18

THE LANGUAGE AND LOGIC OF STATISTICS

INTRODUCTION

When approaching any new subject we often have to learn the language of that subject matter. To better understand the logic of statistics, we need to start with some basic vocabulary, including the terms *constant*, *variable*, *population*, *sample*, *parameter*, and *statistic*. We will also examine a model that will explain how these concepts fit together and, at the same time, explain the difference between descriptive and inferential statistics. Finally, we will examine scales of measurement, or several different ways to think about numbers or data.

The basic goal of statistics is to summarize data. Let's say that you just came home from a tough day at teaching. Your spouse, who is very interested in your career, asks you, "How did your students do on that social studies test you gave today?" You respond by saying, "Alicia received an 85, Bob a 79, Carmen a 95, Dee a 75, Elaine a 90, and so on." It is likely that your tactful spouse will nod, appreciating all of that information. However, it is also likely that your spouse really only wanted a brief summary of the test results. Did the students do well? This would be especially true if you taught five different social studies classes and had a total of 125 students. Statistics gives us the tools that we need to summarize data.

BASIC LANGUAGE AND LOGIC

Constants and Variables

Let's look at some basic language of statistics starting with constants and variables. A *constant* is any value that remains unchanged. For example, the numbers 2, 10, and 31.7856 are all constants. So are the values of π (3.14159) or Avogadro's number (6.02×10^{23}). If there are 28 students enrolled in your class, then that is also a constant. In statistics it is a typical convention to symbolize constants with letters from the beginning of the Roman alphabet (a, b, c, d) or with Greek letters (α [alpha], μ [mu], σ [sigma], and ρ [rho]).

> ***Definition***
> A **constant** is any value that remains unchanged.

A *variable* is a quantity that can take on any of a set of values. For example, variables for students in a college class would include heights, weights, ages, GPAs, and scores on the first examination. Each of these would tend to vary among students. It is a typical convention to symbolize variables with letters from the end of the Roman alphabet, with x, y, and z often being used most often.

> ***Definition***
> A **variable** is a quantity that can take on any of a set of values.

Populations and Samples

Populations and samples are frequently confused with one another. A *population* is the entire set of people or observations in which you are interested or which are being studied. Examples could include everyone living in the United States (people), the ages of all of the current residents of Brownsville, Texas (observations), or the GPAs for all students enrolled at the University of Connecticut. Although we often think of populations as being quite large, they can also be small. For example, the first examination grades for all the students in your 6th-grade class can also be considered to be a population. The size of the population is the number of people or observations and is typically indicated by the uppercase letter N (for number). For example, the number of students currently enrolled at the college where I teach is 2,951. We express this population size as N = 2,951.

> ***Definition***
> A **population** is the entire set of people or observations in which you are interested or which are being studied.

A *sample* is a subset of the population. We generally use a sample when we are interested in studying a population, but when it would be too impractical, time consuming, or expensive to gather information on all of the members of the population. For example, you might be interested in how 3rd graders in Pennsylvania will perform on a new state test. Since it is cheaper and easier to use only a few schools to first try the test, the 3rd graders in those selected schools can serve as a sample for all 3rd-grade students in Pennsylvania.

For another example, we could use the students in a particular classroom as a sample of all of the students attending that school. However, that brings us to an important distinction. You could use the students in a particular class as either a sample or a population, depending on the question that is asked. If you are interested in all of the students in a school, then the students in one particular class serve as a sample of that larger population. However, if you are interested only in that particular class, then the students in that class are the population. We typically represent the size of the sample with the lowercase letter n. For example, if there are 27 students in a class, we can represent this sample size as n = 27.

> ***Definition***
> A **sample** is a subset of the population.

