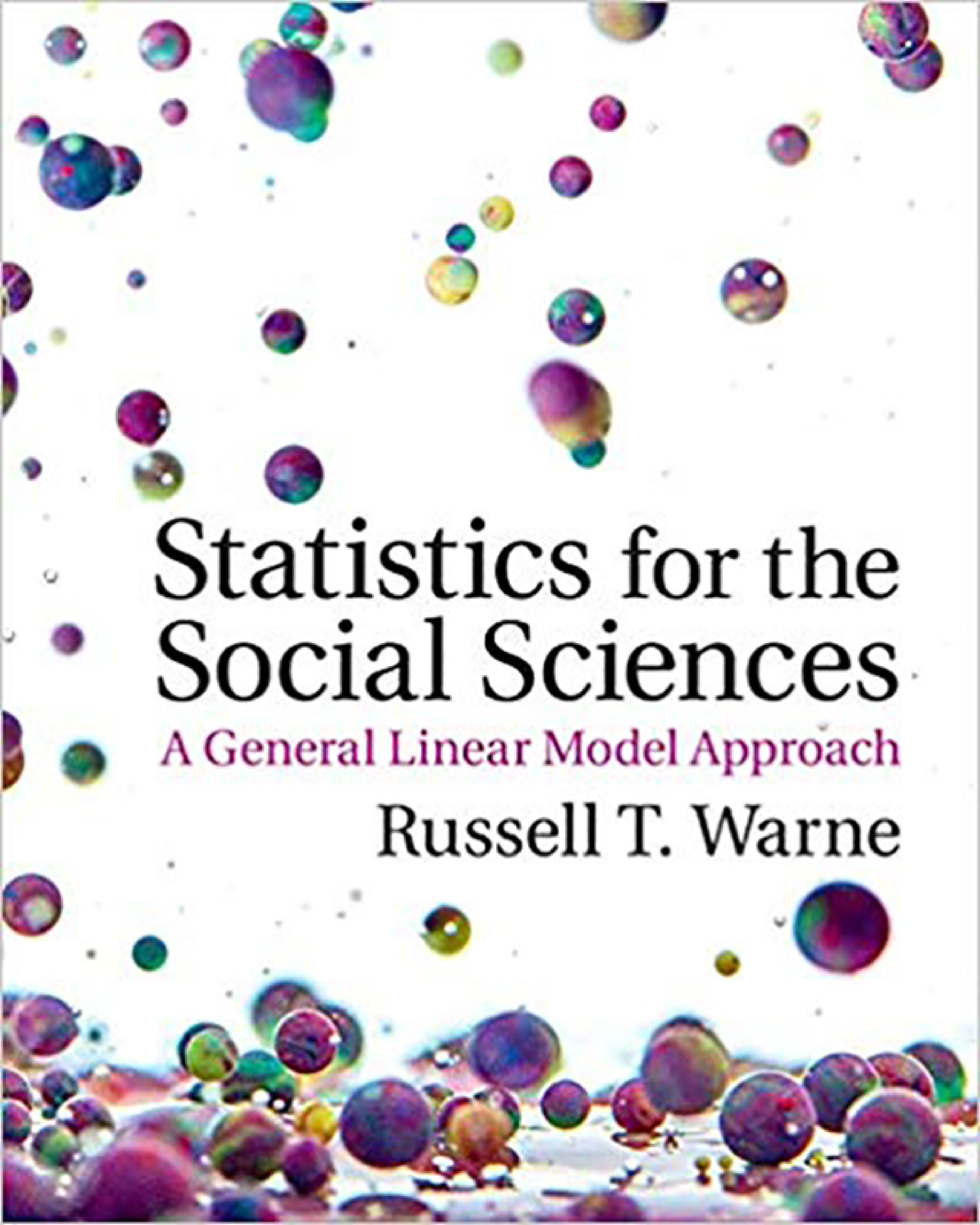

Statistics for the Social Sciences

A General Linear Model Approach

Russell T. Warne

Statistics for the Social Sciences

A General Linear Model Approach

Written by a quantitative psychologist, this textbook explains complex statistics in accessible language to undergraduates in all branches of the social sciences. Built around the central framework of the general linear model (GLM), *Statistics for the Social Sciences* teaches students how different statistical methods are interrelated to one another. With the GLM as a basis, students with varying levels of background are better equipped to interpret statistics and learn more advanced methods in their later courses. Dr. Warne makes statistics relevant to students' varying majors by using fascinating real-life examples from the social sciences. Students who use this book will benefit from clear explanations, warnings against common erroneous beliefs about statistics, and the latest developments in the philosophy, reporting, and practice of statistics in the social sciences. The textbook is packed with helpful pedagogical features including learning goals, guided practice, and reflection questions.

Dr. Russell T. Warne is a quantitative psychologist in the Department of Behavioral Science at Utah Valley University who earned his PhD from Texas A&M University in 2011. Since 2009 he has taught introductory statistics to undergraduates majoring in psychology, sociology, education, anthropology, communications, family science, exercise science, and biology. At Utah Valley University his statistics course is one of his department's highest rated classes, with an average student rating of 4.7 out of 5.

Dr. Warne doesn't just teach statistics. He uses statistics in his research, too. He has published over 40 articles in professional journals in psychology, education, methodology, medicine, sociology, health, business, and the arts. For many of these articles his coauthors were students. Dr. Warne has earned awards for his research from the National Association for Gifted Children, the Southwest Educational Research Association, Texas A&M University, and Utah Valley University. He also uses his statistical training to serve on the editorial board of three scientific journals – *Journal of Psychoeducational Assessment*, *Gifted Child Quarterly*, and *Journal of School Psychology.*

Statistics for the Social Sciences

A General Linear Model Approach

RUSSELL T. WARNE

Utah Valley University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107576971

© Russell Warne 2018

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2018

Printed in the United States of America by Sheridan Books, Inc.

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Warne, Russell T., 1983- author.
Title: Statistics for the social sciences : a general linear model approach / Russell T. Warne.
Description: New York : Cambridge University Press, 2018. | Includes bibliographical references and index.
Identifiers: LCCN 2017023186 | ISBN 9781107576971 (paperback)
Subjects: LCSH: Statistics–Textbooks. | Social sciences–Statistical methods–Textbooks.
Classification: LCC QA276 .W367 2017 | DDC 519.5–dc23 LC record available at https://lccn.loc.gov/2017023186

ISBN 978-1-107-57697-1 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

For Katie.

The likelihood of finding a woman as wonderful as her is $P = .000001$.

Contents

Preface

If you had told me 15 years ago that I would write a statistics textbook, I would have laughed. Sitting in that first statistics course a decade and a half ago as an undergraduate psychology major, I struggled with the concepts that my professor was teaching. I didn't understand why statistical methods were relevant to my major or my career goals. The terminology was difficult, and the explanations that I received were confusing. I tried my hardest, and I passed the course – but not with an impressive grade. My goal in writing this textbook is to help students in the social sciences avoid the unpleasant experience I had in my first statistics course. I've had 15 years to think about what went wrong in the course, and I have designed this textbook to help students have the best possible introduction to statistics.

