# Statistics in Plain English

Fourth Edition

TIMOTHY C. URDAN

# Statistics in Plain English

This introductory textbook provides an inexpensive, brief overview of statistics to help readers gain a better understanding of how statistics work and how to interpret them correctly. Each chapter describes a different statistical technique, ranging from basic concepts like central tendency and describing distributions to more advanced concepts such as *t* tests, regression, repeated-measures ANOVA, and factor analysis. Each chapter begins with a short description of the statistic and when it should be used. This is followed by a more in-depth explanation of how the statistic works. Finally, each chapter ends with an example of the statistic in use, and a sample of how the results of analyses using the statistic might be written up for publication. A glossary of statistical terms and symbols is also included. Using the author's own data and examples from published research and the popular media, the book is a straightforward and accessible guide to statistics.

New features in the fourth edition include:

- sets of work problems in each chapter with detailed solutions and additional problems online to help students test their understanding of the material;
- new "Worked Examples" to walk students through how to calculate and interpret the statistics featured in each chapter;
- new examples from the author's own data, published research, and the popular media to help students see how statistics are applied and written about in professional publications;
- many more examples, tables, and charts to help students visualize key concepts, clarify concepts, and demonstrate how statistics are used in the real world;
- a more logical flow, with correlation directly preceding regression, and a combined glossary appearing at the end of the book;
- a "Quick Guide to Statistics, Formulas, and Degrees of Freedom" at the start of the book, plainly outlining each statistic and when students should use them;
- greater emphasis on (and description of) effect size and confidence interval reporting, reflecting their growing importance in research across the social science disciplines;
- an expanded website at www.routledge.com/cw/urdan with PowerPoint presentations, chapter summaries, a new test bank, interactive problems and detailed solutions to the text's work problems, SPSS datasets for practice, links to useful tools and resources, and videos showing how to calculate statistics, how to calculate and interpret the appendices, and how to understand some of the more confusing tables of output produced by SPSS.

*Statistics in Plain English, Fourth Edition* is an ideal guide to statistics and research methods, for courses that use statistics taught at the undergraduate or graduate level, or as a reference tool for anyone interested in refreshing their memory on key statistical concepts. The research examples are from psychology, education, and other social and behavioral sciences.

**Timothy C. Urdan** is Professor of Psychology at Santa Clara University.

"Urdan's text provides exactly the right depth of information needed to understand statistical concepts from both a theoretical and practical perspective. His style of writing is extremely engaging and helpful as he balances his explanations of theory and analyses with many real world examples that help situate the practical application of the concepts and analyses."

**—Robyn Cooper, Drake University, USA**

"The book explains clearly many concepts in statistics and is written in an unintimidating and readable style. Both undergraduate and graduate students will find it helpful as an introduction to statistics."

**—Bridget Sheng, Western Illinois University, USA**

"This is a straight-forward and yet comprehensive treatment, providing students with a basic understanding of statistics in the social sciences without bogging them down with too many equations and complex examples."

**—Robert M. Bernard, Concordia University, USA**

"I know of no textbook on the market that even comes close to satisfying breadth of coverage/content readability/value for price point as does the Urdan text. … My students continually tell me it is very helpful and easy to understand. … The Website is a very useful tool … Urdan writes very clearly and explains concepts at a level that is appropriate for multiple audiences, including undergraduate, graduate, and even academic audiences … I like that Urdan uses a variety of examples as opposed to competing texts that limit their examples to specific disciplines."

**—Catherine A. Roster, University of New Mexico, USA**

"It is very clear and wherever possible uses normal, plain English. ... Simple terms are preferred to jargon and where key terms are introduced they are done so clearly. … The online support materials … provide useful additional tools for both students and lecturers. … From the examples used it seems to be aimed at a variety of social science students … Psychology, Childhood and Youth and Crime Studies. … I would like to use the book and would encourage students to do so. It contains a lot they need to know in order to understand … statistical tests."

**—Nick Lund, Manchester Metropolitan University, UK**

"Problem sets ... [are] a welcome addition. … The interactive analyses, podcasts, PowerPoints and test questions are all a positive. … The author understands the material quite well and does a good job explaining terms and concepts. … The pedagogical aids … trump … what exists currently for adoption."

**—Nicholas Corsaro, University of Cincinnati, USA**

"I really like the 'writing it up section'. I think that adds a lot to what my students need in regards to statistics. … This … text will be purchased by people needing a baseline level of statistical knowledge. … I enjoyed the text and writing style. … The book would be beneficial to the students. … Mr. Urdan does not attempt to take it too deep and keeps the book succinct enough to make it usable in a beginner stats class."

**—Andrew Tinsley, Eastern Kentucky University, USA**

"I am highly impressed with the latest edition (4th) of Tim Urdan's wonderful text. This author continues to deliver a highly readable yet nuanced text on research methods and fundamental statistics. I use this book in a number of graduate courses for clinical psychology students. It would also serve advanced undergraduates. In addition to clear writing and covering all the basics, I believe another essential strength is the author's commitment to training minds for critical thinking."

**—Jamie K. Lilie, Argosy University, USA**

# Statistics in Plain English

## Fourth Edition

Timothy C. Urdan

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

Fourth edition published 2017
by Routledge
711 Third Avenue, New York, NY 10017, USA

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN, UK

*Routledge is an imprint of the Taylor & Francis Group, an Informa business*



Third edition published 2010 by Routledge

*Library of Congress Cataloging in Publication Data*
Names: Urdan, Timothy C.
Title: Statistics in plain English / Timothy C. Urdan, Santa Clara University.
Description: 4th edition. | New York, NY : Routledge, 2016. |
Includes bibliographical references and index.
Identifiers: LCCN 2016001650 (print) | LCCN 2016005945 (ebook) | ISBN 9781138838338 (hardback) | ISBN 9781138838345 (pbk.) | ISBN 9781315723112 ()
Subjects: LCSH: Statistics–Textbooks. | Mathematical statistics–Textbooks.
Classification: LCC QA276.12.U75 2016 (print) | LCC QA276.12 (ebook) |
DDC 519.5–dc23 LC record available at http://lccn.loc.gov/2016001650

ISBN: 978-1-13883-833-8 (hbk)
ISBN: 978-1-13883-834-5 (pbk)
ISBN: 978-1-31572-311-2 (ebk)

Typeset in ACaslon
by Out of House Publishing

# Brief Contents

**Preface** xi
**Acknowledgments** xv
**About the Author** xvii
**Quick Guide to Statistics, Formulas, and Degrees of Freedom** xix

**Chapter 1** Introduction to Social Science Research Principles and Terminology 1

**Chapter 2** Measures of Central Tendency 13

**Chapter 3** Measures of Variability 21

**Chapter 4** The Normal Distribution 33

**Chapter 5** Standardization and *z* Scores 43

**Chapter 6** Standard Errors 57

**Chapter 7** Statistical Significance, Effect Size, and Confidence Intervals 73

**Chapter 8** *t* Tests 93

**Chapter 9** One-Way Analysis of Variance 113

**Chapter 10** Factorial Analysis of Variance 133

**Chapter 11** Repeated-Measures Analysis of Variance 149

**Chapter 12** Correlation 165

**Chapter 13** Regression 183

**Chapter 14** The Chi-Square Test of Independence 205

**Chapter 15** Factor Analysis and Reliability Analysis: Data Reduction Techniques 213

**Appendices**
**Appendix A**: Area Under the Normal Curve Beyond *z* 229
**Appendix B**: Critical Values of the *t* Distributions 231
**Appendix C**: Critical Values of the *F* Distributions 233
**Appendix D**: Critical Values of the Studentized Range Statistic (for Tukey HSD Tests) 239
**Appendix E**: Critical Values of the $\chi^2$ Distributions 243

