given lessons in sign language over the course of six months. MRI's taken at that point reveal more neural connections in the sign language group. In this study, the independent variable was
   a. Whether or not the monkeys were trained in sign language.
   b. The monkeys' neural connections.
   c. The experimental group
   d. The control group

10. Who makes up the control group in this example?
   a. The monkeys who learn sign language
   b. The monkeys who do not learn sign language
   c. All of the monkeys in the study
   d. The monkeys who are given MRIs

11. What is the treatment in this example?
   a. Whether or not the monkeys receive MRIs
   b. Whether or not the monkeys are taught sign language
   c. The neural connections that develop during the study
   d. Any changes in the motor cortex of the monkeys.

12. Personality is a complex contrast. In fact, entire courses can be taught on personality theory. Suppose one study measured how extraverted participants were. Another study measured whether participants were Type A or Type B personalities. These two different approaches demonstrate two different
   a. Hypotheses
   b. Independent variables
   c. Dependent variables
   d. Operational definitions

13. Which of the following would be considered a construct?
   a. Age
   b. Number of counselors
   c. Intelligence
   d. Income

14. Suppose a survey asked people to indicate their gender (male/female). What scale of measurement does this represent?
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio

15. A researcher measured participants' spatial ability on a spatial aptitude test. What scale of measurement does this represent?
   a. Nominal
   b. Ordinal
   c. Interval
   d. Ratio

16. Suppose a bunch of your friends were competing to earn coins in a video game. If someone performed well, they could earn up to a 1000 coins in a round. If they performed very badly, they could earn no coins at all. What scale of measurement does this represent?
    a. Nominal
    b. Ordinal
    c. Interval
    d. Ratio

17. A group of law school students were challenged to read a description of 20 different crimes, and put them in order from most severe to least severe. What scale of measurement does their order represent?
    a. Nominal
    b. Ordinal
    c. Interval
    d. Ratio

18. Students in Psychology of Learning had to train their rats to learn to press a lever in order to receive a food reward. They kept track of the number of training sessions they needed in order to get their rat to master the task. What scale of measurement does their number of training sessions represent?
    a. Nominal
    b. Ordinal
    c. Interval
    d. Ratio

19. A clinical psychologist specializes in treating individuals with phobias. She keeps a list of all of the different things that clients report as feared objects or situations. This list of phobias represents what scale of measurement?
    a. Nominal
    b. Ordinal
    c. Interval
    d. Ratio

20. Which of the following represents a discrete scale of measurement?
    a. Income
    b. Aptitude score
    c. Personality type
    d. Commuting distance

Homework

1. A group of students in Methods of Social Research wanted to study the relationship between GPA and participation in athletics among ENC traditional undergraduates. Other researchers have found that students involved in division three athletics often have higher GPAs than the rest of the student body and so the group of students in Methods wanted to know whether this was true for ENC students. They were able to gather data on 250 ENC students and found that, among this group, the 75 athletes had slightly higher GPAs on average than the 175 students who were not athletes.
    a. In this study, who makes up the sample? Who makes up the population? (2 pts)
    b. Suppose they found the overall average GPA for ENC athletes was 3.35 and for the ENC non-athletes was 3.21. Are these statistics or parameters? (1 pt)

2. Imagine that a social work agency wants to explore the potential need for a homeless shelter in a particular neighborhood. Since it is difficult to have an accurate estimate of the number of homeless individuals when no shelter currently exists, they collect data on the average income in that city overall, and compare that to the average income in the neighborhood where the shelter is being considered. A lower average income does not automatically mean there are more homeless individuals, but these things are often related, so they would like to show that the average income for the neighborhood is significantly lower than the overall average for the whole city. Suppose they find the average income for the city is $65,000, and the average for the neighborhood in question is $48,000.
   a. What is the sample average? What is the population average? (2 pts)
   b. What is the hypothesis that they would like to test? (1 pt)

3. Suppose you work for a restaurant that wants you to get a set of 100 customer feedback forms completed by customers over the course of the next 30 days. Provide an example of one problematic non-random way to distribute the customer feedback forms, and then provide an example of a better way to distribute them. (1 pt for non-random example, and 1 pt for better example)

4. A public school teacher is completing a research project for her master's thesis that involves assessing the benefits of using smart boards in the classroom. She believes smart boards facilitate learning and lead to higher grades. In order to support her research project, her principal gives her permission to determine who will receive a new smart board and who will stay on a waiting list. She randomly determines which classes will receive smart boards and then compares the average scores of students in each group on a standardized test appropriate for their respective grade levels.
   a. What is her independent variable in this study? (1 pt)
   b. What is her dependent variable? (1 pt)
   c. What is her hypothesis? (1 pt)
   d. Who is in her experimental group? Her control group? (2 pts)
   e. In her analysis, she also checked to see if there were significant differences in the performance of students in each of the different grade levels. If students all were performing at their best, their averages on the grade-specific tests should all be the same. The variable of 'grade level' is not a true independent variable in this analysis. Why not? (1 pt)

5. Suppose a law student wants to carry out a study to determine which law schools produce the most successful graduates. What are three ways she could operationally define 'successful'? (1 pt)

6. For each of the following variables, decide which scale of measurement is being illustrated (nominal, ordinal, interval or ratio). (1/2 pt each)
   - Elementary school (by name)
   - Agreement on a Likert scale from strongly agree to strongly disagree
   - Credit card debt (in dollars)
   - Performance rating on a scale of Poor, fair, good, excellent
   - Emotional Intelligence score (scores range from 50 through 150)
   - Graduate schools ranked by level of difficulty
   - Favorite course
   - Commute time (in minutes)

Taking It Further

7. For each of the following scenarios, decide whether a true experimental variable or a quasi-experimental variable has been used. (1 pt each)

   Scenario A: Avi was interested in exploring the relationship between college major and salary five years after graduation. He collected data from 500 graduates of his college that indicated their salaries at the five year point. He then divided people up according to 6 different academic divisions (Arts and Letters, Education, Humanities, Social Sciences, Natural Sciences and Religion), which allowed him to compare the average salaries for people whose majors fell in each of those divisions.

Scenario B: Michal was interested in the use of peer modeling as an intervention for elementary school aged children with behavior problems. She carried out a study in which three different kinds of peer modeling interventions were used. She randomly assigned her 60 children with behavior problems to each of the three conditions, and then compared their results after the programs were completed.

Scenario C: Professor Quinn wanted to compare the effectiveness of four different teaching strategies in her Lifespan Development course (lecture driven, discussion driven, activity driven, or a blended approach). Over the course of four semesters, she tried out each approach and then examined the resulting final grades. She chose which approach to use for each class randomly, using each approach only once.

8. George carried out some research on the effectiveness of different types of music therapy in the treatment of different mood disorders. He collected progress scores from clients receiving any one of three different kinds of music therapy: listening therapy, music lesson therapy, or performance therapy. Participants were clients receiving therapy for either depression or bipolar disorder. Ultimately, George hopes to learn whether certain kinds of music therapy are more effective as treatments for mood disorders generally, whether depression or bipolar disorder is more responsive to music therapy, and whether the most effective type of music therapy varies by disorder.
    a. In this study, what are the three hypotheses that could be tested with a two-way ANOVA? Which are tests of main effects, and which is the test of an interaction? (3 pts)
    b. Name the two independent variables. Is each variable a true independent variable, or a quasi-independent variable? Why? (4 pts)