Although populations (at any given point in time) are stable, samples typically fluctuate. For example, let's say that we would like to determine the average age of students at a particular college. If the student directory from the college computer system contained age information, it could likely give you that average with just a few simple commands. However, if that information were not available by way of the computer, it would be difficult and time consuming to gather the ages of all of the students to calculate the average. In that case you would be likely to find a sample of students, gather their ages, compute an average, and use that as an estimate of the population average. The average age of the population (at a given point in time) would be a fixed number, a constant. Let's say that it is 20.2 years. The sample value, for any given sample, however, is unlikely to be exactly 20.2. It will typically be somewhat higher or lower. This will be especially true if the sample is from a class of first-year students or from an evening class consisting primarily of nontraditional adult students. If you gather many different samples from the same population and compute the mean on each sample, the means will vary from one sample to another.

Obviously, you would have to be cautious in how you gathered the sample. One type of sample is known as a *random sample*. In a random sample every member of the population has an equal likelihood of being selected for that sample. Essentially, you have to use some type of random process—one completely free of any possible bias. One way to develop a random sample is to assign every member of the population a unique number starting at 1 and continuing to N (the number of members in the population). You then use some type of random number generator to select numbers, as in a lottery, until you have filled the sample. Statisticians have found that, in the long run, random samples tend to do the best job of mimicking the population. They are the most likely to be representative of the population and unbiased.

> ***Definition***
> A **random sample** is one in which every member of the population has an equal likelihood of being selected.

Now, let's return to the idea of *sampling fluctuation*. Remember, earlier we said that the average age of a population would be stable, a constant. We also said that the average age of any one sample is likely to be different from the average of most other samples as well as the average of the population (even for random samples). Therefore, we say that the *sample average* displays sampling fluctuation; it varies among samples.

> ***Definition***
>
> **Sampling fluctuation** refers to the fact that any measure taken on a particular sample from a population is likely to vary among samples from that same population.

It turns out that small samples tend to display more sampling fluctuation than do larger samples. Let's say that you were to choose a sample size of three (n = 3) to estimate the average age of students at a college. You might stand outside a first-year lecture course and question the first three students who leave the classroom. If their ages were 18, 17, and 19, your average would be 18. On the other hand, let's say that you stand outside another classroom where there are more nontraditional students. The first three students to leave that classroom are 25, 20, and 30. In that case the average age would be 25. However, neither of these samples is typical. With small samples, atypical observations (very high or very low) will affect the average more than they would with larger samples. If your sample consisted of two 20-year-old and one 35-year-old student, then the sample mean would be 25. However, with a larger sample there is a greater likelihood that the atypical observation (the 35-year-old student) will be offset by more typical observations (students in their late teens and early 20s). For example, if our 35-year-old student turned up in a sample of 10 students, that student would be less likely to pull up the average. Let's say that the ages for our 10 students are 19, 35, 18, 18, 20, 21, 20, 19, 22, and 20. Then the average age for our sample would be 21.2, only one year above our hypothetical average age of 20.2.

In general, averages computed on larger samples will fluctuate less from the population average than will averages computed from smaller samples. This is an important point. Although you cannot avoid sampling fluctuation, larger samples will show less sampling fluctuation than smaller samples. Since you would like your samples to be representative of the population, you would like less sampling fluctuation. This brings us to our first general principle of statistics: *larger samples are preferred over smaller samples because they are less influenced by sampling fluctuation.*

Parameters and Statistics

Now that you understand the differences between populations and samples, we need to move on to distinguish parameters from statistics. *Parameters* are descriptive indices (e.g. a mean or standard deviation) of a population. The average age of the college population that we have been discussing is such a parameter. Parameters (at any given point in time) are stable, fixed, unvarying. Therefore, they are constants. They are conventionally indicated with either uppercase Roman letters or Greek letters. Parameters that we will be using will include N, P, μ, σ, and ρ.

Definition

Parameters are descriptive indices (e.g. a mean or standard deviation) of a *population.*

Statistics are descriptive indices (e.g. a mean or standard deviation) of samples. Statistics will vary from sample to sample. Therefore, they are variables. Statistics are typically denoted with lowercase Roman letters. Statistics that we will be using will include n, p, s, and r.