**Bibliography** 245

**Glossary of Terms** 247

**Glossary of Symbols** 257

**Index** 259

# Contents

| | | |
|---|---|---|
| **Preface** | | xi |
| **Acknowledgments** | | xv |
| **About the Author** | | xvii |
| **Quick Guide to Statistics, Formulas, and Degrees of Freedom** | | xix |
| **Chapter 1** | Introduction to Social Science Research Principles and Terminology | 1 |
| | Populations, Samples, Parameters, and Statistics | 1 |
| | Inferential and Descriptive Statistics | 2 |
| | Sampling Issues | 3 |
| | Types of Variables and Scales of Measurement | 4 |
| | Research Designs | 5 |
| | Making Sense of Distributions and Graphs | 7 |
| | Wrapping Up and Looking Forward | 12 |
| | Work Problems | 12 |
| **Chapter 2** | Measures of Central Tendency | 13 |
| | Measures of Central Tendency in Depth | 14 |
| | Example: The Mean, Median, and Mode of Skewed Distributions | 15 |
| | Writing it Up | 19 |
| | Wrapping Up and Looking Forward | 19 |
| | Work Problems | 20 |
| **Chapter 3** | Measures of Variability | 21 |
| | Range | 21 |
| | Variance | 22 |
| | Standard Deviation | 22 |
| | Measures of Variability in Depth | 22 |
| | Why Have Variance? | 26 |
| | Examples: Examining the Range, Variance, and Standard Deviation | 27 |
| | Worked Examples | 30 |
| | Wrapping Up and Looking Forward | 32 |
| | Work Problems | 32 |
| **Chapter 4** | The Normal Distribution | 33 |
| | Characteristics of the Normal Distribution | 33 |
| | Why Is the Normal Distribution So Important? | 33 |
| | The Normal Distribution in Depth | 34 |
| | The Relationship Between the Sampling Method and the Normal Distribution | 35 |
| | Skew and Kurtosis | 36 |
| | Example 1: Applying Normal Distribution Probabilities to a Normal Distribution | 37 |
| | Example 2: Applying Normal Distribution Probabilities to a Nonnormal Distribution | 39 |
| | Wrapping Up and Looking Forward | 40 |
| | Work Problems | 40 |
| **Chapter 5** | Standardization and $z$ Scores | 43 |
| | Standardization and $z$ Scores in Depth | 43 |

| | | |
|---|---|---|
| | Interpreting *z* Scores | 44 |
| | Examples: Comparing Raw Scores and *z* Scores | 51 |
| | Worked Examples | 53 |
| | Wrapping Up and Looking Forward | 54 |
| | Work Problems | 54 |
| **Chapter 6** | Standard Errors | 57 |
| | What Is a Standard Error? | 57 |
| | Standard Errors in Depth | 58 |
| | How to Calculate the Standard Error of the Mean | 60 |
| | The Central Limit Theorem | 61 |
| | The Normal Distribution and *t* Distributions: Comparing *z* Scores and *t* Values | 62 |
| | The Use of Standard Errors in Inferential Statistics | 65 |
| | Example: Sample Size and Standard Deviation Effects on the Standard Error | 66 |
| | Worked Examples | 68 |
| | Wrapping Up and Looking Forward | 70 |
| | Work Problems | 70 |
| **Chapter 7** | Statistical Significance, Effect Size, and Confidence Intervals | 73 |
| | Statistical Significance in Depth | 74 |
| | Limitations of Statistical Significance Testing | 80 |
| | Effect Size in Depth | 82 |
| | Confidence Intervals in Depth | 84 |
| | Example: Statistical Significance, Confidence Interval, and Effect Size for a One-Sample *t* Test of Motivation | 86 |
| | Wrapping up and Looking Forward | 90 |
| | Work Problems | 90 |
| **Chapter 8** | *t* Tests | 93 |
| | What Is a *t* Test? | 93 |
| | *t* Distributions | 93 |
| | The One-Sample *t* Test | 93 |
| | The Independent Samples *t* Test | 94 |
| | Dependent (Paired) Samples *t* Test | 94 |
| | Independent Samples *t* Tests in Depth | 95 |
| | The Standard Error of the Difference between Independent Sample Means | 96 |
| | Determining the Significance of the *t* Value for an Independent Samples *t* Test | 98 |
| | Paired or Dependent Samples *t* Tests in Depth | 100 |
| | Example 1: Comparing Boys' and Girls' Grade Point Averages | 102 |
| | Example 2: Comparing Fifth- and Sixth-Grade GPAs | 103 |
| | Writing it Up | 104 |
| | Worked Examples | 105 |
| | Wrapping Up and Looking Forward | 110 |
| | Work Problems | 110 |
| **Chapter 9** | One-Way Analysis of Variance | 113 |
| | ANOVA vs. Independent *t* Tests | 113 |
| | One-Way ANOVA in Depth | 114 |
| | Deciding if the Group Means Are Significantly Different | 117 |
| | Post-Hoc Tests | 118 |
| | Effect Size | 120 |

| | | |
|---|---|---|
| | Example: Comparing the Sleep of 5-, 8-, and 12-Year-Olds | 122 |
| | Writing it Up | 126 |
| | Worked Example | 126 |
| | Wrapping Up and Looking Forward | 129 |
| | Work Problems | 130 |
| **Chapter 10** | Factorial Analysis of Variance | 133 |
| | When to Use Factorial ANOVA | 133 |
| | Some Cautions | 134 |
| | Factorial ANOVA in Depth | 134 |
| | Interpreting Main Effects in the Presence of an Interaction Effect | 138 |
| | Testing Simple Effects | 140 |
| | Analysis of Covariance | 140 |
| | Illustration of Factorial ANOVA, ANCOVA, and Effect Size with Real Data | 141 |
| | Example: Performance, Choice, and Public vs. Private Evaluation | 143 |
| | Writing it Up | 145 |
| | Wrapping Up and Looking Forward | 145 |
| | Work Problems | 146 |
| **Chapter 11** | Repeated-Measures Analysis of Variance | 149 |
| | When to Use Each Type of Repeated-Measures Technique | 149 |
| | Repeated-Measures ANOVA in Depth | 151 |
| | Repeated-Measures Analysis of Covariance (ANCOVA) | 154 |
| | Adding an Independent Group Variable | 156 |
| | Example: Changing Attitudes about Standardized Tests | 158 |
| | Writing it Up | 162 |
| | Wrapping Up and Looking Forward | 162 |
| | Work Problems | 163 |
| **Chapter 12** | Correlation | 165 |
| | When to Use Correlation and What it Tells Us | 165 |
| | Pearson Correlation Coefficients in Depth | 167 |
| | A Brief Word on Other Types of Correlation Coefficients | 177 |
| | Example: The Correlation Between Grades and Test Scores | 177 |
| | Writing it Up | 179 |
| | Wrapping Up and Looking Forward | 180 |
| | Work Problems | 180 |
| **Chapter 13** | Regression | 183 |
| | Simple vs. Multiple Regression | 183 |
| | Variables Used in Regression | 184 |
| | Regression in Depth | 184 |
| | Multiple Regression | 190 |
| | Example: Predicting the Use of Self-Handicapping Strategies | 197 |
| | Writing it Up | 199 |
| | Worked Examples | 200 |
| | Wrapping Up and Looking Forward | 202 |
| | Work Problems | 203 |
| **Chapter 14** | The Chi-Square Test of Independence | 205 |
| | Chi-Square Test of Independence in Depth | 206 |

Example: Generational Status and Grade Level 209
Writing it Up 210
Worked Example 210
Wrapping Up and Looking Forward 212
Work Problems 212

