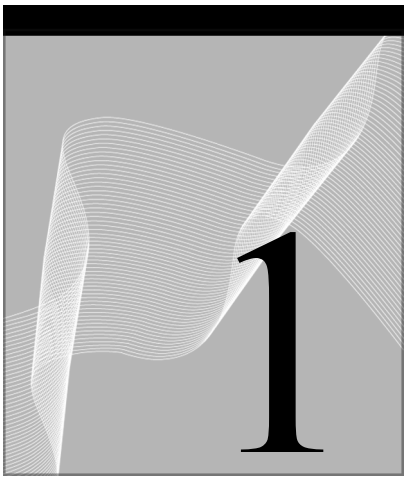


Basic Concepts in Research and Data Analysis

Introduction: A Common Language for Researchers 2

Steps to Follow When Conducting Research 3
The Research Question 3
The Hypothesis 4
Defining the Instrument, Gathering Data, Analyzing Data, and Drawing Conclusions 5

Variables, Values, and Observations 6
Variables 6
Values 7
Quantitative Variables versus Classification Variables 7
Observations 7

Scales of Measurement and JMP Modeling Types 9
Nominal Scales 9
Ordinal Scales 9
Interval Scales 10
Ratio Scales 11
Modeling Types in JMP 12

Basic Approaches to Research 12
Nonexperimental Research 12
Experimental Research 13

Descriptive versus Inferential Statistical Analysis 16
Descriptive Analyses: What Is a Parameter? 16
Inferential Analyses: What Is a Statistic? 16

Hypothesis Testing 17
Types of Inferential Tests 18
Types of Hypotheses 19
The *p* Value 22
Fixed Effects versus Random Effects 23
Summary 25
References 25

Overview. This chapter reviews basic concepts and terminology from research design and statistics. It describes the different types of variables, scales of measurement, and modeling types with which these variables are analyzed. The chapter reviews the differences between nonexperimental and experimental research and the differences between descriptive and inferential analyses. Finally, it presents basic concepts in hypothesis testing. After completing this chapter, you should be familiar with the fundamental issues and terminology of data analysis, and be prepared to learn about using JMP for data analysis.

Introduction: A Common Language for Researchers

Research in the social sciences is a diverse topic. In part, this is because the social sciences represent a wide variety of disciplines, including (but not limited to) psychology, sociology, political science, anthropology, communication, education, management, and economics. Further, within each discipline, researchers can use a number of different methods to conduct research. These methods can include unobtrusive observation, participant observation, case studies, interviews, focus groups, surveys, ex post facto studies, laboratory experiments, and field experiments.

Despite this diversity in methods used and topics investigated, most social science research still shares a number of common characteristics. Regardless of field, most research involves an investigator gathering data and performing analyses to determine what the data mean. In addition, most social scientists use a common language in conducting and reporting their research: researchers in psychology and management speak of "testing null hypotheses" and "obtaining significant *p* values."

The purpose of this chapter is to review some of the fundamental concepts and terms that are shared across the social sciences. You should familiarize (or refamiliarize) yourself

with this material before proceeding to the subsequent chapters, as most of the terms introduced here will be referred to again and again throughout the text. If you are currently taking your first course in statistics, this chapter provides an elementary introduction. If you have already completed a course in statistics, it provides a quick review.

Steps to Follow When Conducting Research

The specific steps to follow when conducting research depend, in part, on the topic of investigation, where the researchers are in their overall program of research, and other factors. Nonetheless, it is accurate to say that much research in the social sciences follows a systematic course of action that begins with the statement of a research question and ends with the researcher drawing conclusions about a null hypothesis. This section describes the research process as a planned sequence that consists of the following six steps:

1. Developing a statement of the research question
2. Developing a statement of the research hypothesis
3. Defining the instrument (questionnaire, unobtrusive measures)
4. Gathering the data
5. Analyzing the data
6. Drawing conclusions regarding the hypothesis.

The preceding steps reference a fictitious research problem. Imagine that you have been hired by a large insurance company to find ways of improving the productivity of its insurance agents. Specifically, the company would like you to find ways to increase the dollar amount of insurance policies sold by the average agent. You begin a program of research to identify the determinants of agent productivity.

The Research Question

The process of research often begins with an attempt to arrive at a clear statement of the research question (or questions). The research question is a statement of what you hope to have learned by the time you complete the program of research. It is good practice to revise and refine the research question several times to ensure that you are very clear about what it is you really want to know.

For example, in the present case, you might begin with the question

> "What is the difference between agents who sell more insurance and agents who sell less insurance?"

An alternative question might be

> "What variables have a causal effect on the amount of insurance sold by agents?"

Upon reflection, you realize that the insurance company really only wants to know what things management can do to cause the agents to sell more insurance. This realization eliminates from consideration certain personality traits or demographic variables that are not under management's control, and substantially narrows the focus of the research program. This narrowing, in turn, leads to a more specific statement of the research question, such as

> "What variables under the control of management have a causal effect on the amount of insurance sold by agents?"

Once you have defined the research question more clearly, you are in a better position to develop a good hypothesis that provides an answer to the question.