Definition

Statistics are descriptive indices (e.g. a mean or standard deviation) of *samples.*

Generally, we are more interested in parameters than statistics because we typically want to know about the characteristics of populations. There are two ways that we can go about determining population parameters. Those two routes are represented in the Figure 18.1 (Games & Klare, 1967).

The first route is the simplest and is represented on the far left side of the diagram. It involves simply and directly describing the characteristics of the population. If the population of interest is reasonably small, you can simply gather the required information from all members of the population, and from the information that you gather on all the members of the population you can calculate the population parameter (e.g. the average age). As a psychologist, I find that this is rarely feasible. Most of the populations in which I am interested, as a psychologist, are large. However, as a teacher, this route is readily available to me since most of the time the population in which I am interested is a single classroom. Therefore, teachers can frequently use Route 1.

Route 2 is somewhat more complicated. This is the route that you must take when it is impractical to make observations on all members of the population. In this case you start with the population and develop a random sample from that population. You then

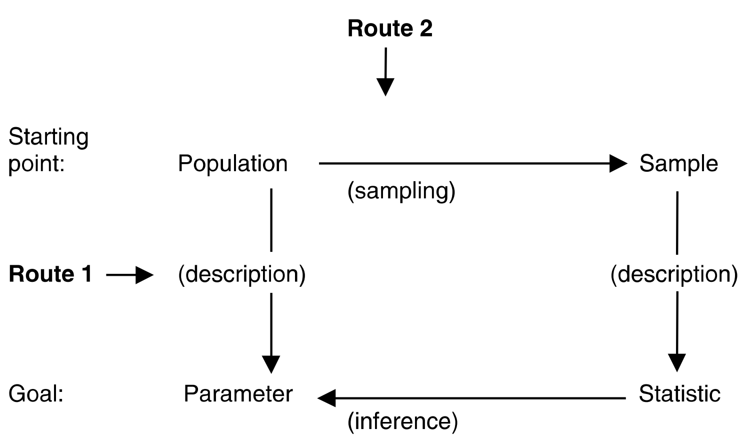

Figure 18.1 Finding population parameters.

make observations on the desired characteristic of all members of that sample and calculate the sample statistic (e.g. the sample mean). Finally, you use inferential statistics to make a probability statement concerning the likelihood that the sample statistic is representative of the corresponding population parameter. You need to ask, "Is the statistic a good estimate of the population parameter?"

This brings us to the differences between descriptive and inferential statistics. *Descriptive statistics* involves directly calculating parameters from populations or statistics from samples. (When you calculate a parameter directly from a population, that process should be referred to as descriptive **parameters**. Nevertheless, the field is known as descriptive **statistics**.) You are simply describing the appropriate population or sample. Look back at Figure 18.1. Route 1 from the diagram is an example of descriptive statistics. *Inferential statistics* involves making estimates of population parameters from sample statistics, and is represented by Route 2 on Figure 18.1. Researchers use both descriptive and inferential statistics frequently. However, for teachers, the populations in which they are interested are most typically single classrooms. Since it is relatively easy to gather information on all members of the population (the students in one's classroom), teachers are typically able to follow Route 1 and use descriptive statistics. Since this text is primarily aimed at classroom teachers, you will be dealing exclusively with descriptive statistics. If you were to take a statistics course in a psychology department or an advanced education course on statistics, you would also learn about inferential statistics since the focus would be more on research, using samples to make estimates about populations.

Definitions

The process of directly describing the characteristics of either populations or samples is referred to as **descriptive statistics**.

The process of using samples to estimate the characteristics of populations is referred as **inferential statistics**.

MEASUREMENT SCALES

Data may show up in a variety of forms. Therefore, over the years we have developed a variety of scales of measurement—terms that we use to describe data.