**Chapter 15** Factor Analysis and Reliability Analysis: Data Reduction Techniques 213
Factor Analysis in Depth 213
A More Concrete Example of Exploratory Factor Analysis 216
Confirmatory Factor Analysis: A Brief Introduction 220
Reliability Analysis in Depth 221
Writing it Up 224
Work Problems 224
Wrapping Up 226

Appendices

Appendix A: Area Under the Normal Curve Beyond *z* 229

Appendix B: Critical Values of the *t* Distributions 231

Appendix C: Critical Values of the *F* Distributions 233

Appendix D: Critical Values of the Studentized Range Statistic (for Tukey HSD Tests) 239

Appendix E: Critical Values of the $\chi^2$ Distributions 243

Bibliography 245

Glossary of Terms 247

Glossary of Symbols 257

Index 259

# Preface

## Why Use Statistics?

As a researcher who uses statistics frequently, and as an avid listener of talk radio, I find myself yelling at my radio daily. Although I realize that my cries go unheard, I cannot help myself. As radio talk show hosts, politicians making political speeches, and the general public all know, there is nothing more powerful and persuasive than the personal story, or what statisticians call anecdotal evidence. When a parent of an autistic child appears on television, crying and proclaiming that vaccinations caused the child to develop autism, one cannot help but be moved. Similarly, when a politician claims that public schools are failing and the solution is to break up teacher unions and create more charter schools, the argument can be compelling. But as researchers, we must go beyond the personal story or impassioned argument and look for broader evidence. How can we tell whether vaccines cause autism? How do we decide whether public schools are failing, whether teacher unions help or hinder valuable reform, and whether charter schools do a better job of educating students? To answer these questions, we need good data, and then we need to analyze the data using the appropriate statistics.

Many people have a general distrust of statistics, believing that crafty statisticians can "make statistics say whatever they want" or "lie with statistics." In fact, if a researcher calculates the statistics correctly, he or she cannot make them say anything other than what they say, and statistics never lie. Rather, crafty researchers can interpret what the statistics *mean* in a variety of ways, and those who do not understand statistics are forced to either accept the interpretations that statisticians and researchers offer or reject statistics completely. I believe a better option is to gain an understanding of how statistics work and then use that understanding to interpret the statistics one sees and hears for oneself. The purpose of this book is to make it a little easier to understand statistics.

## Uses of Statistics

One of the potential shortfalls of anecdotal data is that they are idiosyncratic. One of our cognitive shortcomings, as people, is that we tend to believe that if something is true for us, it must be fact. "I eat a multivitamin every day and I haven't been sick for 20 years!" "My grandmother smoked a pack a day for 50 years and she lived until she was 96!" "My parents spanked me and I turned out fine!" Although these statements may (or may not!) be true for the individuals who uttered them, that does not mean they are true for everyone, or even for most people. Statistics allow researchers to collect information, or data, from a large number of people and then summarize their typical experience. Do *most* people who take multivitamins live healthier lives? Do most people who smoke a pack a day live shorter lives than people who do not smoke? Is there any association between whether one is spanked and how one "turns out," however that is defined? Statistics allow researchers to take a large batch of data and *summarize* it into a couple of numbers, such as an average. Of course, when many data are summarized into a single number, a lot of information is lost, including the fact that different people have very different experiences. So it is important to remember that, for the most part, statistics do not provide useful information about each individual's experience. Rather, researchers generally use statistics to make *general* statements about a population. Although personal stories are often moving or interesting, it is also important to understand what the *typical* or *average* experience is. For this, we need statistics.

Statistics are also used to reach conclusions about general differences between groups. For example, suppose that in my family, there are four children, two men and two women. Suppose that the women in my family are taller than the men. This personal experience may lead me to the conclusion that women are generally taller than men. Of course, we know that, on average, men are taller than women. The reason we know this is because researchers have taken large, random samples of men and women and compared their average heights. Researchers are often interested in making such comparisons: Do cancer patients survive longer using one drug than another? Is one method of teaching children to read more effective than another? Do men and women differ in their enjoyment of a certain movie? To answer these questions, we need to collect data from randomly selected samples and compare these data using statistics. The results we get from such comparisons are often more trustworthy than the simple observations people make from nonrandom samples, such as the different heights of men and women in my family.

Statistics can also be used to see if scores on two variables are related and to make predictions. For example, statistics can be used to see whether smoking cigarettes is related to the likelihood of developing lung cancer. For years, tobacco companies argued that there was no relationship between smoking and cancer. Sure, some people who smoked developed cancer. But the tobacco companies argued that (a) many people who smoke never develop cancer, and (b) many people who smoke tend to do other things that may lead to cancer development, such as eating unhealthy foods and not exercising. With the help of statistics in a number of studies, researchers were finally able to produce a preponderance of evidence indicating that, in fact, there is a relationship between cigarette smoking and cancer. Because statistics tend to focus on overall patterns rather than individual cases, this research did not suggest that *everyone* who smokes will develop cancer. Rather, the research demonstrated that, on average, people have a greater chance of developing cancer if they smoke cigarettes than if they do not.

With a moment's thought, you can imagine a large number of interesting and important questions that statistics about relationships can help you to answer. Is there a relationship between self-esteem and academic achievement? Is there a relationship between the physical appearance of criminal defendants and their likelihood of being convicted? Is it possible to predict the violent crime rate of a state from the amount of money the state spends on drug treatment programs? If we know the father's height, how accurately can we predict the son's height? These and thousands of other questions have been examined by researchers using statistics designed to determine the relationship between variables in a population. With the rise of the internet, data is now being collected constantly. For example, most casual users of the internet and social media provide information about their age, gender, where they live, how much money they make, how much they spend, what they like to buy, who their friends are, which websites they visit, what they like, whether they are married or single, etc. With the help of statistics, data analysts determine what advertisements you should see when you visit a website, how to attract you to certain websites, and how to get you to encourage your friends (without your knowledge) to like or buy various products. More than ever before, statistics and data are deeply affecting many aspects of your life. With this in mind, wouldn't it be nice to understand a bit more about how these statistics work?

## How to Use this Book

If you are new to statistics, this book can provide an easy introduction to many of the basic, and most commonly used, statistics. Or, if you have already taken a course or two in statistics, this book may be useful as a reference book to refresh your memory on statistical concepts you have encountered in the past. It is important to remember that this book is much less detailed than a traditional statistics textbook. It was designed to provide a relatively short and inexpensive introduction to statistics, with a greater focus on the conceptual part of statistics than the computational, mathematical part. Each of the concepts discussed in this book is more complex than the presentation in this book might suggest, and a thorough understanding of these concepts may be acquired only with the use of a more traditional, more detailed textbook.

With that warning firmly in mind, let me describe the potential benefits of this book, and how to make the most of them. As a researcher and a teacher of statistics, I have found that statistics textbooks often contain a lot of technical information that can be intimidating to nonstatisticians. Although, as I said previously, this information is important, sometimes it is useful to have a short, simple description of a statistic, when it should be used, and how to make sense of it. This is particularly true for students taking only their first or second statistics course, those who do not consider themselves to be "mathematically inclined," and those who may have taken statistics years ago and now find themselves in need of a little refresher. My purpose in writing this book is to provide short, simple descriptions and explanations of a number of statistics that are easy to read and understand.

To help you use this book in a manner that best suits your needs, I have organized each chapter into sections. In the first section, a brief (1 to 2 pages) description of the statistic is given, including what the statistic is used for and what information it provides. The second section of each chapter contains a slightly longer (3 to 12 pages) discussion of the statistic. In this section, I provide a bit more information about how the statistic works, an explanation of how the formula for calculating the statistic works, the strengths and weaknesses of the statistic, and the conditions that must exist to use the statistic. Each chapter includes an example or two in which the statistic is calculated and interpreted. The chapters conclude with illustrations of how to write up the statistical information for publication and a set of work problems.