Decades of research in educational and cognitive psychology have shown that students learn more material when they have a way to organize what they are expected to learn. I have therefore given students and instructors a way to think about statistical methods called the general linear model (GLM). The GLM underlies most of the statistical analyses that are used in applied and social science research, and having this general framework of understanding statistics will avoid the tendency of students to see statistics as a disconnected set of procedures. My hope is that the GLM can help students understand statistical methods so that their commonalities and differences make sense.

Another common finding in the research on learning is that students retain more knowledge when the material is relevant to them. Whenever possible, I have used examples of real data from across the social sciences to illustrate statistical concepts. Empirical research suggests that using real data to teach statistical concepts increases student learning (e.g., Allen & Baughman, 2016). I also believe that using real data shows students how the equations and theory of statistics have been used to create real knowledge in their fields. Additionally, the examples from the actual work in the social sciences are far more likely to be interesting and relevant to students than the made-up examples that my professor gave me when I was an undergraduate.

My final goal with this textbook is to reduce anxiety for students. When I took my first statistics course (and the second and the third), I was overwhelmed. Statistics was intimidating, and my anxiety made it more difficult to learn what I was supposed to. To avoid this, I have made several efforts to make statistics accessible and manageable for students. First, I have made explanations simple and straightforward. Technical terminology is defined clearly, and a glossary in the back of the book provides easy reference for students. Second, because formulas can be intimidating for some students, every formula is labeled, and the various symbols are explained. I also have shown why the formulas are set up in the way that they are and how they are linked with their interpretation. Finally, I have included detailed examples of statistical problems that are worked out step-by-step. By following these examples, students will better understand the steps and interpretation of statistical analyses, and solve problems and complete homework assignments more efficiently.

What Makes this Textbook Different

- **A foundation in the research on student learning**. Although I consider myself a quantitative psychologist, my doctoral program was housed in a college of education, and my PhD is in educational psychology. That means that I am familiar with the research on how people learn, and I have applied this research to the structure and writing of every chapter. Although I rarely make my debt to cognitive and educational psychology explicit, readers can be assured that every decision behind this book was made from an evidence-based perspective.

- **Teaching statistical concepts the way that experts think about them**. Statistical novices see statistics as an unrelated hodgepodge of analysis methods. On the other hand, statistical experts see statistics as all being united by general concepts, such as their correlational nature or the universality of effect sizes. Yet, no other author of a textbook for social science undergraduates teaches about these general concepts – most of which are united in the GLM. I think that students will master and retain more material by learning about the GLM in their first statistics course. Moreover, learning the GLM makes future statistical procedures (e.g., nonparametric statistics, multivariate statistics) easier to learn because they are also members of the GLM – albeit more complex ones.
- **Practical application**. The use of examples from real studies in the social sciences was done with an eye on showing students how statistics can lead to practical knowledge in the social sciences. Browsing through the book will show that I have included examples from psychology, sociology, family studies, anthropology, education, social work, and more. This is the result of my effort to make statistics relevant to students from a broad cross-section of majors.
- **Discussion of controversies**. For the sake of simplicity, most authors of undergraduate textbooks try to avoid controversies in statistics. I can sympathize with this viewpoint, but I think that it is not beneficial to students. Social science statistics is a field that has been punctuated by periodic controversies for the last 120 years. I believe that avoiding these controversies encourages simplistic thinking and a naive view of statistics. Instead, I feel that undergraduate students can be more nuanced thinkers and more professional in their interpretations if they are aware of some of the disagreements and arguments among professionals. Therefore, I have included discussions of the shortcomings of null hypothesis testing, the interpretation of effect sizes, questionable research practices, and more. I believe that this enlivens the textbook and prepares students for joining the social science community.

For Students

If you are a student using this textbook, there are a few things you can do to improve your performance in your statistics course.

First, make this course a priority. Devote the time and attention necessary to do well in your statistics course. Although I have tried to make this textbook accessible, some concepts are simply counterintuitive or strange, and very few people can grasp the logic and mathematics of statistics without effort. Read the text carefully and with full attention. Cognitive psychologists have shown that human beings can only effectively focus attention on one thing at a time, and it is going to be difficult to learn statistics with any distractions, including the TV, your phone, or social media. Make sure you can devote all of your attention to what you are reading.

Second, study carefully. As you read, highlight important sentences. Take notes in the margins, and jot down questions that come to your mind as you read. Coming to class prepared with questions and notes will make you more able to excel in the course. I also recommend that you practice the concepts you read about by working out the questions at the end of each chapter – even if they are not assigned as homework. My students have found these questions to be particularly useful in cementing their understanding of statistical concepts.

Third, participate in class. Ask those questions you prepared in advance. My experience in teaching statistics is that when a student asks a question, there are almost always classmates who have the same question. Also, don't hesitate to ask your instructor to repeat a concept or to restate it in other words. You won't look stupid – it's just part of learning statistics. When I learn a new statistical concept, I have to hear it stated a few different ways before it "clicks" and makes sense to me. Most of my students are the same.

Fourth, pay close attention to the Guided Examples throughout this textbook and make sure you clearly understand every step of a procedure. Get out a piece of scratch paper and have a go at the Guided Examples first, to ensure you know how to find the correct answer and aren't just reading passively. This will make homework and tests go more smoothly.

Fifth, if your instructor does not require you to do so, take advantage of the software guides so that you can learn how to perform statistical procedures with a computer. Learning about the mathematics of statistics is valuable. But the reality of the twenty-first century is that almost all statistical work is done by computer. Using the Software Guides to learn how to perform and interpret statistical procedures will give you a valuable skill. More employers and graduate schools are seeking people who can analyze data, and mastering a software program can give you an advantage when you apply for jobs or graduate school.

Finally, if it has been a long time since you have had to do math or if you have forgotten your algebra, then a basic review may be helpful. I wrote this book expecting that students would have already mastered decimals, fractions, percentages, graphs, linear equations, and basic algebra. If you do not remember these concepts or you need a review, I recommend getting an algebra refresher from Khan Academy (www.khanacademy.org/math/algebra) or from Udacity (www.udacity.com/course/intro-algebra-review–ma004). Your university's tutoring center or your instructor may also be useful resources.