Categorical Data

One type of data is known as *categorical* data, sometimes also known as *nominal* data. This simply involves sorting observations into logical groups and naming those groups (hence the term "nominal" for naming). An example would include sorting the students in a class into males and females. Therefore, your class may contain 14 males and 12 females. If you decide to play an educational game in the classroom, you might sort the students into two groups, Group 1 and Group 2. Some students are given the label as members of Group 1, whereas other students are labeled as members of Group 2. Other examples of categorical data would include forming the students within your class into groups based on eye color or career aspirations. The groups can be labeled with either names (e.g. male, female) or numbers (e.g. 1, 2). However, when we use numbers to label the characteristics of categorical groups, the numbers are simply being used as substitutes for names and, therefore, do not possess the typical mathematical properties of numbers. It would not make any sense to try to manipulate the numbers mathematically since the results would be nonsensical.

> ***Definition***
>
> **Categorical** or **nominal** data involves sorting a larger group into smaller groups, not based on numbers (e.g. males and females).

Ranked Data

Another form of data is known as *ranked* or *ordinal* data. Let's say that we line up the students from a class in order by height. We could label the tallest student number 1, the next tallest number 2, and soon we have ranked all of the students. In this case the numbers really represent relative placement (first, second, third, etc.). With ranked data we frequently have unequal intervals between the numbers (unequal spaces between the rankings). For example, let's suppose that the tallest student in our 1st-grade class is Paul who is 48 inches tall. The second tallest is Sally who is 47 inches tall. Tom is third at 42 inches high. In this case number 1, Paul, is only one inch taller than number 2, Sally. However, number 2, Sally, is five inches taller than number 3, Tom. That is what we mean by unequal intervals. If we only examine the ranked data we know that Paul is taller than Sally, who is taller than Tom. However, we don't know how far apart they are in actual height.

For another example of ranked data, consider 50 high school students participating in a 10K (10 kilometer) race. As the runners come across the finish line, they are ranked 1st place, 2nd place, 3rd place, and so on. However, the distance between the finishers may vary. The 2nd-place runner may only be a split second behind the winner, whereas the 3rd-place runner might be a minute or more behind the first two finishers.

> ***Definition***
>
> **Ranked** or **ordinal** data involve sorting a group by rank (e.g. setting up a class of students by height or the order in which runners complete a race).

Since ranked or ordinal data do show unequal intervals, we are somewhat reluctant to perform mathematical manipulations on that type of data. For example, an average rank and an average height could mean something different for students in your classroom. Let's try an example to demonstrate this point. Let's say that we have a small college classroom with 11 students, some of whom play basketball. Let's look at their heights which I have arranged in order (see Figure 18.2).

If we look at the average rank in the group, it would be Roberto who is ranked sixth. He is 69 inches (5 ft. 9 in.) tall. However, if we calculate the average height, we find that the average is 70.9 inches. So, in reality, Tanya is actually closer to the average in height for

Name	Height (inches)	Rank
Mark	80	1
Amal	79	2
Kurt	78	3
Paul	77	4
Tanya	71	5
Roberto	69	6
Terri	68	7
Cheryl	67	8
Tony	66	9
Ping	63	10
Trish	62	11

Figure 18.2 The height of 11 students from a college class.

the class than is Roberto. Mathematical manipulations of ordinal or ranked data can lead to erroneous conclusions.

Numerical Data

A third type of data is known as *numerical* or *metric* data. When we use numerical data the numbers are meaningful. If we return to our previous example of our 1st-grade classroom, we could record the students' actual heights. You may recall that Paul was 48 inches tall, Sally was 47 inches tall, and Tom was 42 inches tall. Using each student's actual height, in inches, would be numerical data. All numerical data displays equal intervals, and can therefore be mathematically manipulated. Some statisticians insist on a further breakdown of numerical data into interval and ratio scales. With *ratio scales* there is a meaningful zero point. For example, height would be ratio data. Someone who is 60 inches tall would be twice as tall as someone who is 30 inches tall. However, not all scales have a meaningful zero point. For example, IQ scales have no meaningful zero point. We cannot say that someone with an IQ of 100 is twice as smart as someone with an IQ of 50. IQ scales were simply not set up that way. Scales without a meaningful zero point are known as *interval scales*. Within the context of this text, we are not likely to run into too many instances where the difference between interval and ratio scales will affect what we do or how we interpret data. The one exception will include many standardized tests such as IQ tests.