The Hypothesis

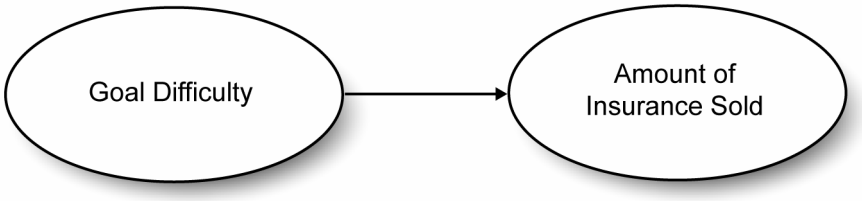
A *hypothesis* is a statement about the predicted relationships among events or variables. A good hypothesis in the present case might identify which specific variable has a causal effect on the amount of insurance sold by agents. For example, the hypothesis might predict that the agents' level of training has a positive effect on the amount of insurance sold. Or, it might predict that the agents' level of motivation positively affects sales.

In developing the hypothesis, you can be influenced by any of a number of sources, such as an existing theory, related research, or even personal experience. Let's assume that you are influenced by *goal-setting theory*. This theory states, among other things, that higher levels of work performance are achieved when difficult work-related goals are set for employees. Drawing on goal-setting theory, you now state the following hypothesis:

> "The difficulty of the goals that agents set for themselves is positively related to the amount of insurance they sell."

Notice how this statement satisfies the definition for a hypothesis: it is a statement about the relationship between two variables. The first variable could be labeled Goal Difficulty, and the second, Amount of Insurance Sold. Figure 1.1 illustrates this relationship.

Figure 1.1 Hypothesized Relationship between Goal Difficulty and Amount of Insurance Sold



The same hypothesis can also be stated in a number of other ways. For example, the following hypothesis makes the same basic prediction:

> "Agents who set difficult goals for themselves sell greater amounts of insurance than agents who do not set difficult goals."

Notice that these hypotheses have been stated in the present tense. It is also acceptable to state hypotheses in the past tense. For example, the preceding could have been stated,

> "Agents who set difficult goals for themselves sold greater amounts of insurance than agents who did not set difficult goals."

You should also note that these two hypotheses are quite broad in nature. In many research situations, it is helpful to state hypotheses that are more specific in the predictions they make. A more specific hypothesis for the present study might be,

> "Agents who score above 60 on the Smith Goal Difficulty Scale sell greater amounts of insurance than agents who score below 40 on the Smith Goal Difficulty Scale."

Defining the Instrument, Gathering Data, Analyzing Data, and Drawing Conclusions

With the hypothesis stated, you can now test it by conducting a study in which you gather and analyze some relevant data. *Data* can be defined as a collection of scores obtained when a subject's characteristics and/or performance are assessed. For example, you could choose to test your hypothesis by conducting a simple correlational study.

Suppose you identify a group of 100 agents and determine

- the difficulty of the goals set for each agent
- the amount of insurance sold by each agent.

Different types of instruments result in different types of data. For example, a questionnaire can assess goal difficulty, but company records measure amount of insurance sold. Once the data are gathered, each agent has one score that indicates difficulty of the goals, and a second score that indicates the amount of insurance the agent sold.

With the data gathered, an analysis helps tell if the agents with the more difficult goals did, in fact, sell more insurance. If yes, the study lends some support to your hypothesis; if no, it fails to provide support. In either case, you can draw conclusions regarding the tenability of the hypotheses, and you have made some progress toward answering your research question. The information learned in the current study might then stimulate new questions or new hypotheses for subsequent studies, and the cycle repeats. For example, if you obtained support for your hypothesis with the current correlational study, you could follow it up with a study using a different method, perhaps an experimental study. The difference between correlational and experimental studies is described later. Over time, a body of research evidence accumulates, and researchers can review this body to draw general conclusions about the determinants of insurance sales.

Variables, Values, and Observations

When discussing data, you often hear the terms *variables*, *values*, and *observations*. It is important to have these terms clearly defined.

Variables

For the type of research discussed here, a *variable* refers to some specific characteristic of a subject that assumes one or more different values. For the subjects in the study just described, amount of insurance sold is an example of a variable—some subjects sold a lot of insurance and others sold less. A different variable was goal difficulty—some subjects had more difficult goals, while others had less difficult goals. Age was a third variable, and gender (male or female) was yet another.

Values

A *value* refers to either a subject's relative standing on a quantitative variable, or a subject's classification within a classification variable. For example, Amount of Insurance Sold is a quantitative variable that can assume many values. One agent might sell $2,000,000 worth of insurance in one year, another sell $100,000 worth of policies, and another sell nothing ($0). Age is another quantitative variable that assumes a wide variety of values. In the sample shown in Table 1.1, these values ranged from a low of 22 years to a high of 56 years.

Quantitative Variables versus Classification Variables

You can see that, in both amount of insurance sold and age, a given value is a type of score that indicates where the subject stands on the variable of interest. The word "score" is an appropriate substitute for the word "value" in these cases because both are *quantitative variables*. They are variables in which numbers serve as values.

A different type of variable is a *classification variable*, also called a *qualitative variable* or *categorical variable*. With classification variables, different values represent different groups to which the subject belongs. Gender is a good example of a classification variable, as it assumes only one of two values—a subject is classified as either male or female. Race is another example of a classification variable, but it can assume a larger number of values—a subject can be classified as Caucasian American, African American, or Asian American, or as belonging to another group. These variables are classification variables and not quantitative variables because values only represent group membership; they do not represent a characteristic that some subjects possess in greater quantity than others.