For Instructors

When using this textbook, instructors should be aware of a few things. First, it may be difficult to teach every concept in this textbook during a single semester. So, most instructors will have to select the concepts they teach. A route I recommend is to skip Chapter 10 (dependent samples *t*-tests), Chapter 14 (nonparametric and advanced methods), and Chapter 15 (which serves as a reference guide for students who wish to encounter more complex statistical methods as they read published articles or hope to execute their first research project). Some instructors may also find that their students are sufficiently prepared that they can skip – or cover very quickly – Chapter 3 (visual models) and Chapter 4 (central tendency and variability). Additionally, parts of some chapters can be skipped, such as the trimmed mean (in Chapter 4), the arbitrary nature of axes (at the end of Chapter 5), null hypothesis testing of correlation coefficients (in Chapter 12), or the last half of Chapter 11, which covers the logic and assumptions of ANOVA and *post hoc* tests. It is not essential for students to master the minutia of every chapter to have a basic understanding of statistical thinking and methods.

Second, even though the chapters are generally arranged in ascending order of complexity of the statistical procedures, some chapters can be taught out of order. Sometimes I prefer to teach

Chapter 12 (correlation and regression) before Chapter 7 (null hypothesis testing and z-tests) because one of the basic themes of the GLM is that all statistical procedures examine the relationship between variables. This is easiest to see in a correlation, so I sometimes teach about correlations first in order to show how other procedures are themselves correlational. Other instructors may prefer to teach Chapter 13 (χ^2 tests) before Chapter 12 (ANOVA) and save the complex hand calculations of ANOVA for the end of the semester. I also sometimes teach unpaired-samples t-tests (Chapter 10) before paired-samples t-tests (Chapter 9) because it makes the concept of paired data easier for students to understand. The point is to be flexible and teach concepts in the order that works best for you and your students.

I have placed boxed texts throughout the book. These contain ideas that are relevant, but that may not be essential to understanding a concept. They range from handy tips (like mnemonic devices and restatements of important material) to discussions of controversies and statistical practices related to the main ideas of the chapter or the book. Instructors may find these boxes useful in generating classroom discussions. Be creative and flexible in how you use these features to increase students' understanding.

Finally, this textbook includes step-by-step guides for conducting statistical analyses using the Statistical Package for the Social Sciences (SPSS) and Microsoft Excel. Although some instructors will prefer to use other programs, I chose these because (1) Microsoft Excel – or very similar spreadsheet programs – will often be available to students after they graduate, and (2) in the social sciences SPSS is one of the most frequent statistical analysis packages students use in graduate school. I hope that these software guides will make the textbook a useful reference guide for students after they finish their statistics course. Some instructors believe that using statistical software is an important component of an introductory class; this textbook is designed to handle that need, and instructors who wish to emphasize software use will likely concentrate on the software guides and skip the step-by-step calculation found in many chapters. On the other hand, some people believe that it is important to understand the mathematics and theory of statistics in order to properly interpret computer program output. Instructors with this viewpoint will see value in hand calculation while recognizing that computers dominate modern statistical process. Additionally, some instructors are forced to emphasize hand calculation for various reasons, such as a lack of computer resources at their institution or an assessment requirement that emphasizes calculation. In my opinion, there are valid reasons to spend time on hand calculation or on software use. This book is designed to handle both methods of learning statistics.

Last Words

This textbook is the textbook that I wish I had had 15 years ago. I remember what it was like for statistics to be non-intuitive and confusing. The struggle was not fun, but in an odd way I'm grateful for my negative experience in my first statistics class because it has made me a better teacher today. Hopefully, this book will help students leave the course with a positive view of statistics. If there is anything I can do to improve the book, please contact me at rwarne@uvu.edu or via Twitter, where my handle is @Russwarne. I would love to hear from my readers.

Have a great semester!

Acknowledgements

Left to my own devices, I could never have produced a textbook of this quality without the input and help of many people. At Cambridge University Press, the acquisitions editor David Repetto convinced me to write this textbook and had faith in a young, untried psychologist. His confidence never wavered, and his belief in my competence was a great source of encouragement when problems arose or when the writing became a challenge. His associates, Claudia Bona-Cohen, Elisa Adams, and Claire Eudall, all provided help and input in drafts of the chapters. Their responses to my drafts and their efficient coordination of the peer reviews of my chapters were immensely helpful in removing inaccuracies. Indeed, I hold a special appreciation for the anonymous peer reviewers – all of whom (even the ones who passionately hated the drafts they read) made suggestions that improved the book greatly. These reviewers saved me from several foolish errors, and my thanks for them know no bounds.

As with any piece of writing, this book is a product of the time in which it was written. The months when I drafted the chapters (March 2014–September 2016) were a particularly fierce time in the "replication crisis" in psychology, a time in which many long-respected findings were brought into question after these studies could not be replicated. Many readers will notice the concern for replicable findings and stable results, which is a result of the current debate about replication. I appreciate the social and biomedical scientists who have worked tirelessly to make their fields aware of the seriousness of these replication problems. I have tried to incorporate their suggestions into the textbook so that the next generation of students can avoid the mistakes of the past (including some that I made in my career).

I am also indebted to several colleagues whose ideas have shaped my thinking about statistical issues. Ross Larsen gave me insight into important issues related to the central limit theorem and has been a valued friend and coauthor since my graduate school days. The statistical thinking of my former professors and mentors, including Bruce Thompson and Oi-mon Kwok, is also apparent in most chapters. One unanticipated influence on the content of the book was Twitter. Ulrich Schimmack, Victoria Savalei, Daniël Lakens, Jelte Wicherts, Uri Simonsohn, Rickard Carlsson, and others provided inspiration and clarification for chapters – often without knowing that they were doing so. I have said it many times, and now I want to say it in print: "Science Twitter is the best Twitter."

The hundreds of students I have taught statistics to over the years were another source of help and inspiration for this book. I am very grateful for the students enrolled in my classes as I wrote this book. These students were my "guinea pigs" as I tried out different ways of explaining statistical concepts. Many students helped me see which examples were effective (or not) and worked out the problems in my explanations and datasets.

A few students were particularly helpful as I worked on this book. Kimberlee Waite graciously allowed me to use her data from her study on sibling relationships among people with and without autism. Zachary Rawlings assembled the key for the end-of-chapter problems, many of which were tested as homework assignments for the students in my classes. However, three students were "all-stars" and deserve unending praise for their help: Kathy L. Youngkin, Becky Bytheway, and Liliana López Lemus. These three students read the drafts of every chapter and met with me to give feedback. They ensured that every chapter of the book was student-centered, comprehensible, and used as little jargon as possible. These three students were phenomenal in ferreting out the weaknesses in my writing and teaching. If this book is accessible to students, it is largely because of their efforts.

On a more personal note, it is essential to acknowledge the help and support of my amazing wife, Katie. She was the one who talked me into committing myself to write this book. Every time I had a "writing day," she let me work undisturbed, even though it was difficult for her sometimes. It was often very challenging to meet every single deadline (especially with seven family trips to the hospital during the 18 months I was writing), but with my wife's unwavering support, it happened. I do not deserve a woman as wonderful as she is.