I should also point out that, on occasion, psychometricians have developed measurement scales that do not fully conform to the definitions that I have provided.

Definitions

Numeric or **metric** data involves data with equal intervals between points on the scale.

Interval data is numeric data without a meaningful zero point.

Ratio data is numeric data with a meaningful zero point.

If you are given a choice about gathering data in either a numeric or in a ranked form, I would recommend that you gather data in the numeric form. When we transform numeric data (e.g. heights) into ranks, we are frequently losing information and may come to an erroneous conclusion as we did in the height example mentioned above. As a general rule, it is better to have more information than less when making decisions.

Discrete Data vs. Continuous Data

There is still one additional distinction that we need to make concerning metric (numeric) data: That is the distinction between discrete and continuous data. *Discrete data* is data that can appear only as whole numbers. Examples would include the number of students in a class and the scores on most examinations. *Continuous data*, on the other hand, can be expressed with fractions or digits after the decimal point. Examples could include height (68.5698 inches) when measured with a very precise instrument, or when my wife instructs me to stop at the grocery store to pick up 1½ pounds of ground sirloin for a recipe we plan to make that evening. Some mathematicians object to treating discrete data as you would continuous data because you can occasionally obtain results that are meaningless. However, a good compromise comes from the idea of rounding. For example, we often treat continuous data as if it were discrete. We typically report someone's height to the nearest inch. Therefore, we could just as easily treat discrete data as if it were rounded off. If a student obtained a score of 76 on an examination (discrete data), we would say that this score of 76 represents any score between 75.50 and 76.49999. We say that we are reporting the score as 76, the midpoint of the interval between 75.50 and 76.4999.

Definitions

Discrete data involves information that can only be represented as whole numbers (e.g. the size of a family).

Continuous data involves information that can be represented as decimals and fractions (e.g. I weigh 162.3 pounds).

Of course, this also means that we need to be careful how we report continuous data. For example, we typically report ages from our last birthday. However, this may lead to errors of as much as six months when computing averages. Therefore, we should report all continuous data, such as ages, to the nearest whole number. For example, if you were 19 years and 7 months old, you should report your age as 20 years. We will use this convention for the remainder of this text. I would further recommend that you use it whenever treating data.

SUMMARY

This chapter provides an introduction to statistical concepts that will be important in understanding measurement principles. A subject matter such as statistics has its own vocabulary. This chapter described a number of important terms. For example, a constant is a value that does not change, whereas a variable is a value that can change from one individual to the next. A population is the entire set of information in which we are interested, whereas a sample is a subset of that population. A parameter describes a population, whereas a statistic describes a sample.

The chapter also described a model that helps us differentiate between descriptive and inferential statistics. When doing research, we frequently rely on inferential statistics, whereas in the classroom we will mostly use descriptive statistics.

In the field of measurement we use several different scales. Data can typically be characterized as categorical, ranked, or numeric. Numeric data can also be further described as either interval data or as ratio data. Finally, statisticians and psychometricians sometimes differentiate between discrete and continuous data.

EXERCISES

1. Which of the following are populations, which are samples, and which could be either? (Mark **P** for a population, an **S** for a sample, and an **E** for either.)

_______ a. The ages of all the students in your high school graduating class on the day you graduated for a report to the newspaper on the characteristics of your graduating class

_______ b. The total number of hours that five students from a class of 30 students study one night to find the average study time per night for the students in your class

_______ c. The entire freshman class at Texas State University, San Marcos

_______ d. Because you are doing a report on the school cafeteria, you pass out surveys randomly to students as they walk into lunch on Monday

_______ e. The grade-point average for every student in a class for a study on the characteristics of the students in that class

_______ f. The exam scores for the students in an English class when computing the characteristics of that exam

_______ g. The age of every person in the United States when calculating the average age of Americans

_______ h. The number of students living in the freshmen dormitories

_______ i. In order to compare last year's 5th-grade math curriculum to the new 5th-grade curriculum, you gather data from one 5th-grade classroom

_______ j. The shoe sizes for all of the male professors on your campus

2. Which of the following are parameters, which are statistics, and which could be either? (Mark **P** for a parameter, an **S** for a statistic, and an **E** for either.)