Before reading the book, it may be helpful to note three of its features. First, some of the chapters discuss more than one statistic. For example, in Chapter 2, three measures of central tendency are described: the mean, median, and mode. Second, some of the chapters cover statistical concepts rather than specific statistical techniques. For example, in Chapter 4, the normal distribution is discussed. There is also a chapter on statistical significance, effect size, and confidence intervals. Finally, you should remember that the chapters in this book are not necessarily designed to be read in order. The book is organized such that the more basic statistics and statistical concepts are in the earlier chapters, whereas the more complex concepts appear later in the book. However, each chapter in the book was written to stand on its own. This was done to enable you to use each chapter as needed. If, for example, you had no problem understanding *t* tests when you learned about them in your statistics class, but find yourself struggling to understand one-way analysis of variance, you may want to skip the *t* test chapter (Chapter 8) and go directly to the analysis of variance chapter (Chapter 9). If you are brand new to statistics, however, keep in mind that some statistical concepts (e.g., *t* tests, ANOVA) are easier to understand if you first learn about the mean, variance, and hypothesis testing.

## New Features in this Edition

This fourth edition of *Statistics in Plain English* includes a number of features not available in the previous editions. Each of the 15 chapters now includes a set of work problems, with solutions and additional work problems provided on the website that accompanies this book. In addition, a section called "Worked Examples" has been added to most of the chapters in the book. In this section, I work through all of the steps to calculate and interpret the statistic featured in the chapter. There is also a link to a video of me calculating each statistic at the end of each Worked Examples section. These videos are available on the website that accompanies this book. Each chapter has also been revised to clarify confusing concepts, add or revise graphs and tables, and provide more examples of how the statistics are used in the real world. Throughout the book, there is a greater emphasis on, and description of, effect size. The supporting materials provided on the website at www.routledge.com/cw/urdan have also been updated, including many new and improved videos showing how to calculate statistics, how to read and interpret the appendices, and how to understand some of the more confusing tables of output produced by SPSS. PowerPoint summaries of each chapter, answers to the work problems, and a set of interactive work problems are also provided on the website. Finally, I have included a "Quick Guide to Statistics, Formulas, and Degrees of Freedom" at the beginning of the book, plainly outlining each statistic and when students should use them.

# Acknowledgments

I would like to sincerely thank the reviewers who provided their time and expertise reading previous drafts of this book and offered very helpful feedback, including but not limited to:

Jason Abbitt, Miami University
Stephen M. Barkan, University of Maine
Robert M. Bernard, Concordia University
Danny R. Bowen, The Southern Baptist Theological Seminary
Nicholas Corsaro, University of Cincinnati
Heather Chapman, Weber State University
Seo-eun Choi, Arkansas State University
Robyn Cooper, Drake University
Margaret Cousins, University of Chester
Robert Crosby, California Baptist University
Clare Davies, University of Winchester
James Green, University of North Alabama
Henriette Hogh, University of Chichester
Matthew Jerram, Suffolk University
Jacyln Kelly, City University of New York, Graduate Center
Jo Ann Kelly, Walsh University
Yeonsoo Kim, University of Nevada, Las Vegas
Jennifer M. Kitchens, Lebanon Valley College
Franz Kronthaler, University of Applied Science HTW Chur
Jamie K. Lilie, Argosy University
Nick Lund, Manchester Metropolitan University
Gay McAlister, Southern Methodist University
Catherine A. Roster, University of New Mexico
Bridget Sheng, Western Illinois University
Joshua Stephens, Cleveland State University
Nathaniel Straight, Loyola University New Orleans
James Swartz, University of Illinois at Chicago
Christine Tartaro, Richard Stockton College of NJ
Andrew Tinsley, Eastern Kentucky University
Jon Yearsley, University College Dublin
Akane Zusho, Fordham University

I could not fit all of your suggestions into this new edition, but I incorporated many of them and the book is better as a result of your hard work. Thanks to Debra Riegert and Fred Coppersmith at Routledge/Taylor & Francis for your patience, help, and guidance. As always, students and colleagues at Santa Clara University made valuable contributions to this book, so thank you to Caitlin Courshon, Bhaumik Dedhia, and Elwood Mills. Thanks, again, to Ella, Nathaniel, and Jeannine for your patience while I worked through your vacation time to finish this book. Finally, thanks to you readers for using this book. We are in this statistics struggle together.

# About the Author

**Timothy C. Urdan** is a professor in the Department of Psychology at Santa Clara University. He received his Ph.D. in Education and Psychology from the University of Michigan, where he received several honors including the School of Education Merit Award, the Horace H. Rackham Predoctoral Fellowship, and the Burke Aaron Hinsdale Scholar Award. He is an associate editor of the *Merrill-Palmer Quarterly* and serves on the editorial board for *Contemporary Educational Psychology*, the *Journal of Educational Psychology*, and the *American Educational Research Journal*. Dr. Urdan is the co-editor of two book series, *Adolescence and Education* and *Advances in Motivation and Achievement*, and also co-edited the *APA Educational Psychology Handbook*. He is a fellow of the American Psychological Association.

# Quick Guide to Statistics, Formulas, and Degrees of Freedom

| Statistic | Symbol | When you use it | Formula | Degrees of freedom (*df*) |
|---|---|---|---|---|
| Mean | $\bar{X}$, μ | To find the average of a distribution. | $\bar{X} = \frac{\Sigma X}{n}, \ \mu = \frac{\Sigma X}{N}$ | |
| Standard deviation (sample) | *s* | To use sample data to estimate the average deviation in a distribution. It is a measure of variability. | $s = \sqrt{\frac{\Sigma(X - \bar{X})^2}{n - 1}}$ | |
| Standard deviation (population) | σ | To find the average deviation in a distribution. It is a measure of variability. | $\sigma = \sqrt{\frac{\Sigma(X - \mu)^2}{N}}$ | |
| Standard score for individual (*z* score) | *z* | To find the difference between an individual score and the mean in standard deviation units. | $z = \frac{X - \mu}{\sigma}$ or $z = \frac{X - \bar{X}}{s}$ | |
| Standard score for mean (*z* score) | *z* | To find the difference between a sample mean and a population mean in standard error units. | $z = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}}$ | |
| Standard error of the mean | $s_{\bar{x}}$, $\sigma_{\bar{x}}$ | To find the average difference between the population mean and the sample means when samples are of a given size and randomly selected. | $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ or $s_{\bar{x}} = \frac{s}{\sqrt{n}}$ | |
| One-sample *t* test | *t* | To determine whether the difference between a sample mean and the population mean is statistically significant. | $t = \frac{\bar{X} - \mu}{s_{\bar{x}}}$ | $n - 1$ |
| Independent samples *t* test | *t* | To determine whether the difference between two independent sample means is statistically significant. | $t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{x}1-\bar{x}2}}$ | $n_1 + n_2 - 2$ |
| Dependent (paired) samples *t* test | *t* | To determine whether the difference between two dependent (i.e., paired) sample means is statistically significant. | $t = \frac{\bar{X} - \bar{Y}}{s_{\bar{D}}}$ | $N - 1$, where *N* is the number of pairs of scores. |
| One-way ANOVA | *F* | To determine whether the difference between two or more independent sample means is statistically significant. | $F = \frac{MS_b}{MS_e}$ | $k - 1$, $n - k$, where *k* is the number of groups and *n* is the number of cases across all samples. |