I used to think it was a cliché for authors to state that the credit for the strengths of their book belongs to many people, but that the faults belong to the author alone. Now that I have finished my first book, I understand why so many authors say this. Every person listed in these acknowledgements made the book better; their influence was always positive. As a result, I must frankly admit that the book's shortcomings are solely my responsibility. This isn't magnanimity. Rather, it is reality.

Examples

Chapter(s)	Example	Field
1	Baby naming	Sociology
1	Birth dates and hockey teams	Sports
1	Freud's theories were not scientific	Psychology/philosophy of science
1	Working memory theory is falsifiable	Psychology
1	Personality and job performance	Psychology/business
1, 3, 4, 9	Sibling relationships and autism	Psychology/family studies
1	Parental spending on children	Psychology/sociology/family studies
1	Why do people commit crimes?	Sociology/criminology
1	Changing schools and dropout rates	Education
1	Mate selection (matching hypothesis)	Psychology/sociology/family science
1	Darwin's theory of evolution	Biology
1	Newton's laws of motion	Physics
2	Affection	Psychology/family science
2	Perception of stimuli and actual stimuli intensity	Psychology
2	Levels of data examples (temperature, weight, gender, etc.)	Social sciences
2	Educational and psychological test scores and rating scales	Education/psychology
3	Murder rate in Canadian provinces	Sociology/criminology
3	Deceptive truck ad	Business/marketing
3	Deceptive political poll	Political science/journalism
3	Advanced placement tests (interactive map link)	Education
3	Baby Name Voyager (interactive chart link)	Sociology
4	Outliers in math ability	Education/psychology
4	Resilient children in poverty	Psychology/sociology
4	65% of households have below-average income	Economics/sociology
5	Quetelet's data	Biology
5	Normally distributed variables	Biology/psychology/education

Chapter(s)	Example	Field
5	Non-normally distributed variables	Economics/sociology/psychology
5	Skewed data	Education
5	Bimodal distribution example	Psychology
5	Social workers' clients' unemployment time	Social work
6	Dice rolls	Probability/mathematics
6	Incorrect beliefs in gambling	Psychology
6	Probability of a baby's sex	Probability/mathematics/biology
6	Drawing cards from a playing pack	Probability/mathematics
6	Mood states of people with bipolar disorder	Psychology
7, 8	Impact of anti-seizure medicine exposure *in utero*	Medicine/psychology
7	Invisible dog	Philosophy of science
7	Importance of effect sizes	Psychology/education
7	Reporting impossible *p*-values	Educational psychology
7	Social workers' job satisfaction	Social work
8	Are people taller today?	Human biology/economics
8	First graders' IQ scores	Psychology
9	Quality of life improvement after therapy	Psychology
9	Husbands and wives producing paired scores	Family science
9	Matching on gender and handedness	Neuroscience
9	Medical test errors	Medicine
9	Selecting students for educational programs	Education
9	Effectiveness of play therapy	Psychology
10	Acceptable behavior online for young adults in relationships	Sociology/psychology/family science/communications
10	Schizophrenia in children	Psychology
10	Frequently occurring unlikely events	Sports/probability
10	Are males or females smarter?	Psychology
11	Cross-cultural comparisons of indecisiveness	Anthropology/psychology
11	Stroop Test	Psychology/neuroscience
11	Comparing life stressor for unmarried, married, and recently divorced subjects	Family science

Chapter(s)	Example	Field
11, 14	Aspirin's impact on surviving a heart attack	Medicine
12	Video games and intelligence	Psychology
12	National corruption and legislature functioning	Sociology
12	Talking speed and schizophrenia	Psychology
12	Test anxiety	Psychology
12	Depression and job satisfaction	Psychology
13	Air temperature	Earth science
12	ADHD and creativity	Psychology
12	Sugar consumption and behavioral changes	Psychology/medicine
12	Husbands' and wives' personality traits	Psychology/family science
12	High-school grades and first-year college grades	Education/psychology
12	Height at birth and in adulthood	Medicine/physical anthropology
12	Height and weight	Medicine/physical anthropology
12	Lifelong stability of intelligence	Psychology
12	Stability of extraversion	Psychology
12	Husband and wives' religiousness	Family science/psychology
13	Intergenerational height increases	Biology/physical anthropology
13	Test score increases and decreases in children	Psychology/education
13	Regression to the mean in patients with high blood pressure	Medicine
13	Movie sequels are not very good	Entertainment
13	The *Sports Illustrated* curse	Sports
12	Company earnings over time	Economics
13	High-school size and math achievement scores	Education
13	Football players and restriction of range	Sports
12, 13	Job testing	Psychology/sociology
13	Test validity for different groups	Psychology
12	Vocabulary size differences in children from different income groups	Psychology/sociology
14	The *Titanic*	History/sociology
14	Therapy outcomes	Psychology

Chapter(s)	Example	Field
15	Physical activity in gym classes	Education/sociology/health psychology
15	Gene–environment interaction in breast feeding	Psychology
15	Personality–sex interaction in school grades	Education/psychology
15	Simpson's paradox in gifted education	Education/psychology
15	Antisocial personality disorder symptoms	Psychology

1 Statistics and Models

"Why do I need to take statistics?"

Every semester that I teach statistics I have students asking this question. It's a fair question to ask. Most of my students want to work as therapists, social workers, psychologists, anthropologists, sociologists, teachers, or in other jobs that seem far removed from formulas and numbers. As students majoring in the social sciences, almost all of them are more interested in people than they are in numbers. For many students (including me, a long time ago), statistics is something to get out of the way so that they can devote their attention to the stuff they *really* want to do (Rajecki, Appleby, Williams, Johnson, & Jeschke, 2005).

The purpose of this chapter is to show you why students majoring in the social sciences have to take statistics. This chapter will also explore some introductory concepts and terminology that are necessary to understand the book. The most important concept introduced in this chapter is models, which many experts use to understand their statistical results. Because my goal is to help you think about statistics like a professional, I will use that perspective throughout the book.

Learning Goals

- Explain why students in the behavioral sciences need statistics to be successful in their majors.
- Outline the general process for designing and executing a research study.
- Describe the benefits of using models to understand reality.
- Demonstrate why using a model – even if it is not completely accurate – is useful for people working in the social sciences.
- Identify the three types of models and describe their characteristics.

Why Statistics Matters

Although many students would not choose to take a statistics course, nearly every social science department requires its students to take a statistics course (e.g., Norcross et al., 2016; Stoloff et al., 2010). Why? Apparently, the professors in these departments think that statistics is essential to their students' education, despite what their students may think.

The main reason that many students must take statistics is that research in the social sciences is dominated by methodologies that are statistics-based; this family of methods is called **quantitative research**. Researchers who use quantitative research convert their data into numbers for the purpose of analysis, and the numbers are then analyzed by **statistical methods**. Numerical data are so important that one social scientist even argued that "progress in science is impossible without numbers and measurement as words and rhetoric are not enough" (Bouchard, 2014, p. 569).