Observations

In discussing data, researchers often make references to observational units (or *observations*), which can be defined as the individual subjects (or other objects) that serve as the source of the data. Within the social sciences, a person is usually the observational unit under study (although it is also possible to use some other entity, such as an individual school or organization, as the observational unit). In this text, the person is the observational unit in all examples. Researchers often refer to the number of observations (or cases) included in their data, which simply refers to the number of subjects who were studied. For a more concrete illustration of the concepts discussed so far, consider the data in Table 1.1.

Table 1.1 Insurance Sales Data

Observation	Name	Gender	Age	Goal Difficulty Score	Rank	Sales
1	Bob	M	34	97	2	$598,243
2	Walt	M	56	80	1	$367,342
3	Jane	F	36	67	4	$254,998
4	Susan	F	24	40	3	$80,344
5	Jim	M	22	37	5	$40,172
6	Mack	M	44	24	6	$0

This table reports information about six research subjects: Bob, Walt, Jane, Susan, Jim, and Mack—the data table includes six observations. Information about a given observation (subject) appears as a *row* running from left to right across the table. The first *column* of the data set (running vertically) indicates the observation number, and the second column reports the name of the subject who constitutes or identifies that observation. The remaining five columns report information on the five research variables under study.

- The Gender column reports subject gender, which assumes either "M" for male or "F" for female.
- The Age column reports the subject's age in years.
- The Goal Difficulty Score column reports the subject's score on a fictitious goal difficulty scale. Assume that each participant completed a 20-item questionnaire that assessed the difficulty of the work goals. Depending on how they respond to the questionnaire, subjects receive a score that can range from a low of 0 (meaning that the subject's work goals are quite easy to achieve) to a high of 100 (meaning that they are quite difficult to achieve).
- The Rank column shows how the supervisor ranked the subjects according to their overall effectiveness as agents. A rank of 1 represents the most effective agent, and a rank of 6 represents the least effective.
- The Sales column lists the amount of insurance sold by each agent (in dollars) during the most recent year.

The preceding example illustrates a very small data table with six observations and five research variables (Gender, Age, Goal Difficulty, Rank, and Sales). Gender is a classification variable and the others are quantitative variables. The numbers or letters that appear within a column represent some of the values that these variables can have.

Scales of Measurement and JMP Modeling Types

One of the most important schemes for classifying a variable involves its *scale of measurement*. Researchers generally discuss four scales of measurement: nominal, ordinal, interval, and ratio. In JMP, scales of measurement are designated using three *modeling types*. Modeling types are discussed later, in the section "Modeling Types in JMP."

Before analyzing a data set, it is important to determine each variable's scale of measurement (modeling type) because certain types of statistical procedures require certain scales of measurement. For example, one-way analysis of variance generally requires that the independent variable be a nominal-level variable and the dependent variable be an interval or ratio (continuous) variable. In this text, each chapter that deals with a specific statistical procedure indicates what scale of measurement is required by the variables under study. Then, you must decide whether your variables meet these requirements.

Nominal Scales

A nominal scale is a classification system that places people, objects, or other entities into mutually exclusive categories. A variable measured using a nominal scale is a classification variable that indicates the group to which each subject belongs. The examples of classification variables provided earlier (Gender and Race) also serve as examples of nominal variables. They tell us to which group a subject belongs, but they do not provide any quantitative information about the subjects. That is, the Gender variable might tell us that some subjects are males and other are females, but it does not tell us that some subjects possess more of a specific characteristic relative to others. However, the remaining three scales of measurement provide some quantitative information.

Ordinal Scales

Values on an ordinal scale represent the rank order of the subjects with respect to the variable being assessed. For example, the preceding table includes one variable called Rank that represents the rank ordering of subjects according to their overall effectiveness as agents. The values on this ordinal scale represent a hierarchy of levels with respect to the construct of *effectiveness*. That is, we know that the agent ranked "1" was perceived as being more effective than the agent ranked "2," that the agent ranked "2" was more effective than the one ranked "3," and so forth.

Caution: An ordinal scale has a limitation, in that equal differences in scale values do not necessarily have equal quantitative meaning. For example, look at the following rankings:

Rank	Name
1	Walt
2	Bob
3	Susan
4	Jane
5	Jim
6	Mack

Notice that Walt is ranked "1" while Bob is ranked "2." The rank difference between these two rankings is 1 (2 – 1 = 1), so there is one unit of rank difference between Walt and Bob. Now notice that Jim is ranked "5" while Mack is ranked "6." The rank difference between them is also 1 (6 – 5 = 1), so there is also 1 unit of difference between Jim and Mack. Putting the two together, the rank difference between Walt and Bob is equal to the rank difference between Jim and Mack. However, that does not necessarily mean that the difference in overall effectiveness between Walt and Bob is equal to the difference in overall effectiveness between Jim and Mack. It is possible that Walt is just barely superior to Bob in effectiveness, while Jim is substantially superior to Mack. These rankings reveal very little about the quantitative differences between the subjects with regard to the underlying construct (effectiveness, in this case). An ordinal scale simply provides a rank order of the subjects.