| Statistic | Symbol | When you use it | Formula | Degrees of freedom (*df*) |
|---|---|---|---|---|
| Cohen's *d* (effect size) | $d$ | To determine the size of an effect (e.g., difference between sample means) in standard deviation units. | $d = \frac{\bar{X}_1 - \bar{X}_2}{\hat{s}}$ | |
| Confidence interval for sample mean | CI | To create an interval within which one is 95% or 99% certain the population parameter (i.e., the population mean) is contained. | $CI_{95} = \bar{X} \pm (t_{95})(s_{\bar{x}})$ | |
| Correlation coefficient (Pearson) | $r$ | To calculate a measure of association between two intervally scaled variables. | $r = \frac{\Sigma(z_x z_y)}{N}$ | |
| Coefficient of determination | $r^2$ | To determine the percentage of variance in one variable that is explained by the other variable in a correlational analysis. It is a measure of effect size. | $r^2$ | |
| *t* test for correlation coefficient | $t$ | To determine whether a sample correlation coefficient is statistically significant. | $t = (r)\sqrt{\frac{N-2}{1-r^2}}$ | $N - 2$, where $N$ is the number of cases in the sample. |
| Regression coefficient | $b$ | To determine the amount of change in the *Y* variable for every change of one unit in the *X* variable in a regression analysis. | $b = r \times \frac{s_y}{s_x}$ | |
| Regression intercept | $a$ | To determine the predicted value of *Y* when *X* equals zero in a regression analysis. | $a = \bar{Y} - b\bar{X}$ | |
| Predicted value of *Y* | $\hat{Y}$ | To determine the predicted value of *Y* for a given value of *X* in a regression analysis. | $\hat{Y} = bX + a$ | |
| Chi-square | $\chi^2$ | To examine whether the frequency of scores in various categories of two categorical variables are different from what would be expected. | $\chi^2 = \sum\left(\frac{(O-E)^2}{E}\right)$ | $R - 1, C - 1$ |

# Chapter 1

# Introduction to Social Science Research Principles and Terminology

When I was in graduate school, one of my statistics professors often repeated what passes, in statistics, for a joke: "If this is all Greek to you, well that's good." Unfortunately, most of the class was so lost we didn't even get the joke. The world of statistics and research in the social sciences, like any specialized field, has its own terminology, language, and conventions. In this chapter, I review some of the fundamental research principles and terminology, including the distinction between samples and populations, methods of sampling, types of variables, the distinction between inferential and descriptive statistics, and a brief word about different types of research designs.

## Populations, Samples, Parameters, and Statistics

A **population** is an individual or group that represents *all* the members of a certain group or category of interest. A sample is a subset drawn from the larger population (see Figure 1.1). For example, suppose that I wanted to know the average income of the current full-time employees at Google. There are two ways that I could find this average. First, I could get a list of every full-time employee at Google and find out the annual income of each member on this list. Because this list contains every member of the group that I am interested in, it can be considered a population. If I were to collect these data and calculate the **mean**, I would have generated a **parameter**, because a parameter is a value generated from, or applied to, a population. Another way to generate the mean income of the full-time employees at Google would be to randomly select a subset of employee names from my list and calculate the average income of this subset. The subset is known as a **sample** (in this case it is a **random sample**), and the mean that I generate from this sample is a type of **statistic**. Statistics are values derived from sample data, whereas parameters are values that are either derived from, or applied to, population data.

It is important to keep a couple of things in mind about samples and populations. First, a population does not need to be large to count as a population. For example, if I wanted to know the average height of the students in my statistics class this term, then all of the members of the class (collectively) would comprise the population. If my class only has five students in it, then my population only has five cases. Second, populations (and samples) do not have to include people. For example, suppose I want to know the average age of the dogs that visited a veterinary clinic in the last year. The population in this study is made up of dogs, not people. Similarly, I may want to know the total amount of carbon monoxide produced by Ford vehicles that were assembled in the United States during 2005. In this example, my population is cars, but not all cars—it is limited to Ford cars, and only those actually assembled in a single country during a single calendar year.

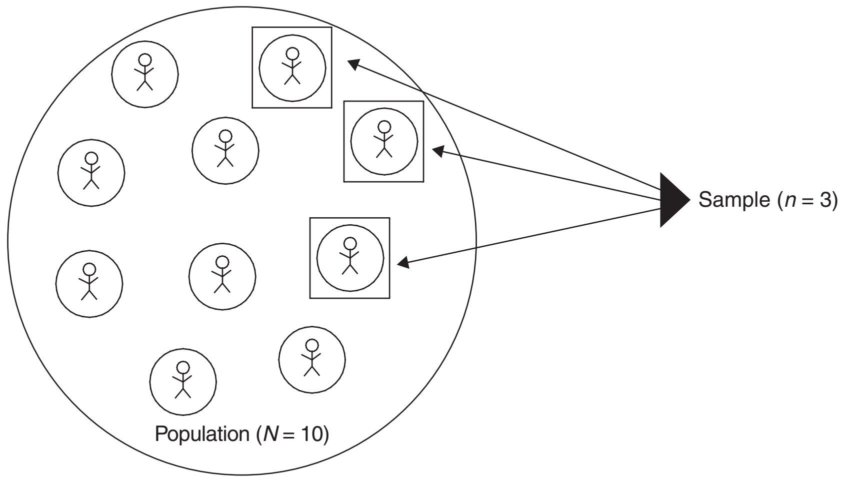


**FIGURE 1.1** A population and a sample drawn from the population.

Third, the researcher generally defines the population, either explicitly or implicitly. In the examples above, I defined my populations (of dogs and cars) explicitly. Often, however, researchers define their populations less clearly. For example, a researcher may conduct a study with the aim of examining the frequency of depression among adolescents. The researcher's sample, however, may include only a group of 15-year-olds who visited a mental health service provider in Connecticut in a given year. This presents a potential problem, and leads directly into the fourth and final little thing to keep in mind about samples and populations: Samples are not necessarily good representations of the populations from which they were selected. In the example about the rates of depression among adolescents, notice that there are two potential populations. First, there is the population identified by the researcher and implied in the research question: adolescents. But notice that adolescents is a very large group, including all human beings, in all countries, between the ages of, say, 13 and 20. Second, there is the much more specific population that was defined by the sample that was selected: 15-year-olds who visited a mental health service provider in Connecticut during a given year. In Figure 1.1, I offer a graphic that illustrates the concept of a sample of 3 individuals being selected from a population of 10 individuals.

## Inferential and Descriptive Statistics

Why is it important to determine which of these two populations is of interest in this study? Because the consumer of this research must be able to determine how well the results from the sample **generalize** to the larger population. Clearly, depression rates among 15-year-olds who visit mental health service providers in Connecticut may be different from other adolescents. For example, adolescents who visit mental health service providers may, on average, be more depressed than those who do not seek the services of a psychologist. Similarly, adolescents in Connecticut may be more depressed, as a group, than adolescents in California, where the sun shines and Mickey Mouse keeps everyone smiling. Perhaps 15-year-olds, who have to suffer the indignities of beginning high school without yet being able to legally drive, are more depressed than their 16-year-old, driving peers. In short, there are many reasons to suspect that the adolescents who were *not* included in the study may differ in their depression rates from adolescents who were included in the study. When such differences exist, it is difficult to apply the results garnered from a sample to the larger population. In research terminology, the results may not generalize from the sample to the population, particularly if the population is not clearly defined.