Quantitative methods – and therefore statistics – dominate most of the behavioral sciences: psychology, sociology, education, criminal justice, economics, political science, and more. Most researchers working in these fields use statistics to test new theories, evaluate the effectiveness of therapies, and learn about the concepts they study. Even workers who do not conduct research must understand statistics in order to understand how (and whether) to apply scientific knowledge in their daily work. Without statistics a practitioner risks wasting time and money by using ineffective products, therapies, or procedures. In some cases this could lead to violations of ethics codes, accusations of malpractice, lawsuits, and harm to clients or customers. Even students who do not become scientists may need statistics to verify whether an anecdotal observation (e.g., that their company sells more products after a local sports team wins a game than after a losing one) is true. Thus, a mastery of statistics is important to many people, not just researchers and social scientists.

There are four main ways that practitioners use statistics in their work in the social sciences:

1. Separating good research from bad
2. Evaluating the conclusions of researchers
3. Communicating findings to others
4. Interpreting research to create practical, real-world results.

There is some overlap among these four points, so some job tasks will fall into more than one category. Nevertheless, this is still a useful list of ways that professionals use statistics.

Separating good research from bad is important for any practitioner. The quality of the research published in scientific journals varies greatly. Some articles become classics and spark new avenues of research; others report shoddy research. Thus, the fact that a study was published in a scientific journal is not, by itself, evidence of good-quality scientific work. A knowledge of statistics is one of the most important tools that a person can have in distinguishing good research from bad. Having the ability to independently judge research prevents practitioners from being susceptible to fads in their field or from wasting resources on practices that provide few benefits.

The benefits of separating good research from bad research are important for the general public, too (not just practitioners). Most people rely on reports from the news media and the Internet to learn about scientific findings. However, most journalists are not trained scientists and do not have the skills needed to distinguish between a high-quality study and a low-quality one (Yettick, 2015). Readers with statistical training will be able to make these judgments themselves, instead of relying on the judgment of a journalist or social media contacts.

Statistical savviness can also help people in *evaluating researchers' conclusions*. Ideally, the conclusions in a scientific article are supported by the data that the researchers collected. However, this is not always the case. Sometimes researchers misinterpret their data because they either used the wrong statistical procedures or did not understand their results. Having statistical competence can prevent research consumers from being at the mercy of the authors and serve as an independent check on researchers.

Another way that people employed in the social sciences use statistics is in *communicating findings and results to others*, such as their clients and colleagues. Increased global competition now means that stakeholders are demanding evidence that the services they receive are effective. Government entities, insurance companies, school districts, and customers are now more likely than ever to demand that people working in the social sciences use "evidence-based practices," meaning that practitioners are expected to use techniques and tools that are supported by scientific evidence (Capraro & Thompson, 2008). Workers who can understand statistics are at an advantage in this type of environment because they will be able to collect and analyze the data that show that their work is effective. But without statistical data, even the best therapist, teacher, or sociologist could appear to be ineffective – perhaps even incompetent.

Finally, people working in the social sciences must be able to use statistics in *translating research into day-to-day practice*. Because most social science research is quantitative, this means understanding statistical analyses and interpreting them in a way that is relevant to their work. Without statistics, practitioners will not be able to understand or implement new therapies, interventions, or techniques. In time these practitioners' work could become outdated or obsolete.

The Quantitative Research Process. The quantitative research process may take many forms, but generally it requires the researcher to follow these steps:

1. Form a research question or research hypothesis
2. Define the population of interest
3. Select a sample of population members
4. Choose variables to measure and operationalize them
5. Select independent and dependent variables
6. Collect data on the variables
7. Analyze the data with statistics and test the hypothesis to determine whether it is supported by the data.

This is a simplified outline, but it is really all that we need to know to understand statistics. Most students take a research methods course after their statistics course that will explain the research process in detail (Norcross et al., 2016; Stoloff et al., 2010). (And if your department doesn't require a research methods course, you should definitely sign up for one anyway!)

The first step in the scientific research process is to form a **research question** or **research hypothesis**. A **research question** is the question that a research study is designed to answer. For example, in a fascinating sociological study, Lieberson and Mikelson (1995) were interested in the ways that some parents invent original names for their children. They had a central research question: "Do parents who create names do so in such a way that the names still convey their child's gender?" (Lieberson & Mikelson, 1995, p. 933). The researchers understood that – from a sociological perspective – new parents do not randomly create names. Rather, a child's name often communicates information, and one of the most fundamental pieces of information that a name can convey is the child's gender. Therefore, they wanted to learn if invented names communicate information, so they posed their research question and designed their study to answer it. (The results indicated that strangers can indeed usually guess the gender of a child with an original, unique name.)

A **hypothesis** is similar to a research question; but rather than a question, a research hypothesis is a testable belief that researchers have about the outcome of their scientific study. For example, one Canadian research team (Barnsley & Thompson, 1988) studied the impact that a child's birth date had on the likelihood that they would play hockey on an elite youth team. In Canada many youth hockey leagues require students to reach a certain age by January 1, meaning that players with birthdays early in the year would be larger and taller than players with later birthdays. A previous study (Barnsley, Thompson, & Barnsley, 1985) had shown that there were more than about four times more professional hockey players born in January, February, and March than were born in October, November, and December. Therefore, the researchers hypothesized that this trend towards hockey players having birthdays in the beginning of the year would also be apparent in elite youth teams. (Their hypothesis was supported by the research.)

These two examples show an important distinction between research questions and research hypotheses. Research questions tend to be more exploratory, and researchers usually ask research questions when they have little or no basis on which to make an educated guess about the results of their study. On the other hand, research hypotheses are expected results that a researcher has, and this expectation is often formed on the basis of previous research or theory.

A research hypothesis is more than just a belief about the world. To be scientific, a hypothesis must be falsifiable, or possible to disprove. If it is impossible to design a study that would produce results that would disprove the hypothesis, then the hypothesis is not scientific (Popper, 1935/2002). For example, one of the reasons that Freud's theory of psychoanalysis is unscientific was that it is impossible to devise an experiment that would produce results that would show the theory to be untrue. Instead, Freud (and, in later years, his followers) always had some sort of explanation for behaviors that seemed – at first glance – to contradict his theory (Cioffi, 1998). They built up an entire body of ideas that could explain away any apparent evidence that disproved the theory. For example, Freud claimed that male clients who said they did not have a desire to murder their father and marry their mother were using a "defense mechanism" called denial, and the very use of this defense mechanism supported Freud's theory. But not using the defense mechanism would also support psychoanalytic theory. Therefore, there was no way to disprove the theory, making it unscientific.