Interval Scales

With an interval scale, equal differences between scale values do have equal quantitative meaning. For this reason, an interval scale provides more quantitative information than the ordinal scale. A good example of an interval scale is the Fahrenheit degree scale used to measure temperature. With the Fahrenheit scale, the difference between 70 degrees and 75 degrees is equal to the difference between 80 degrees and 85 degrees: The units of measurement are equal throughout the full range of the scale.

However, the interval scale also has an important limitation: it does not have a true zero point. A *true zero point* means that a value of zero on the scale represents zero quantity of the construct being assessed. The Fahrenheit scale does not have a true zero point. When a Fahrenheit thermometer reads 0 degrees, that does not mean there is absolutely no heat present in the environment.

Researchers in the social sciences often assume that many of their man-made variables are measured on an interval scale. For example, in the preceding study involving insurance agents, you probably assume that scores from the goal difficulty questionnaire constitute an interval-level scale. That is, you assume that the difference between a score of 50 and 60 is approximately equal to the difference between a score of 70 and 80. Many researchers also assume that scores from an instrument such as an intelligence test are measured at the interval level of measurement.

On the other hand, some researchers are skeptical that instruments such as these have true equal-interval properties, and prefer to call them *quasi-interval* scales. Disagreement about the level of measurement achieved with such instruments continues to be a controversial topic within the social sciences.

However, it is clear that neither of the preceding instruments has a true zero. A score of 0 on the goal difficulty scale does not indicate the complete absence of goal difficulty, and a score of 0 on an intelligence test does not indicate the complete absence of intelligence. A true zero point is found only with variables measured on a ratio scale.

Ratio Scales

Ratio scales are similar to interval scales in that equal differences between scale values have equal quantitative meaning. However, ratio scales also have a true zero point, which gives them an additional property. With ratio scales, it is possible to make meaningful statements about the ratios between scale values. For example, the system of inches used with a common ruler is an example of a ratio scale. There is a true zero point because zero inches does in fact indicate a complete absence of length. With this scale, it is possible to make meaningful statements about ratios. It is appropriate to say that an object four inches long is twice as long as an object two inches long. Age, as measured in years, is also on a ratio scale—a 10-year-old house is twice as old as a 5-year-old house. Notice that it is not possible to make these statements about ratios with the interval-level variables discussed above. You would not say that a person with an IQ of 160 is twice as intelligent as a person with an IQ of 80 because there is no true zero point on the IQ scale.

Although ratio-level scales might be easiest to find in the physical properties of objects, such as height and weight, they are also common in the type of research discussed in this manual. For example, the study discussed previously included the variables for age and amount of insurance sold (in dollars). Both of these have true zero points, and are measured as ratio scales.

Modeling Types in JMP

In JMP, each variable has a modeling type that designates its scale of measurement. In a JMP analysis, the modeling types of the variables convey their scale of measurement. The JMP modeling types are called *nominal*, *ordinal*, and *continuous*. Nominal and ordinal modeling types have the same characteristics as those described above for the same scales of measurement. The continuous modeling type in JMP encompasses the characteristics of both the ratio and interval scales of measurement. Modeling types are used by JMP analysis platforms to help determine the correct analysis that needs to be done.

The discussions that follow refer to JMP modeling types, which are discussed in detail in Chapter 3, "Working with JMP Data."

Basic Approaches to Research

Nonexperimental Research

Much research can be described as being either nonexperimental or experimental in nature. In *nonexperimental* research (also called *nonmanipulative*, *correlational,* or *observational* research), the researcher studies the naturally occurring relationship between two or more naturally occurring variables. A *naturally occurring variable* is a variable that is not manipulated or controlled by the researcher. It is measured as it normally exists.

The insurance study described previously is a good example of nonexperimental research in that you measured two naturally occurring variables (goal difficulty and amount of insurance sold) to determine whether they were related. Another example of nonexperimental research would be an investigation of the relationship between IQ and college grade point average (GPA).

With nonexperimental designs, researchers sometimes refer to response variables and predictor variables.

- A *response variable* is an outcome variable or criterion variable, whose values you want to predict from one or more predictor variables. The response variable is often the main focus of a study because it is mentioned in the statement of the research problem. In the previous example, the response variable is Amount of Insurance Sold. In some experimental research, the response variable is also called the dependent variable.

- A *predictor variable* is the variable used to predict values of the response. In some studies, you might even believe that the predictor variable has a causal effect on the response. In the insurance study, for example, the predictor variable was Goal Difficulty. Because you believed that Goal Difficulty could positively affect insurance sales, you conducted a study in which Goal Difficulty was the predictor and Sales was the response. You do not necessarily have to believe that there is a causal relationship between two variables to conduct a study such as this—you might only be interested in determining whether it is possible to predict one variable from the other. In experimental research, the predictor variable is also known as the independent variable.

Notice that nonexperimental research, which investigates the relationship between just two variables, does not provide evidence concerning cause-and-effect relationships. The reason for this can be seen by reviewing the insurance sales study. If a psychologist conducts this study and finds that the agents with the more difficult goals also tend to sell more insurance, it is not necessarily true that having difficult goals causes them to sell more insurance. Perhaps selling a lot of insurance increases the agents' self-confidence, and this causes them to set higher work goals for themselves. Under this second scenario, it is the insurance sales that had a causal effect on goal difficulty.