So why is generalizability important? To answer this question, I need to introduce the distinction between **descriptive** and **inferential** statistics. Descriptive statistics apply only to the members of a sample or population from which data have been collected. In contrast, inferential statistics refer to the use of sample data to reach some conclusions (i.e., make some inferences) about the characteristics of the larger population that the sample is supposed to represent. Although researchers are sometimes interested in simply describing the characteristics of a sample, for the most part we are much more concerned with what our sample tells us about the population from which the sample was drawn. In the depression study, the researcher does not care so much about the depression levels of the sample *per se*. Rather, the data from the sample is used to reach some conclusions about the depression levels of adolescents *in general*. But to make the leap from sample data to inferences about a population, one must be very clear about whether the sample accurately represents the population. If the sample accurately represents the population, then observations in the sample data should hold true for the population. But if the sample is not truly representative of the population, we cannot be confident that conclusions based on our sample data will apply to the larger population. An important first step in this process is to clearly define the population that the sample is alleged to represent.

## Sampling Issues

There are a number of ways researchers can select samples. One of the most useful, but also the most difficult, is **random sampling.** In statistics, the term *random* has a much more specific meaning than the common usage of the term. It does not mean haphazard. In statistical jargon, *random* means that every member of a defined population has an equal chance of being selected into a sample. The major benefit of random sampling is that any differences between the sample and the population from which the sample was selected will not be systematic. Notice that in the depression study example, the sample differed from the population in important, *systematic* (i.e., nonrandom) ways. For example, the researcher most likely systematically selected adolescents who were more likely to be depressed than the average adolescent because she selected those who had visited mental health service providers. Although randomly selected samples may differ from the larger population in important ways (especially if the sample is small), these differences are due to chance rather than to a systematic bias in the selection process.

**Representative sampling** is another way of selecting cases for a study. With this method, the researcher purposely selects cases so that they will match the larger population on specific characteristics. For example, if I want to conduct a study examining the average annual income of adults in San Francisco, by definition my population is “adults in San Francisco.” This population includes a number of subgroups (e.g., different ethnic and racial groups, men and women, retired adults, disabled adults, parents, single adults, etc.). These different subgroups may be expected to have different incomes. To get an accurate picture of the incomes of the adult population in San Francisco, I may want to select a sample that represents the population well. Therefore, I would try to match the percentages of each group in my sample with those in my population. For example, if 15 percent of the adult population in San Francisco is retired, I would select my sample in a manner that included 15 percent retired adults. Similarly, if 55 percent of the adult population in San Francisco is male, 55 percent of my sample should be male. With random sampling, I may get a sample that looks like my population or I may not. But with representative sampling, I can ensure that my sample looks similar to my population on some important variables. This type of sampling procedure can be costly and time-consuming, but it increases my chances of being able to generalize the results from my sample to the population.

Another common method of selecting samples is called **convenience sampling**. In convenience sampling, the researcher generally selects participants on the basis of proximity, ease of access, and willingness to participate (i.e., convenience). For example, if I want to do a study on the achievement levels of eighth-grade students, I may select a sample of 200 students from the nearest middle school to my office. I might ask the parents of 300 of the eighth-grade students in the school to participate, receive permission from the parents of 220 of the students, and then collect data from the 200 students that show up at school on the day I hand out my survey. This is a convenience sample. Although this method of selecting a sample is clearly less labor-intensive than selecting a random or representative sample, that does not necessarily make it a bad way to select a sample. If my convenience sample does not differ from my population of interest *in ways that influence the outcome of the study*, then it is a perfectly acceptable method of selecting a sample.

To illustrate the importance of the sampling method in research, I offer two examples of problematic sampling methods that have led to faulty conclusions. First, a report by the American Chemical Society (2002) noted that several beach closures in southern California were caused by faulty sampling methods. To test for pollution levels in the ocean, researchers often take a single sample of water and test it. If the pollution levels are too high in the sample, the beach is declared unsafe and is closed. But water conditions change very quickly, and a single sample may not accurately represent the overall pollution levels of water at the beach. More samples, taken at different times during the day and from different areas along the beach, would have produced results that more accurately represented the true pollution levels of the larger area of the beach, and there would have been fewer beach closures.

The second example involves the diagnosis of heart disease in women. For decades, doctors and medical researchers considered heart disease to be a problem only for men. As a result, the largest and most influential studies included only male samples (Doshi, 2015). Two consequences of this failure to include women in the samples that were researched were that doctors were less likely to order testing for heart disease for their female patients than their male patients, and the symptoms of heart disease and cardiac failure among women, which are often different from those of men, were not understood. Many women who could have had their heart disease treated early or their symptoms of cardiac arrest quickly diagnosed died because women were not included in the samples for research on heart disease and heart attacks. The population of people with heart disease clearly includes women, so the samples that included only men were not representative of the population.

## Types of Variables and Scales of Measurement

In social science research, a number of terms are used to describe different types of variables. A **variable** is pretty much anything that can be codified and have more than a single value (e.g., income, gender, age, height, attitudes about school, score on a measure of depression, etc.). A **constant**, in contrast, has only a single score. For example, if every member of a sample is male, the "gender" category is a constant. Types of variables include **quantitative** (or **continuous**) and **qualitative** (or **categorical**). A quantitative variable is one that is scored in such a way that the numbers, or values, indicate some sort of amount. For example, height is a quantitative (or continuous) variable because higher scores on this variable indicate a greater amount of height. In contrast, qualitative variables are those for which the assigned values do not indicate more or less of a certain quality. If I conduct a study to compare the eating habits of people from Maine, New Mexico, and Wyoming, my "state" variable has three values (e.g., 1 = Maine, 2 = New Mexico, 3 = Wyoming). Notice that a value of 3 on this variable is not *more* than a value of 1 or 2—it is simply *different*. The labels represent qualitative differences in location, not quantitative differences. A commonly used qualitative variable in social science research is the **dichotomous variable**. This is a variable that has two different categories (e.g., male and female).

In social science research, there are four different scales of measurement for variables: nominal, ordinal, interval, and ratio. A **nominally scaled variable** has different categories (e.g., male and female, Experimental Group 1, Experimental Group 2, Control Group, etc.). **Ordinal variables** are those whose values are placed in meaningful order, but the distances between the values are not equal. For example, if I wanted to know the 10 richest people in America, in order from the wealthiest to the 10th richest, the wealthiest American would receive a score of 1, the next

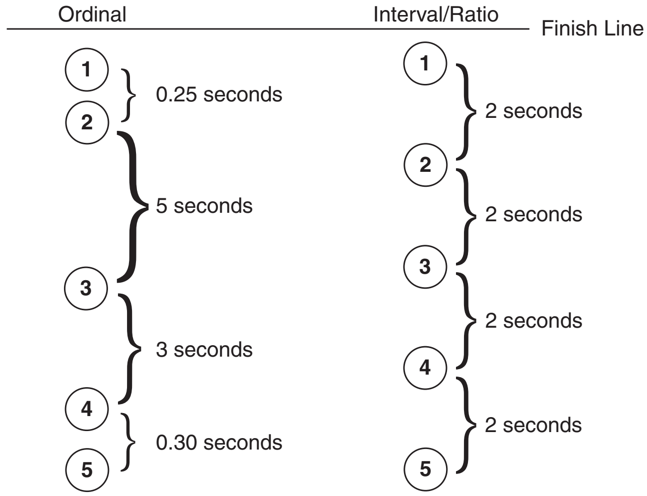


**FIGURE 1.2** Difference between ordinal and interval/ratio scales of measurement.

richest a score of 2, and so on through 10. Notice that while this scoring system tells me where each of the wealthiest 10 Americans stands in relation to the others (e.g., Bill Gates is 1, Michael Bloomberg is 8, etc.), it does not tell me how much *distance* there is between each score. So while I know that the wealthiest American is richer than the second wealthiest, I do not know if he has one dollar more or one billion dollars more. Variables measured with an **interval** scale have values that have order, but they also have equal distances between each unit on the scale. For example, businesses often survey their customers to gain information about how satisfied they are with the service they received. They may be asked to rate the service on a scale from 1 to 10, and this kind of rating scale is an interval scale of measurement. On such surveys, the distance between each number is presumed to be equal, such that a score of 10 would indicate twice as much satisfaction than a score of 5.[1] Variables measured using a **ratio** scale of measurement have the same properties as intervally scaled variables, but they have one additional property: Ratio scales can have a value of zero, while interval scales do not. A great deal of social science research employs measures with no zero value, such as attitude and beliefs surveys (e.g., "On a scale from 1 to 5, how much do you like orange soda?"). Examples of ratio scaled variables include temperatures (e.g., Celsius, Farenheit), income measured in dollars, measures of weight and distance, and many others. Figure 1.2 illustrates a critical difference between ordinal and interval or ratio scales of measurement: Ordinal scales don't provide information about the distance between the units of measurement, but interval and ratio scales do.