In contrast, a falsifiable theory could be found to be untrue. If the theory were untrue, then evidence could emerge that would disprove it – forcing scientists to suggest other interpretations of the data. On the other hand, when a falsifiable theory withstands attempts to disprove it, the theory is strengthened and becomes more believable. For example, in one famous article, Miller (1956) proposed that people's working memory had a limited capacity, which was "seven, plus or minus two" (p. 81) and that people would have a great deal of difficulty remembering a longer list of items without a memory aid (e.g., writing things down, using a mnemonic device). This theory is quite easy to test: all it requires is finding the limits of what a person can remember in a short period of time and seeing if they do more poorly when a list exceeds 5–9 items. Despite efforts to disprove this theory, it has held up well and led to an increased understanding of human memory (e.g., Baddeley, 2003) and improved educational practices (e.g., Paas, Renkl, & Sweller, 2003). Likewise, Barnsley and Thompson's (1988) hypothesis that elite youth hockey players would be born earlier in the calendar year is easily falsifiable. For example, if these elite hockey players had birthdays evenly distributed throughout the year, or if most birthdays were at the end of the year, then it would falsify the hypothesis. Thus, the hypothesis that better hockey players are born earlier in the year is a scientific hypothesis.

The second step in this process is to define the **population** of interest. A population consists of every person, event, or object that a researcher could wish to study. The choice of a population is completely at a researcher's discretion. For example, criminal justice researchers could define their population as all crimes committed in Australia in a year. Family science researchers may define their population as all children whose biological parents are divorced or had never married. Psychologists may define their population as every person who suffers from anorexia nervosa. In the social sciences, populations usually consist of people, although they do not have to.

The third step in the quantitative research process is to select a sample from the population. Many populations have millions of people in them, and the constraints of money and time may make it unfeasible to gather data from every person in the population. Moreover, studying people means that often we don't have data on every person in a population because those people may refuse to participate in a study, and it is not ethical to force them to participate. Because of these limitations,

Sidebar 1.1 Example of a research design

Here is an example that should clarify the nature of quantitative research. Many psychologists are interested in personality because they believe it consists of a set of stable traits that influence how a person acts in different circumstances. In one study (Barrick, Stewart, & Piotrowski, 2002), psychologists chose to measure extraversion – which is the degree that a person is outgoing – and job performance. They wanted to test their belief – i.e., their research hypothesis – that more-extraverted people would also be better at their jobs. The operational definitions of these two variables were rather straightforward: an extraversion score obtained from a pencil-and-paper test and a supervisor's rating (on a scale from 1 to 7) of the person's job performance. Psychologists believe that personality is a stable trait through the life span, so it seems unlikely that people's job performance would affect or change their personality. It is much more likely that differences in people's personalities would lead to differences in job performance. Therefore, in this example, extraversion (a personality trait) is an independent variable, and job performance is a dependent variable.

After operationalizing their variables and deciding which variable would be the independent variable and which would be the dependent variable, the researchers collected their data. They found that their belief about extraverted people being better at their jobs was supported by the results of their statistical analyses. This is an oversimplified description of the study. The authors were interested in more than two variables, and some of their statistical analyses were more sophisticated than anything in this book. But this summary serves as a good explanation of the quantitative research process.

---

quantitative researchers in the social sciences almost always select a **sample** of population members to study. A sample is a subset of the population that the researcher collects data from. Ideally, a researcher has a sample that is representative of the population at large so that the data gathered from the sample can teach the researcher about the entire population.

In the fourth step of the process, after selecting a sample, a researcher collects data on specific variables. A **variable** is a characteristic that sample or population members can differ on. For example, in Sidebar 1.1 (see box), the researchers were interested in extraversion and job performance. These are both variables because some sample members in the study were more extraverted (i.e., outgoing) than others, and some sample members were better at their jobs than others. Because there is variation among the sample members, both of these characteristics are variables. On the other hand, a characteristic that is the same for all sample or population members is called a **constant**. In the example a constant may be the species that the sample members belonged to. Because all of them were human and there is no variability among sample members on this trait, this is a constant.

The fifth step of the research process is to choose independent and dependent variables. The **dependent variable** is the outcome variable in the study. The **independent variable** is the variable that is believed to cause changes in the dependent variable. Often social scientists create a design that permits them to have total control over the independent variable, but this is not always possible. After the variables are chosen, it is necessary to operationalize the variable. An **operationalization** is a researcher's definition of a variable that permits a researcher to collect quantitative data on it. Operationalization is very common in the social sciences because many of the things that social scientists are interested in – personality, anger, political opinions, attitudes, racism, and more – are not inherently expressed as numbers. Therefore, it is necessary for researchers to create a definition of the variable that can allow them to collect numerical data. An example of an operationalization is counting the number of times a teacher reprimands a child as a measure of how disruptive the child is in class. Other common operationalizations in the social sciences include scores on a test (as in Sidebar 1.1), ratings of the subject's behavior, and measuring the time a person takes to complete a task.

The majority of quantitative research in the social sciences can be classified as experimental research or correlational research (Cronbach, 1957). These are not the only two types of research designs in the social sciences, but they make up the bulk of research. In **experimental research** a researcher creates controlled changes in an independent variable in order to learn how those changes affect the dependent variable. On the other hand, in **correlational research** the researcher does not manipulate or control the independent variable; rather, the researcher measures people's existing levels of the independent variable. The study in Sidebar 1.1 (Barrick et al., 2002) is a correlational study because the authors did not change or manipulate people's levels of extraversion.

Both kinds of research designs have their benefits and drawbacks. Experimental research designs give scientists more control over their subjects and data, and often experimental research permits more conclusions about cause-and-effect relationships (see Sidebar 12.4). However, because it is frequently laboratory-based, experimental research is often criticized for being too artificial. On the other hand, correlational research is often very applicable to real-life situations, but because of the poor control over the data, researchers usually cannot draw strong conclusions about the nature of the relationships between variables (Cronbach, 1957).

This classification and the discussion of experimental and correlational research is simplified. Other common designs include descriptive research and quasi-experimental research. Descriptive research describes variables and often does not attempt to investigate relationships among variables. Opinion polls are a common example of this type of research. Quasi-experimental methods are a hybrid between experimental and correlational designs, where researchers attempt to manipulate a variable, but may not have complete control over it (as when subjects choose whether to participate in an intervention or control group). In a research methods class, you will learn more about these four research designs and perhaps additional important ones, such as single-case designs and qualitative research.

The sixth step in the quantitative research process is to collect data. The mechanics of data collection is beyond the scope of this book, but it is necessary to mention that it is one of the steps of quantitative research. Finally, the seventh and last step is to analyze data with statistics. Most of this book will be concerned with this last step in the quantitative research process.