As this example shows, with nonexperimental research it is often possible to obtain a single result that is consistent with a number of contradictory causal explanations. Hence, a strong inference that variable A had a causal effect on variable B is rarely if ever valid when you conduct simple correlational research with just two variables. To obtain stronger evidence of cause and effect, researchers either analyze the relationships between a larger number of variables using sophisticated statistical procedures that are beyond the scope of this text, or drop the nonexperimental approach entirely and use experimental research methods instead. The nature of experimental research is discussed in the following section.

Experimental Research

Most experimental research can be identified by three important characteristics:

- Subjects are randomly assigned to experimental conditions.
- The researcher manipulates an independent predictor variable.
- Subjects in different experimental conditions are treated similarly with regard to all variables except the independent variable.

To illustrate these concepts, assume that you conduct an experiment to test the hypothesis that goal difficulty positively affects insurance sales. Assume that you identify a group of 100 agents to serve as subjects. You randomly assign 50 agents to a "difficult goal" condition. Subjects in this group are told by their superiors to make at least 25 cold calls (unexpected sales calls) to potential policyholders per week. The other 50 agents have been randomly assigned to the "easy goal" condition. They have been told to make just 5 cold calls to potential policyholders per week. The design and results of this experiment are illustrated in Table 1.2.

After one year, you determine how much new insurance each agent sold that year. Assume that the average agent in the difficult goal condition sold $156,000 of new policies, while the average agent in the easy goal condition sold just $121,000 worth.

Table 1.2 Design of the Experiment Used to Assess the Effects of Goal Difficulty

Group	Treatment Conditions under the Independent Variable (Goal Difficulty)	Results Obtained with the Dependent Variable (Amount of Insurance Sold)
Group 1 (N = 50)	Difficult Goal Condition	$156,000 in Sales
Group 2 (N = 50)	Easy Goal Condition	$121,000 in Sales

It is possible to use some of the terminology associated with nonexperimental research when discussing this experiment. For example, it is appropriate to continue to refer to Amount of Insurance Sold as being a response variable, because this is the outcome variable of interest. You can also refer to Goal Difficulty as the predictor variable because you believe that this variable, to some extent, predicts the amount of insurance sold.

Notice, however, that Goal Difficulty is now a different kind of variable. In the nonexperimental study, Goal Difficulty was a naturally occurring variable that could take on a wide variety of values (whatever score the subject received on the goal difficulty questionnaire). In the present experiment Goal Difficulty is a *manipulated variable*, which means that you (as the researcher) determine what value of the variable is to be assigned to each subject. In this experiment, Goal Difficulty assumes only one of two values—subjects are in either the difficult goal group or the easy goal group. Therefore, Goal Difficulty is now a classification variable, with a nominal modeling type.

Although it is acceptable to speak of predictor and response variables within the context of experimental research, it is more common to speak in terms of independent variables and dependent variables.

- An *independent variable* is that variable whose values (or levels) the experimenter selects to determine what effect this independent variable has on the dependent variable. The independent variable is the experimental counterpart to a predictor variable.
- A *dependent variable* is some aspect of the subject's behavior assessed to reflect the effects of the independent variable. The dependent variable is the experimental counterpart to a response variable.

In the example shown in Table 1.2, Goal Difficulty is the independent variable and Sales is the dependent variable.

Remember that the terms predictor variable and response variable can be used with almost any type of research, but that the terms independent and dependent variable should be used only with experimental research.

Researchers often refer to the different *levels* of the independent variable. These levels are also referred to as *experimental conditions* or *treatment conditions* and correspond to the different groups to which a subject can be assigned. The present example includes two experimental conditions, a "difficult goal condition" and an "easy goal condition."

With respect to the independent variable, you can speak in terms of the experimental group versus the control group. Generally speaking, the *experimental group* receives the experimental treatment of interest, while the *control group* is an equivalent group of subjects that does not receive this treatment. The simplest type of experiment consists of one experimental group and one control group. For example, the present study could have been redesigned so that it consisted of an experimental group that was assigned the goal of making 25 cold calls (the difficult goal condition) and a control group in which no goals were assigned (the no-goal condition).

You can expand the study by creating more than one experimental group. You could do this in the present case by assigning one experimental group the difficult goal of 25 cold calls and the second experimental group the easy goal of just 5 cold calls, and include a third group as the control group assigned no goals.

Descriptive versus Inferential Statistical Analysis

To understand the difference between descriptive and inferential statistics, you must first understand the difference between populations and samples.

- A *population* is the entire collection of a carefully defined set of people, objects, or events. For example, if the insurance company in question employed 10,000 insurance agents in the U.S., then those 10,000 agents would constitute the population of agents hired by that company.
- A *sample* is a subset of the people, objects, or events selected from that population. For example, the 100 agents used in the experiment described earlier constitute a sample.

Descriptive Analyses: What Is a Parameter?

A *parameter* is a descriptive characteristic of a population. For example, if you found the average amount of insurance sold by all 10,000 agents in this company (the population of agents in this company), the resulting average (also called the *mean*) would be a population parameter. To obtain this average, you first need to tabulate the amount of insurance sold by each and every agent. When calculating this mean, you are engaging in descriptive statistical analysis. *Descriptive statistical analysis* focuses on the exhaustive measurement of population characteristics. You define a population, assess each member of that population, and compute a summary value (such as a mean or standard deviation) based on those values.