One useful way to think about these different kinds variables is in terms of how much information they provide. While nominal variables only provide labels for the different categories of the variable, ordinal variables offer a bit more information by telling us the order of the values. Variables measured using interval scales provide even more information, telling us both the order of the values and the distance between the values. Finally, variables measured with ratio scales add just a little bit more information by including the value of zero in its range of possible values. Figure 1.3 provides a graphic to help you think about the information provided by each of these four types of variables.

## Research Designs

There are a variety of research methods and designs employed by social scientists. Sometimes researchers use an **experimental design**. In this type of research, the experimenter divides the cases in the sample into different groups and then compares the groups on one or more variables of interest. For example, I may want to know whether my newly developed mathematics

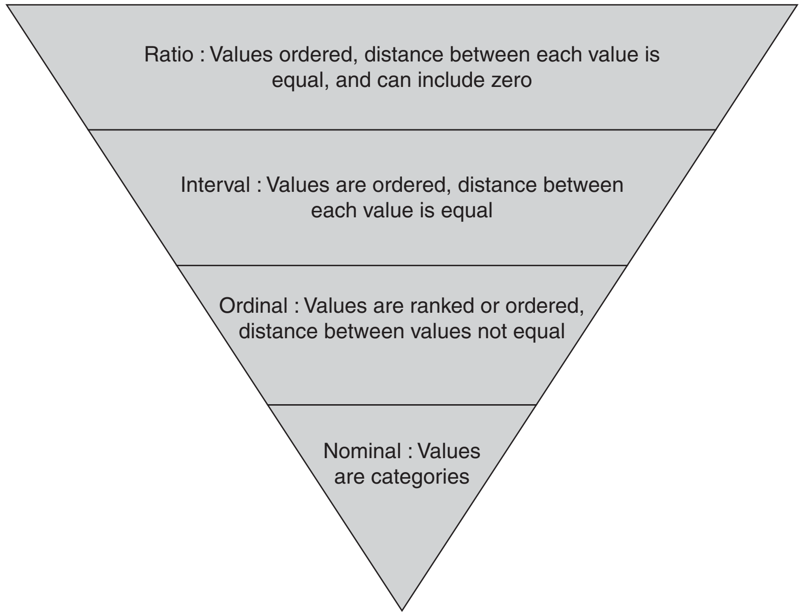


**FIGURE 1.3** Hierarchical arrangement of scales of measurement.

curriculum is better than the old one. I select a sample of 40 students and, using **random assignment**, teach 20 students a lesson using the old curriculum and the other 20 using the new curriculum. Then I test each group to see which group learned more mathematical concepts. This method of random assignment to groups and testing the effects of a particular treatment is known as a **randomized control trial (RCT)** experiment. It has been used for years in medical and laboratory research in the physical and social sciences, and in recent years has been used to great effect in the social sciences outside of laboratory settings. For example, Walton and Cohen (2011) used an RCT design to examine whether a brief intervention could increase first-year college students' feeling of belonging and decrease their feelings of isolation at university. They randomly assigned about half of their participants to receive their intervention while the other half did not, then they compared the two groups and found that those who received the intervention had better psychological and academic outcomes. By assigning students to the two groups using random assignment, it is hoped that any important differences between the two groups will be distributed evenly between the two groups, and that any differences in test scores between the two groups is due to differences in the effectiveness of the two curricula used to teach them. Of course, this may not be true.

A **quasi-experimental research design** is quite similar to an experimental design. Both of these designs typically involve manipulating some variable to see if that variable has an effect on some outcome. In addition, both research designs include some sort of random assignment. The major difference is that in a quasi-experimental design, the research usually occurs outside of the lab, in a naturally occurring setting. In the earlier example used to illustrate the experimental design, I could have had the students come to my research lab and conducted my study in a controlled setting so that I could tightly control all of the conditions and make them identical between the two groups, except for the math curriculum that I used. In a quasi-experimental study, I might find two existing classrooms with 20 students each and ask the teacher in one classroom to use the old math curriculum and the teacher in another classroom to use the new curriculum. Instead of randomly assigning students to these two classrooms (which is difficult to do in a real school), I might randomly select which classroom gets the new curriculum and which one uses the old. I could take steps to try to minimize the differences between the two classrooms (e.g., conduct the study in two different classes of students that are taught by the same teacher, try to find two classrooms that are similar in terms of the gender composition of the students, etc.), but generally speaking it is more difficult to control the conditions in a quasi-experimental design than an experimental design. The benefit of a quasi-experimental design, however, is that it allows the researcher to test the effects of experimental manipulations in more natural, real-world conditions than those found in a research laboratory.

**Correlational research designs** are also a common method of conducting research in the social sciences. In this type of research, participants are not usually randomly assigned to groups. In addition, the researcher typically does not actually manipulate anything. Rather, the researcher simply collects data on several variables and then conducts some statistical analyses to determine how strongly different variables are related to each other. For example, I may be interested in whether employee productivity is related to how much employees sleep (at home, not on the job!). So I select a sample of 100 adult workers, measure their productivity at work, and measure how long each employee sleeps on an average night in a given week. I may find that there is a strong relationship between sleep and productivity. Now, logically, I may want to argue that this makes sense, because a more rested employee will be able to work harder and more efficiently. Although this conclusion makes sense, it is too strong a conclusion to reach based on my correlational data alone. Correlational studies can only tell us whether variables are related to each other—they cannot lead to conclusions about *causality*. After all, it is possible that being more productive at work *causes* longer sleep at home. Getting one's work done may relieve stress and perhaps even allows the worker to sleep in a little longer in the morning, both of which create longer sleep.

Experimental research designs are good because they allow the researcher to isolate specific **independent variables** that may cause variation, or changes, in **dependent variables.** In the example above, I manipulated the independent variable of the mathematics curriculum and was able to reasonably conclude that the type of math curriculum used affected students' scores on the dependent variable, the test scores. The primary drawbacks of experimental designs are that they are often difficult to accomplish in a clean way and they often do not generalize to real-world situations. For example, in my study above, I cannot be sure whether it was the math curricula that influenced the test scores or some other factor, such as a pre-existing difference in the mathematical abilities of my two groups of students, or differences in the teacher styles that had nothing to do with the curricula, but could have influenced the test scores (e.g., the clarity or enthusiasm of the teacher). The strengths of correlational research designs are that they are often easier to conduct than experimental research, they allow for the relatively easy inclusion of many variables, and they allow the researcher to examine many variables simultaneously. The principle drawback of correlational research is that such research does not allow for the careful controls necessary for drawing conclusions about causal associations between variables.