Qualitative Research. Although quantitative research is the predominant methodology in the social sciences, it is important to mention the principal alternative to quantitative research: a methodology called **qualitative research**. Researchers who specialize in qualitative research believe that it is not beneficial to convert aspects of their sample members into numbers, because most people do not experience their world through numbers. For example, in response to the Barrick et al. (2002) study (see Sidebar 1.1), a qualitative researcher would say that it is nonsensical and too simplistic to convert a person's job performance into a number ranging from 1 to 7. Instead, a qualitative researcher might interview the subject's supervisor to learn which aspects of the job the person excels at and how the person views the work experience, and would then analyze the text of the interview transcript to learn about the nuances of the person's experience. Qualitative methods are popular in anthropology, family science, and some branches of education, psychology, and political science. However, because qualitative research is a methodology that intrinsically rejects numerical data, we will not discuss it further in this book. Nevertheless, you should be aware that qualitative methodologies are valuable in answering questions that quantitative methods cannot.

Check Yourself!

- What are the four reasons a student majoring in the social sciences needs to take a statistics course?
- What are some of the possible consequences for a practitioner who does not understand statistics?
- Explain the difference between an independent variable and a dependent variable.
- What is the difference between qualitative and quantitative research?
- Why is operationalization necessary in the quantitative research process?
- What is the relationship between a sample and a population?

Two Branches of Statistics

As the science of quantitative data analysis, statistics is a broad field, and it would be impossible for any textbook to cover every branch of statistics while still being of manageable length. In this book we will discuss two branches of statistics: descriptive statistics and inferential statistics. **Descriptive statistics** is concerned with merely describing the data that a researcher has on hand. Table 1.1 shows an excerpt from a real collection of data from a study (Waite, Cardon, & Warne, 2015) about the sibling relationships in families where a child has an autism spectrum disorder. (We will discuss this study and its data in much more detail in Chapters 3 and 10.) Each row in the dataset represents a person and each column in the dataset represents a variable. Therefore, Table 1.1 has 13 people and 6 variables in it. Each piece of information is a **datum** (plural: **data**), and because every person in the table has a value for every variable, there are 84 data in the table (13 people multiplied by 6 variables = 78 data). A compilation of data is called a **dataset**.

Sidebar 1.2 Terminology

Remember that the term "statistics" has two major meanings. The first meaning of "statistics" is the science of data analysis. The second meaning of "statistics" is the procedures of an analysis, which are often used to estimate population parameters (Urbina, 2014).

Even though the dataset in Table 1.1 is small, it is still difficult to interpret. It takes a moment to ascertain, for example, that there are more females than males in the dataset, or that most people are satisfied with their relationship with their sibling with autism. Table 1.1 shows just an excerpt of the data. In the study as a whole, there were 45 variables for 13 subjects, which totals to 585 data. No person – no matter how persistent and motivated they are – could understand the entire dataset without some simplification. This is actually a rather small dataset. Most studies in the social sciences have much larger sample sizes. The purpose of descriptive statistics is to describe the

Table 1.1 Example of quantitative dataset (from **Waite, Cardon, & Warne,** 2015)

Gender	Are satisfied with sibling relationship	Believe that the sibling understands the respondent well	Believe that the sibling understands respondent's interests	Subject worries about their sibling with autism	Believe that parents treated the sibling with autism differently
Female	5	4	3	5	1
Female	5	4	4	1	3
Female	5	5	5	4	5
Female	5	4	2	5	3
Male	3	4	3	5	4
Female	3	1	1	4	5
Male	4	2	2	5	4
Female	4	4	3	4	4
Male	2	2	2	5	5
Female	4	3	3	4	3
Female	5	3	4	5	4
Male	3	2	2	4	5
Female	5	3	4	1	4

Note. In this table, 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree.

dataset so that it is easier to understand. For example, we could use descriptive statistics to say that in the range of scores on the variable that measures people's satisfaction with their sibling relationship, the average score is 4.1, while the average score on the variable measuring whether the sibling with autism understands the respondent's interests is 2.9. Chapters 2–5 are concerned with descriptive statistics.

On the other hand, if a researcher only has sample data on hand, descriptive statistics tell the researcher little about the population. A separate branch of statistics, termed **inferential statistics**, was created to help researchers use their sample data to draw conclusions (i.e., inferences) about the population. Inferential statistics is a more complicated field than descriptive statistics, but it is also far more useful. Few social scientists are interested just in the members of their sample. Instead, most are interested in their entire population, and so many social scientists use inferential statistics to learn more about their population – even though they don't have data from every population member. In fact, they usually only have data from a tiny portion of population members. Inferential statistics spans Chapters 6–15 of this book.

An example of a use of inferential statistics can be found in a study by Kornrich (2016). This researcher used survey data to examine the amount of money that parents spend on their children. He divided his sample into five groups, ranked from the highest income to the lowest income. He then found the average amount of money that the parents in each group spent on their children and used inferential statistics to estimate the amount of money each group in the population would spend on their children. Unsurprisingly, richer parents spent more money on their children, but Kornrich (2016) also found that the gap in spending on children between the richest 20% and poorest 20% of families had widened between 1972 and 2010. Because Kornrich used inferential statistics, he could draw conclusions about the general population of parents – not just the parents in his sample.

Check Yourself!

- What is the purpose of descriptive statistics?
- What is the purpose of inferential statistics?
- How are inferential statistics different from descriptive statistics?

Models

This book is not organized like most other textbooks. As the title states, it is built around a **general linear model (GLM)** approach. The GLM is a family of statistical procedures that help researchers ascertain the relationships between variables. Chapter 7 explains the GLM in depth. Until then, it is important to understand the concept of a model.

When you hear the word "model," what do you think of? Some people imagine a fashion model. Others think of a miniature airplane model. Still others think of a prototype or a blueprint. These are all things that are called "models" in the English language. In science, **models** are "simplifications of a complex reality" (Rodgers, 2010, p. 1). Reality is messy and complicated. It is hard to understand. In fact, reality is so complex – especially in the social sciences – that in order for people to comprehend it, researchers create models.

An example from criminology can illustrate the complexity of reality and the need for models. One of the most pressing questions in criminology is understanding who will commit crimes and why. In reality, it is impossible to comprehend every influence that leads to a person's decision to commit a crime (or not). This would mean understanding the person's entire personal history, culture, thoughts, neighborhood, genetic makeup, and more. Andrews and Bonta (2010) have developed the risk-need-responsivity (RNR) model of criminal conduct. Although not its only purpose, the RNR model can help users establish the risk that someone will commit a crime. Andrews and Bonta do not do this by attempting to understand every aspect of a person. Rather, they have chosen a limited number of variables to measure and use those to predict criminal activity. Some of these variables include a history of drug abuse, previous criminal behavior, whether the person is employed, the behavior of their friends, and the presence of certain psychological diagnoses (all of which affect the probability that someone will commit a crime). By limiting the number of variables they measure and use, Andrews and Bonta have created a model of criminal behavior that has been successful in identifying risk of criminal behavior and reducing offenders' risk of future reoffending after treatment (Andrews, Bonta, & Wormith, 2011). This model – because it does not contain every possible influence on a person's criminal behavior – is simplified compared to reality.