Most people think of populations as being very large groups, such as all of the people in the U.S. However, a group does not have to be large to be a population; it only has to be the entire collection of the people or things being studied. For example, a teacher might define as a population all 23 students taking an English course, and then calculate the average score of these students on a measure of class satisfaction. The resulting average is a parameter.

Inferential Analyses: What Is a Statistic?

A *statistic* is a numerical value that is computed from a sample, describes some characteristic of that sample such as the mean, and can be used to make inferences about the population from which the sample is drawn. For example, if you were to compute the average amount of insurance sold by your sample of 100 agents, that average would be a statistic because it summarizes a specific characteristic of the sample. Remember that the

word "statistic" is generally associated with samples, while "parameter" is generally associated with populations.

In contrast to descriptive statistics, *inferential* statistical analysis involves using information from a sample to make inferences, or estimates, about the population. For example, assume that you need to know how much insurance is sold by the average agent in the company. Suppose it is impossible (or very difficult) to obtain the necessary information from all 10,000 agents and then calculate a mean. An alternative is to draw a random (and ideally representative) sample of 100 agents and determine the average amount sold by this subset. If this group of 100 sold an average of $179,322 worth of policies last year, then your best guess of the amount of insurance sold by all 10,000 agents would be $179,322. You have used characteristics of the sample to make inferences about characteristics of the population. Using some simple statistical procedures, you can even compute confidence intervals around the estimate, which allows you to make statements such as

> "There is a 95% chance that the actual population mean lies somewhere between $172,994 and $185,650."

This is the real value of inferential statistical procedures—they allow you to review information obtained from a relatively small sample and then to make inferences about a population.

Hypothesis Testing

Most of the procedures described in this manual are inferential procedures that let you test specific hypotheses about the characteristics of populations. As an illustration, consider the simple experiment, described earlier, in which 50 agents are assigned to a difficult goal condition and 50 other agents to an easy goal condition. Assume that, after one year, the difficult-goal agents sold an average of $156,000 worth of insurance, while the easy-goal agents sold only $121,000 worth. On the surface, this seems to support your hypothesis that agents sell more insurance when they have difficult goals. But, even if goal setting had no effect at all, you don't really expect the two groups of 50 agents to sell exactly the same amount of insurance. You expect one group to sell somewhat more than the other due to chance alone. The difficult-goal group did sell more insurance, but did it sell *enough* more to make you confident that the difference was due to placing the agents in different goal groups?

What's more, you can argue that you don't even care about the amount of insurance sold by these two small samples. What really matters is the amount of insurance sold by the larger populations they represent. Define the first population as "the population of agents

assigned difficult goals" and the second as "the population of agents assigned easy goals." Your real research question is whether the first population sells more than the second. To address this question, you need hypothesis testing.

Types of Inferential Tests

Generally speaking, there are two types of tests conducted when using inferential procedures:

- tests of group differences
- tests of association.

With a *test of group differences*, you want to know whether two populations differ with respect to their mean scores on some response variable. The present experiment leads to a test of group differences because you want to know whether the average amount of insurance sold in the population of difficult-goal agents is different from the average amount sold in the population of easy-goal agents. A different example of testing group differences might involve a study in which the researcher wants to know whether Caucasian Americans, African Americans, and Asian Americans differ with respect to their mean scores on a locus of control scale. (Locus of control refers to the extent to which people believe that their own actions determine the rewards they obtain.) Notice that in both cases, two or more distinct populations are being compared with respect to their mean scores on a single response variable.

With a *test of association*, there is a single population of individuals and you want to know whether there is a relationship between two or more variables within this population. Perhaps the best-known test of association involves testing the significance of a correlation coefficient. Assume that you conduct a simple correlational study in which you ask 100 agents to complete the 20-item goal difficulty questionnaire. Remember that, with this questionnaire, subjects can receive a score that ranges from a low of 0 to a high of 100. You can then correlate these goal difficulty scores with the amount of insurance sold by the agents that year. Here, the goal difficulty scores constitute the predictor variable and the amount of insurance sold serves as the response. Obtaining a strong positive correlation between these two variables means that the more difficult the agents' goals, the more insurance they tend to sell. This is called a test of association because you determine whether there is an association, or *relationship*, between the predictor and response variables. Notice also that there is only one population being studied—there is no experimental manipulation that creates a difficult-goal population versus an easy-goal population.

For the sake of completeness, it is worth mentioning that there are some relatively sophisticated procedures that also let you test whether the association between two variables is the same across two or more populations. Analysis of covariance (ANCOVA) is one procedure that allows such a test.

For example, you could form a hypothesis that the association between self-reported goal difficulty and insurance sales is stronger in the population of agents assigned difficult goals than it is in the population assigned easy goals. To test this hypothesis, you randomly assign a group of insurance agents to either an easy-goal condition or a difficult-goal condition (as described earlier). Each agent completes the 20-item self-report goal difficulty scale and is then given the group assignment (treatment) to make more or fewer cold calls. Subsequently, you could record each agent's sales. Analysis of covariance allows you to determine whether the relationship between questionnaire scores and sales is stronger in the difficult-goal population than it is in the easy-goal population.

ANCOVA also allows you to test a number of additional hypotheses and is beyond the scope of this text. For more information about ANCOVA in JMP, see the *JMP Statistics and Graphics Guide* (2003).

