

1.2: Importance of Statistics

Learning Objectives

- Motivate the importance of statistical literacy in our daily lives
- Review the scientific method
- Define variable
- Define independent variable
- Define dependent variable
- Outline the basic process of statistics-based research

General Overview of Why/How We Study Statistics

Most of us are probably aware of the human desire to live a good life. In trying to achieve this goal, we must make choices based upon the world around us. To help us make these decisions, we can examine things that are consistent and predictable and use them to anticipate future events. This, paired with the knowledge that not everything in life can be classified as consistent or predictable, can help us develop tools that [facilitate](#) our growth.

What are some examples of the challenges we find in life that draw us to develop some understanding of these tools and how the world works? Perhaps a candidate for mayor claims that the rising crime rate in the city is a direct consequence of the policies his political opponent supports; how can we determine if this is correct? How do we decide who we vote for in the upcoming election? Perhaps a salesperson is trying to sell us blackout curtains, claiming that it will lower the heating bill of the house. Will we actually save money by purchasing the curtains? Will a new diet fix our health issues? Can we improve the fuel economy of our vehicles by using a different kind of fuel? Is the probability of severe weather high enough to justify canceling some event? There are many situations that require us to observe and analyze in order to make the best choice.

Having experienced the unpredictability and inconsistency found in both the natural world and our human societies, a legitimate question arises: how do we know that our understanding of the world actually represents reality? The world is complicated. Numerous factors, both seen and unseen, play into every event. People may make claims out of ignorance, or they might have ulterior, even malicious, intentions. Knowing this, how should we act? We could very easily become overwhelmed with doubts, but luckily, throughout most of our lives, we have been developing ways to address them.

The answer is quite simple and familiar. We observe people and events repeatedly throughout our lives, noting different circumstances and outcomes. We analyze our observations to form an initial conclusion. From our initial perceptions, we refine and build trust in our models by repeatedly testing and continually updating them. Eventually, our perceptions of certain individuals and events [garner](#) enough trust and consistency that we naturally rely on them. However, we are always open to additional information that may cast new light upon our previous perceptions. In most of us, this process happens naturally. Hopefully, we recognize this process as the foundation of the scientific method.

We can understand the scientific method as the result of recognizing and [honing](#) the natural process of inquiry outlined above. As we grow in our ability to analyze the world, common errors and methodological inefficiencies are identified and [expunged](#). The scientific method begins with a set of observations that [elicit](#) some interest which then fosters the development of a research question on a particular topic. At this stage, an initial generalization or hypothesis is constructed to provide insight into an answer to the research question; the hypothesis speaks to the relationship between specific aspects of the field of interest. These specific aspects are **variables** (properties or characteristics of some event, object, or person that can take on different values or amounts). The hypothesis must be falsifiable; that is, it could be shown to be incorrect.

Once the hypothesis has been constructed, the hypothesis is tested through experimentation. When we merely observe, we cannot account for the individual influence of each variable at play. The experimental design process is one of the most important steps in the scientific method. Here, the researcher identifies all of the other variables, called **confounding variables**, that may affect the hypothesized relationship. The researcher may devise a plan to negate or control the influence of those confounding variables while systematically changing some of the variables of interest. The variables that are changed systematically by a researcher during an experiment are called **independent variables**. The **dependent variables** are the variables measured as the independent variables are manipulated. The experimental design becomes significantly more complicated as the number of independent and dependent variables increases since knowledge of how each independent variable interacts with each dependent variable would need to be

examined. For this reason, it is preferable to keep the number of independent and dependent variables to a minimum within a particular experiment. More variables can be examined in subsequent experiments.

? Text Exercise 1.2.1

High schools often prepare students for graduation by exploring career options and the various paths into those fields. A common component of such presentations is [job satisfaction](#). In which fields are people most satisfied with their jobs? In which fields are people happiest? A young student might look at a list of careers with high satisfaction, which includes clergy, chiropractors, firefighters, nurses, and dentists (to name a few), and think that picking a career from such a list will result in living a good life. We can understand the collection of job satisfaction data as a form of experimentation. Identify independent, dependent, and confounding variable(s) and assess the connection between profession and happiness.

Answer

In job satisfaction studies, the primary variables of interest are profession and satisfaction. The researchers study particular careers, which are the "values" of the variable (profession), and then measure the dependent variable (satisfaction) as the particular careers change. This makes profession the independent variable and satisfaction the dependent variable. Many factors, such as personal values, interests, and strengths, play a major role in job satisfaction. The degrees of repetition and mindlessness in a job also play a role. These variables, and many other variables left unstated, are confounding variables. Satisfaction in career and life corresponds most directly with a person's individual values, interests, and strengths. We are unique and our vocation, our call in life, will match who we are.

The experimental design process also focuses on data collection and analysis. The researcher must determine how the independent variables will be altered consistently, how the dependent variables will be measured reliably, what analyses are appropriate for the collected data, and how these analyses test the hypothesis. Recall the [definition of statistics](#) provided earlier. Fluency with statistics facilitates this process and helps ensure that our conclusions will be meaningful, not necessarily desired, but meaningful.

Once the experimental design is finished, the experiment will be conducted, the collected data will be summarized and analyzed, and a conclusion will be made regarding the hypothesis. If the initial hypothesis was found to be false, a new hypothesis could be formed incorporating the newest findings. Alternatively, the data could align with the hypothesis, which only increases our confidence in its [veracity](#). The scientific method encourages the sharing of experimental design and conclusions. Significant and immediate confidence can be attained in rejecting hypotheses, while confidence in the truth of hypotheses comes from repeatedly conducting experiments that support the hypotheses.

Whether we explicitly engage in the scientific method or just try to make good decisions, we routinely engage in the process of observation, generalization, testing, and updating. This process requires the collection and analysis of data with the goal of drawing further conclusions and is statistical in nature. Therefore, it [behooves](#) us to take our study of statistics seriously and to utilize it in our daily lives. It will benefit us twofold: in developing our own understanding of the world and in intelligently considering the many claims of others.

Consequently, we should understand the basic process involved in statistic-based research, whether we involve ourselves informally (in a basic inquiry of our day-to-day lives) or formally (in some important inquiry that may have an impact well beyond ourselves). The process of statistic-based research can be broadly discussed in four steps:

1. Establish a research question that is to be explored.
2. Determine appropriate subjects and needed variables to guide in producing data to address the research question. Collect such data using appropriate methods.
3. Summarize the collected data using appropriate statistical processes. Use inferential methods when needed.
4. Carefully use the summarized data to make sound, reasoned conclusion(s). The conclusion(s) should be further analyzed for practical significance, not merely statistical significance.

Although we will focus mainly on steps 3 and 4 of this process in this course, the first two steps are equally important. Establishing a quality research question, as well as determining what data is needed and how to collect that data, can be more challenging to do than the last two steps. As both producers and consumers of statistic-based research, we must be familiar with this process to understand its power and the limitations of such research.

Statistics are often presented in an effort to add credibility to an argument or advice, as can be seen in the numerous advertisements viewed daily. Many of the numbers thrown around do not represent careful statistical analysis. They can be misleading and push us into decisions that we might regret. To be an intelligent consumer of statistics, our first reflex must be to question the statistics that we encounter. We must think about the claims, the numbers, their sources, and most importantly, the procedures used to generate them.

1.2: Importance of Statistics is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by The Math Department at Fort Hays State University.

- 1.2: Importance of Statistics by [David Lane](#) is licensed [Public Domain](#). Original source: <https://onlinestatbook.com>.