This example illustrates an important consequence of creating a model. Because models are simplified, every model is – in some way – wrong. Andrews and Bonta (2010) recognize that their model does not make perfect predictions of criminal behavior every time. Moreover, there are likely some influences not included in the RNR model that may affect the risk of criminal behavior, such as a cultural influence to prevent family shame or the dying request of a beloved relative. Therefore, one can think of a trade-off between model simplicity and model accuracy: simpler models are easier to understand than reality, but this simplification comes at a cost because simplicity makes the model wrong. In a sense, this is true of the types of models most people usually think about. A miniature airplane model is "wrong" because it often does not include many of the parts that a real airplane has. In fact, many model airplanes don't have any engines – a characteristic that definitely is *not* true of real airplanes!

Because every model is wrong, it is not realistic to expect models to be perfectly accurate. Instead, models are judged on the basis of how *useful* they are. A miniature model airplane may be useless in understanding how a full-sized airplane works, but it may be very helpful in understanding the aerodynamic properties of the plane's body. However, a different model – a blueprint of the engine – may be helpful in understanding how the airplane obtains enough thrust and lift to leave the ground. As this example shows, the usefulness of the model may depend on the goals of the researcher. The engineer interested in aerodynamics may have little use for the engine blueprint, even though a different engineer would argue that the engine blueprint is a vital aspect of understanding the airplane's function.

This example also shows one last important characteristic of models: often multiple models can fit reality equally well. In other words, it is possible for different models to fit the same reality, such as the miniature airplane model and the plane engine blueprint (Meehl, 1990). As a result, even if a model explains a phenomenon under investigation very well, it may not be the only model that could fit reality well. In fact, there is no guarantee that the model is even the best possible model. Indeed, many researchers in the social sciences are interested in improving their models because that would lead to an improved understanding of the things they investigate. This improvement can happen by combining two models together, finding improved operationalizations of variables, or eliminating unnecessary parts from a model.

Three Types of Models. There are three kinds of models: (1) statistical models, (2) visual models, and (3) theoretical models. **Statistical models** are models that use a limited number of specific variables and are expressed as mathematical equations. These models are often the result of quantitative research, and they are very common in the social sciences. For example, in a study on the consequences of changing high schools, Gasper, DeLuca, and Estacion (2012) found that the dropout rate was 2.36 times higher for students who attended two high schools between grades 9 and 12 than for those who attended just one high school during those same grades (19.1% compared to 8.1%; see p. 502). This is a simple mathematical model, but it basically tells users that the probability that students will drop out of high school more than doubles if they have changed schools.

This example also illustrates some of the characteristics of models in general: this model is simpler than the real world, and therefore this statistical model is wrong to an extent. The decision to drop out of high school is probably the result of many influences. This statistical model does not take into account a student's history, such as academic difficulties, discipline problems, family environment, extracurricular activities, and more. It's unlikely that most high-school dropouts decide to stop their education solely because they changed schools.

Notwithstanding these limitations, the statistical model is still useful. In their study Gasper et al. (2012) found that over one-fourth of all students changed schools in their 4-year high-school career. For school personnel this can be an indicator that they should be aware of the student's risk of

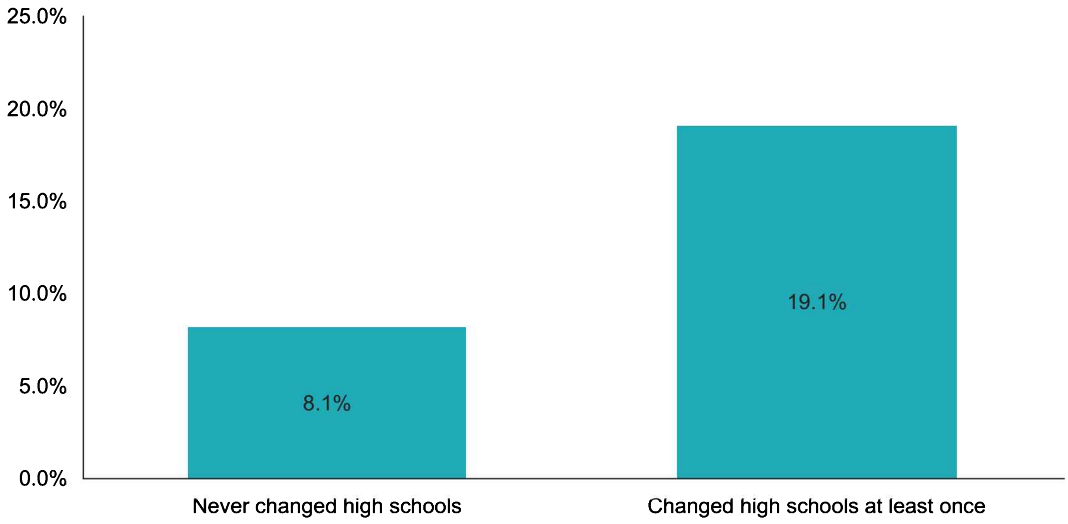

Figure 1.1 A bar graph showing the different rates of dropping out of high school for students who either never changed high schools (*left*) or changed high schools at least once (*right*). This bar graph is an example of a simple visual model. Data from Gaspar et al. (2012, p. 502).

dropping out. Knowing whether a student has changed schools is an easy piece of information to obtain. If it makes teachers and administrators aware of the academic challenges a student faces, it may help them target specific students for extra help or support in an effort to keep them in school.

A **visual model** – which is the second type of model – is a visual representation of data. By creating a picture of the data, a researcher or reader can understand the data at a glance and often without much effort. Thus, visual models simplify data and create a simplified version of the data. An example of a visual model (i.e., a bar graph) is Figure 1.1, which shows the differences in dropout rate between students who stay at the same school and those who change schools. Most readers have experience of visual models, such as line graphs, bar graphs, pie charts, and more. However, in Chapter 3, we will discuss these models and other common visual models in statistics, such as histograms, stem-and-leaf plots, and scatterplots.

The final type of model, the **theoretical model**, is not described in terms of numbers or mathematical equations. Rather, a theoretical model is a causal explanation of reality, usually in verbal terms. An example of a theoretical model in the social sciences is the "matching hypothesis," which states that when selecting a mate people tend to find a person who is about as socially desirable as themselves. This occurs because people self-assess their own desirability and then maximize their chances of finding a desirable mate by (1) avoiding people who would have a high probability of rejecting them, and (2) avoiding potential mates who are much less desirable than themselves because they can likely find someone more desirable (Shaw Taylor, Fiore, Mendelsohn, & Cheshire, 2011). The matching hypothesis explains, for example, why romantic partners tend to have similar levels of physical attractiveness (Feingold, 1988).

But this theoretical model is – like the statistical model example – simplified, and that makes it wrong to an extent because people may become romantically involved with one another for reasons other than "social desirability." The main differences between this theoretical model and a statistical model are that (1) the theoretical model suggests a causal explanation for phenomena, and (2) the theoretical model does not use numbers or equations in its expression.