_______ a. The average age of the students in your class for a report on the characteristics of that class

_______ b. The average reading level for the members of your 3rd-grade class as an estimate of the average reading level for all 3rd graders in your school district

_______ c. The average standardized test scores for the students in your 4th-grade class for a report on the standardized test scores for all 4th-grade students

_______ d. Researchers surveying a random sample of American households for the number of people per American family

_______ e. The average hours of football practice per week for a newspaper article on that particular high school
_______ f. The average height of all of the players on the varsity basketball team for the team roster
_______ g. While finding certain characteristics of every student at your college, the computer information system calculates the most common age of the students is 19
_______ h. The average snowfall, in inches, that your town received in January
_______ i. The average number of shoes that the students in your fifth-period English class own
_______ j. The average IQ score of all the students in your honors program for a report to the school board on the new honors program in your school

3. Determine whether each example is dealing with measurement on a nominal scale, an ordinal scale, an interval scale, or a ratio scale. (Mark **N** for nominal, **O** for ordinal, **I** for interval, and **R** for ratio.)

_______ a. Manuel is the third student to finish the test.
_______ b. Rob is now 48 inches tall.
_______ c. Sam scored a 530 on the SAT verbal exam.
_______ d. On her way home from school, Sarah drives 35 mph.
_______ e. Ben will be the second to the last student to give a speech.
_______ f. A gym class is divided into four groups and Janice is on the blue team.
_______ g. Joe is the youngest in his class.
_______ h. Janice has an IQ score of 106 and Raymond has an IQ score of 98.
_______ i. Tyler had 15 out of 20 items correct on his math test.

SPOTLIGHT ON THE CLASSROOM

Mai Ling is a teacher at Highlands High School, and is the boys' varsity tennis coach. She recently developed a boys' tennis program in her school because previously it was only available to girls. She feels that her program is very strong and that the students involved display a balance between academics and sports. She would like to gather data to show that this is true, but needs your help. What type of data could she gather to demonstrate that the tennis program has not had a negative impact on the academic focus of the boys involved in her program? What would you recommend that she do?

STUDY TIPS

How to Read a Measurement and Statistics Text

One mistake that many students make is that they try to deal with every reading assignment in exactly the same way. They sometimes try to read a physics textbook the same way that they read a novel. However, more experienced and more successful students recognize that reading in one subject matter must be approached differently from reading in another subject matter. Novels that you read for leisure, and even most high school textbooks, have relatively few ideas on each page. College textbooks frequently contain many more ideas per page. Here are some suggestions for reading a measurement and statistics text.

- Although you can easily read 20 or more pages of a novel without a break, a measurement and statistics chapter should probably be broken down into several 8- to 10-page sections. Reading shorter sections of the text will keep you focused and improve your comprehension of the material.
- Read through the technical terminology and the definitions before you read the chapter or section to familiarize yourself with the material. Make word cards to learn the technical terminology.
- This type of material requires a high level of concentration. Read in a quiet place, free from distractions.
- Use a reading/study system that involves previewing the chapter before reading and reviewing the material after reading.
- Examples and box features are frequently as important as, or sometimes even more important than the text.
- Use a text-making system (e.g. highlighting) as you read. It is best to read a full paragraph before you go back and mark the most important material. Mark meaningful phrases rather than whole sentences or key words.
- Mark headings, main ideas, and supporting points.
- Formulas and examples also need to be marked.
- Taking notes from the text can be very helpful. You can take notes in place of highlighting.
- Predict questions that you think the instructor might ask in the margin of your textbook and underline the answers in the text.
- Do the exercises after reading the section or the chapter to apply what you have learned.[1]

NOTE

1. This material has been adapted from Van Blerkom (2009).