Types of Hypotheses

Two different types of hypotheses are relevant to most statistical tests. The first is called the *null hypothesis*, which is often abbreviated as H_0. The null hypothesis is a statement that, in the population(s) being studied, there are either (a) no differences between the group means, or (b) no relationships between the measured variables. For a given statistical test, either (a) or (b) applies, depending on whether the test is to detect group differences or is a test of association.

With a test of group differences, the null hypothesis states that, in the population, there are no differences between any of the groups being studied with respect to their mean scores on the response variable. In the experiment in which a difficult-goal condition is being compared to an easy-goal condition, the following null hypothesis might be used:

> H_0: In the population, the amount of insurance sold by individuals assigned difficult goals does not differ from the amount of insurance sold by individuals assigned easy goals.

This null hypothesis can also be expressed with symbols in the following way:

H_0: $M_1 = M_2$

where

H_0 represents the null hypothesis
M_1 represents mean sales for the difficult-goal population
M_2 represents mean sales for the easy-goal population.

In contrast to the null hypothesis, there is also an *alternative hypothesis* (H_1) that states the opposite of the null. The alternative hypothesis is a statement that there is a difference between the means, or that there is a relationship between the variables, in the population(s) being studied.

Perhaps the most common alternative hypothesis is a *nondirectional alternative hypothesis*. With a test of group differences, a nondirectional alternative hypothesis predicts that the means for the various populations differ, but makes no specific prediction as to which mean will be relatively high and which will be relatively low. In the preceding experiment, the following nondirectional null hypothesis might be used:

H_1: In the population, individuals assigned difficult goals differ from individuals assigned easy goals with respect to the mean amount of insurance sold.

This alternative hypothesis can also be expressed with symbols in the following way:

$H_1: M_1 \neq M_2$

In contrast, a *directional alternative hypothesis* makes a more specific prediction regarding the expected outcome of the analysis. With a test of group differences, a directional alternative hypothesis not only predicts that the population means differ, but also predicts which population means will be relatively high and which will be relatively low.

Here is a directional alternative hypothesis for the preceding experiment.

H_1: The average amount of insurance sold is higher in the population of individuals assigned difficult goals than in the population of individuals assigned easy goals.

This hypothesis can be symbolically represented as follows:

$H_1: M_1 > M_2$

If you believe that the easy-goal population sells more insurance, you replace the "greater than" symbol (>) with the "less than" symbol (<) in the alternative hypothesis, as follows:

H_1: $M_1 < M_2$

Null and alternative hypotheses are also used with tests of association. For the study in which you correlated goal-difficulty questionnaire scores with the amount of insurance sold, you might use the following null hypothesis:

H_0: In the population, the correlation between goal difficulty scores and the amount of insurance sold is zero.

You could state a nondirectional alternative hypothesis that corresponds to this null hypothesis as follows:

H_1: In the population, the correlation between goal difficulty scores and the amount of insurance sold is not equal to zero.

Notice that the preceding is an example of a nondirectional alternative hypothesis because it does not specifically predict whether the correlation is positive or negative, only that it is not zero. A directional alternative hypothesis, on the other hand, might predict a positive correlation between the two variables. You could state such a prediction as follows:

H_1: In the population, the correlation between goal difficulty scores and the amount of insurance sold is greater than zero.

There is an important advantage associated with the use of directional alternative hypotheses compared to nondirectional hypotheses. Directional hypotheses allow researchers to perform *one-sided* statistical tests (also called one-tailed tests), which are relatively powerful. Here, "powerful" means the ability of a test to detect significant differences between group means when differences really do exist. In contrast, non-directional hypotheses allow only *two-sided* statistical tests (also called two-tailed tests), which are less powerful.

Because they lead to more powerful tests, directional hypotheses are generally preferred over nondirectional hypotheses. However, directional hypotheses should be stated only when they can be justified on the basis of theory, prior research, or some other grounds. For example, you should state the directional hypothesis that

"The average amount of insurance sold is higher in the population of individuals assigned difficult goals than in the population of individuals assigned easy goals,"

only if there are theoretical or empirical reasons to believe that the difficult-goal group will indeed score higher on insurance sales. The same should be true when you specifically predict a positive correlation rather than a negative correlation (or vice versa).

The *p* Value

Hypothesis testing is a process to determine whether you can reject a null hypothesis with an acceptable level of confidence. When analyzing data with JMP, you look at the results for two pieces of information that are critical for this purpose:

1. the obtained (calculated) statistic
2. the probability (*p*) value associated with that statistic.

Consider the experiment in which you compared the difficult-goal group to the easy-goal group. One way to test the null hypothesis associated with this study is to perform an independent samples *t* test. When the data analysis for this study is complete, you compute a *t* statistic and its corresponding *p* value. The *p* value indicates the probability that you would obtain the present results if the null hypothesis were true. If the *p* value is very small, you reject the null hypothesis. Recall that the null hypothesis states that there is no difference between groups.

For example, assume that you obtain a *t* statistic of 0.14 and a corresponding *p* value of 0.90. This *p* value means that there are 90 chances in 100 that you would obtain a *t* statistic of 0.14 (or larger) if the null hypothesis were true. Because this probability is so high, you report that there is very little evidence to refute the null hypothesis. In other words, you fail to reject the null hypothesis, and, instead, conclude that there is not sufficient evidence to support a statistically significant difference between the two groups.