## Making Sense of Distributions and Graphs

Statisticians spend a lot of time talking about **distributions**. A distribution is simply a collection of data, or scores, on a variable. Usually, these scores are arranged in order from smallest to largest and then they can be presented graphically. Because distributions are so important in statistics, I want to give them some attention early on in the book, and show you several examples of different types of distributions and how they are depicted in **graphs**. Note that later in this book there are whole chapters devoted to several of the most commonly used distributions in statistics, including the **normal distribution** (Chapters 4 and 5), ***t* distributions** (Chapter 8 and parts of Chapter 7), ***F* distributions** (Chapters 9, 10, and 11), and **chi-square** distributions (Chapter 14).

Let's begin with a simple example. Suppose that I am conducting a study of voters' attitudes and I select a random sample of 500 voters for my study. One piece of information I might want to know is the political affiliation of the members of my sample. So I ask them if they are Republicans, Democrats, or Independents. I find that 45 percent of my sample list

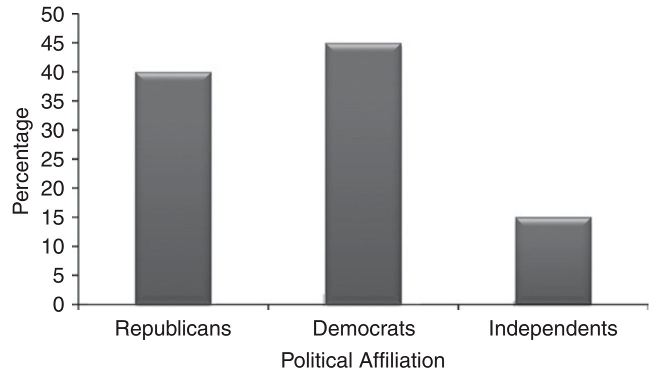


**FIGURE 1.4** Column graph showing the distribution of Republicans, Democrats, and Independents.

themselves as Democrats, 40 percent report being Republicans, and 15 percent identify themselves as Independents. Notice that political affiliation is a nominal, or categorical, variable. Because nominal variables are variables with categories that have no numerical weight, I cannot arrange my scores in this distribution from highest to lowest. The value of being a Republican is not more or less than the value of being a Democrat or an Independent—they are simply different categories. So rather than trying to arrange my data from the lowest to the highest value, I simply leave them as separate categories and report the percentage of the sample that falls into each category.

There are many different ways that I could graph this distribution, including a pie chart, bar graph, column graph, different sized bubbles, and so on. The key to selecting the appropriate graphic is to keep in mind that the purpose of the graph is to make the data easy to understand. For my distribution of political affiliation, I have created two different graphs. Both are fine choices because both of them offer very clear and concise summaries of the distribution and are easy to understand. Figure 1.4 depicts the distribution as a column graph, and Figure 1.5 presents the data in a pie chart. Which graphic is best for these data is a matter of personal preference. As you look at Figure 1.4, notice that the *X* axis (the horizontal one) shows the party affiliations: Democrats, Republicans, and Independents. The *Y* axis (the vertical one) shows the percentage of the sample. You can see the percentages in each group and, just by quickly glancing at the columns, you can see which political affiliation has the highest percentage of this sample and get a quick sense of the differences between the party affiliations in terms of the percentage of the sample. The pie chart in Figure 1.5 shows the same information, but in a slightly more striking and simple manner, I think.

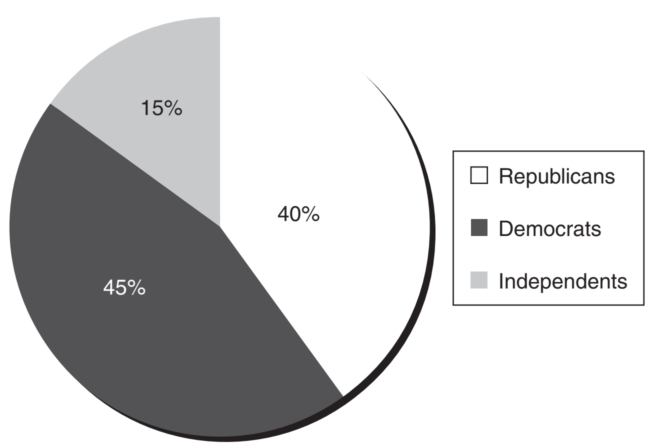


**FIGURE 1.5** Pie chart showing the distribution of Republicans, Democrats, and Independents.

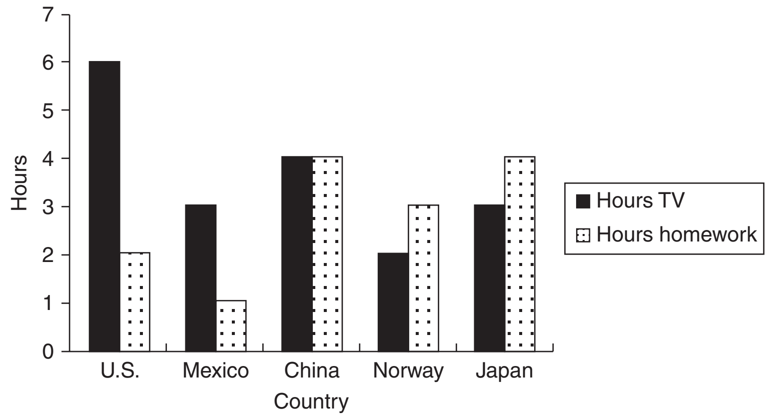


**FIGURE 1.6** Average hours of television viewed and time spent on homework in five countries.

Sometimes, researchers are interested in examining the distributions of more than one variable at a time. For example, suppose I wanted to know about the association between hours spent watching television and hours spent doing homework. I am particularly interested in how this association looks across different countries. So I collect data from samples of high school students in several different countries. Now I have distributions on two different variables across five different countries (the U.S., Mexico, China, Norway, and Japan). To compare these different countries, I decide to calculate the average, or **mean** (see Chapter 2) for each country on each variable. Then I graph these means using a column graph, as shown in Figure 1.6 (note that these data are fictional—I made them up). As this graph clearly shows, the disparity between the average amount of television watched and the average hours of homework completed per day is widest in the U.S. and Mexico and virtually non-existent in China. In Norway and Japan, high school students actually spend more time on homework than they do watching TV, according to my fake data. Notice how easily this complex set of data is summarized in a single graph.

Another common method of graphing a distribution of scores is the line graph, as shown in Figure 1.7. Suppose that I select a random sample of 100 first-year college students who have just completed their first term. I ask them to each tell me the final grades they received in each of their classes and then I calculate a grade point average (GPA) for each of them. Finally, I divide the GPAs into six groups: 1 to 1.4, 1.5 to 1.9, 2.0 to 2.4, 2.5 to 2.9, 3.0 to 3.4, and 3.5 to 4.0. When I count up the number of students in each of these GPA groups and graph these data using a line graph, I get the results presented in Figure 1.7. Notice that along the $X$ axis I have displayed the six different GPA groups. On the $Y$ axis I have the **frequency**, typically denoted by the symbol $f$. So in this graph, the $Y$ axis shows how many students are in each GPA group.

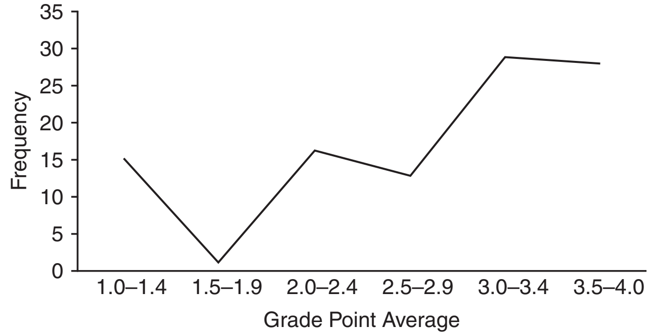


**FIGURE 1.7** Line graph showing frequency of students in different GPA groups.