On the other hand, assume that the research project instead produces a *t* value of 8.45 and a corresponding *p* value of 0.001. The *p* value of 0.001 means that there is only one chance in 1000 that you would obtain a *t* value of 8.45 (or larger) if the null hypothesis were true. This is so unlikely that you are fairly confident that the null hypothesis is *not* true. You therefore reject the null hypothesis and conclude that there is a difference in mean sales between the two populations. In rejecting the null hypothesis, you have tentatively accepted the alternative hypothesis.

Technically, the *p* value does not really provide the probability that the null hypothesis is true. Instead, it provides the probability that you would obtain the present results (the

present *t* statistic, in this case) if the null hypothesis were true. This might seem like a trivial difference, but it is important to know the meaning of the *p* value.

Notice that you are able to reject the null hypothesis only when the *p* value is a small number (0.001, in the above example). But how small must a *p* value be before you can reject the null hypothesis? A *p* value of 0.05 is one of the most commonly accepted cutoff values. Typically, when researchers obtain a *p* value *larger* than 0.05 (such as 0.13 or 0.37), they fail to reject the null hypothesis, and instead conclude that the differences or relationships being studied are not statistically significant (there is no significant difference between groups). When researchers obtain a *p* value *smaller* than 0.05, they reject the null and conclude that differences or relationships being studied are statistically significant (there is a significant difference between groups). The 0.05 level of significance is not an absolute rule that must be followed in all cases, but it is serviceable for most types of investigations likely to be conducted in the social sciences and other areas.

Fixed Effects versus Random Effects

Experimental designs can be represented by mathematical models described as fixed-effects models, random-effects models, or mixed-effects models. The use of these terms refers to the way that the levels of the independent variable (or predictor variable) were selected.

When the researcher arbitrarily selects the levels of the independent variable, the independent variable is called a *fixed-effects factor,* and the resulting model is a *fixed-effects model.* For example, assume that in the current study you arbitrarily decided that the subjects in your easy-goal condition would be told to make just 5 cold calls per week, and that the subjects in the difficult-goal condition would be told to make 25 cold calls per week. In this case, you have *fixed* (arbitrarily selected) the levels of the independent variable. Your experiment therefore represents a fixed-effects model.

In contrast, when the researcher randomly selects levels of the independent variable from a population of possible levels, the independent variable is called a *random-effects factor*, and the model is a *random-effects model*. For example, assume you know that the number of cold calls an insurance agent could possibly place in one week ranges from 0 to 45. This range represents the population of cold calls that you could possibly research. Assume you use some random procedure to select two values from this population of possible calls, and that those two randomly selected values are 12 and 32. In conducting your study, one group of subjects is assigned to make at least 12 cold calls per week, while the second is assigned to make 32 calls. In this case, your study represents a

random-effects model because the levels of the independent variable were randomly selected from all possible levels.

As an illustration of a fixed-effects model, assume that you want to conduct research on the effectiveness of hypnosis in reducing anxiety among subjects who suffer from phobias. Specifically, you want to perform an experiment that compares the effectiveness of 10 sessions of relaxation training versus 10 sessions of relaxation training plus hypnosis. In this study, the independent variable might be labeled "Type of Therapy." Notice that you did not randomly select these two treatment conditions from the population of all possible treatment conditions because you know which treatments you want to compare, and design the study accordingly. This is experimental research and your study represents a fixed-effects model.

To provide a nonexperimental example, assume that you want to conduct a study to determine whether Caucasian Americans score significantly higher than African Americans on internal locus of control. The predictor variable in your study is race, and the response variable is scores on some index of locus of control. Most likely, you chose "Caucasian American" versus "African American" as predictor variable groups because you are particularly interested in these two races. You did not randomly select these groups from all possible races. Therefore, the study is another example of a fixed-effects model.

Random-effects factors do sometimes appear in research. For example, in a repeated-measures investigation more than one measure of the response variable is taken from each subject. Subjects are viewed as a random-effects factor, assuming they were randomly selected. Some studies include both fixed-effects factors and random-effects factors. Those models are called *mixed-effects models*.

This distinction between fixed and random effects has important implications for the types of inferences that can be drawn from statistical tests. When analyzing a fixed-effects model, you can generalize the results of the analysis only to the specific levels of the independent variable manipulated in that study. This means that if you arbitrarily selected 5 cold calls versus 25 cold calls for your two treatment conditions, once the data are analyzed you can draw conclusions only about the population of agents assigned 5 cold calls versus the population assigned 25 cold calls.

On the other hand, if you randomly selected two values for your treatment conditions (say, 12 versus 32 cold calls) from the population of possible numbers of calls, the model is a random-effects model. This means that you can draw conclusions about the entire population of possible values that the independent variable can assume. Inferences are not restricted to just the two treatment conditions investigated in the study. In other

words, you can draw inferences about the relationship between the population of the possible number of cold calls that agents could be assigned and the response variable (insurance sales).

Summary

Regardless of discipline, researchers need a common language to use when discussing their work with others. This chapter has reviewed the basic concepts and terminology of research that will be referred to throughout this text. Now that you can speak the language, you are ready to move on to Chapter 2, "Getting Started with JMP," where you learn how to do simple JMP analyses.

References

SAS Institute Inc. 2003. *JMP Statistics and Graphics Guide.* Cary, NC: SAS Institute Inc.