

# INTRODUCTION TO Statistics

Fundamental Concepts  
and  
Procedures of Data Analysis

Howard M. Reid

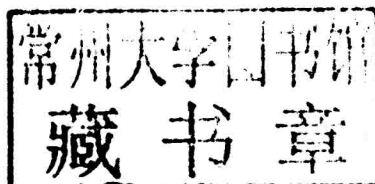


# INTRODUCTION TO STATISTICS

## *Fundamental Concepts and Procedures of Data Analysis*

**Howard M. Reid**

*SUNY Buffalo State*



Los Angeles | London | New Delhi  
Singapore | Washington DC



Los Angeles | London | New Delhi  
Singapore | Washington DC

FOR INFORMATION:

SAGE Publications, Inc.  
2455 Teller Road  
Thousand Oaks, California 91320  
E-mail: [order@sagepub.com](mailto:order@sagepub.com)

SAGE Publications Ltd.  
1 Oliver's Yard  
55 City Road  
London EC1Y 1SP  
United Kingdom

SAGE Publications India Pvt. Ltd.  
B 1/I 1 Mohan Cooperative Industrial Area  
Mathura Road, New Delhi 110 044  
India

SAGE Publications Asia-Pacific Pte. Ltd.  
3 Church Street  
#10-04 Samsung Hub  
Singapore 049483

Acquisitions Editor: Vicki Knight  
Assistant Editor: Kalie Koscielak  
Editorial Assistant: Jessica Miller  
Production Editor: Stephanie Palermi  
Copy Editor: QuADS Prepress (P) Ltd.  
Typesetter: C&M Digitals (P) Ltd.  
Proofreader: Barbara Johnson  
Indexer: Marilyn Augst  
Cover Designer: Candice Harman  
Marketing Manager: Nicole Elliott  
Permissions Editor: Karen Ehrmann

Copyright © 2014 by SAGE Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

Printed in the United States of America

*Library of Congress Cataloging-in-Publication Data*

Reid, Howard M.

Introduction to statistics : fundamental concepts and procedures of data analysis / Howard M. Reid.

pages cm

Includes bibliographical references and index.

ISBN 978-1-4522-7196-5 (paperback : acid-free paper)  
ISBN 978-1-4833-0157-0 (web PDF)

1. Statistics—Textbooks. I. Title.

QA276.12.R45 2013

519.5—dc23 2013006445

This book is printed on acid-free paper.



SFI label applies to text stock

13 14 15 16 17 10 9 8 7 6 5 4 3 2 1

# INTRODUCTION TO STATISTICS

*This book is dedicated to the memory of my father, Malcolm, and my brother, Robert. Both were outgoing, inquisitive, supportive, energetic, intelligent, and hardworking. They were devoted to their families, and each embraced life to the fullest. Tragically, each died much too young. They are greatly missed. And the book is dedicated to my mother, Jeannette, whose love, support, and kindness are a blessing to all who know her.*

# ABOUT THE AUTHOR

**Howard M. Reid** is a professor at State University of New York, Buffalo State. He received his doctorate in experimental psychology from the University of Maine and followed it by postdoctoral study in behavior genetics at The Jackson Laboratory. He is currently conducting studies examining perception of art, is directing a variety of undergraduate student research projects, and is chairing the College senate. He has been an active researcher with broad interests, including published work in operant analysis of behavior, animal models of epilepsy, links between ADHD (attention-deficit/hyperactivity disorder) and laterality, as well as construction of a scale measuring student appreciation of the liberal arts. He has also directed numerous undergraduate research projects. He is the recipient of a number of prestigious awards, including the SUNY Chancellor's Awards in Teaching and Faculty Service and the BSC President's Award for Excellence as an Undergraduate Research Mentor.



# PREFACE

**T**here are many fine introductory statistics books that have been written for undergraduate students. Yet these students often continue to view statistics courses negatively, and many fear they will be unable to master the basic level of understanding that is essential to progress in their majors. The present text is an attempt to rethink what students need to learn in an introductory statistics course and how best to organize the presentation of this material to assist them in acquiring the competence necessary to succeed in their chosen field of study.

Every book is written from some perspective. This book is written from the perspective that a first course in applied statistics is an introduction to a form of critical thinking as much as it is an introduction to a mathematically based discipline. And this book emphasizes what a student will need to remember in future semesters, even years, rather than focusing on a cursory introduction to numerous techniques, many of which will soon be forgotten. Finally, this book is designed to provide a foundation on which students can build as they take further statistics or methodology courses.

As a consequence, while this book covers many of the same topics as in other texts, the presentation, and in some cases the content, differs in significant ways. First, the text is organized to assist students in understanding the logic of statistical procedures and how they are related to each other. And the material is presented so that students gain confidence in their ability to master this subject. Simple concepts and procedures are presented before difficult ones, and concepts build so that by the end of the book readers have gained a clear comprehension of the goals, basic techniques, and the limitations of statistical analysis. Second, an aim of the text is to have the reader not only be exposed to fundamental concepts such as variability but also come to appreciate that these concepts re-occur in a variety of contexts. Repetition and an emphasis on the use of definitional equations enhance students' gaining a deeper understanding. For instance, the reader repeatedly sees that statistical measures of variability are based on deviations, something that can be lost when computational equations are used. Of course, researchers don't commonly employ definitional equations in statistical calculations. (They no longer turn to computational equations either.) Instead, they turn to statistical packages to conduct analyses of their data. Accordingly, a third major thrust is that this text emphasizes the mastery of SPSS through the use of step-by-step directions and numerous figures, all integrated into the chapters. As a result, students see how mastery of SPSS complements their



learning of statistical procedures. Fourth, a concerted effort has been made to integrate the study of statistics with numerous disciplines as well as by including brief historical sections and incorporating relevant quotations. While it is unlikely that every example will resonate with every reader, the goal is to present statistics in a manner that appeals to students coming to the course with diverse backgrounds and interests.

The result is, I believe, a text that will assist students in being able to understand the proper uses, and limitations, of a set of commonly employed statistical procedures. By incorporating progress checks, lists of assumptions, consequences of violating the assumptions, and referral to an overview table of statistical procedures (Appendix K), I have found that my students are better able to match the correct procedure with a research design. This is often a difficult challenge for students. Consequently, one of the goals while writing this text was to present the material in such a way that students would be better able to correctly identify the appropriate procedure to use, not just be able to calculate an answer when told what test to employ.

The book's organization differs significantly from what is commonly seen in other introductory statistics texts. Instead of the almost universal order of presentation which relegates nonparametric tests, seemingly as an afterthought, to a final chapter, the current book is organized so that the procedures that employ nominal data, and which are also the easiest for students to calculate, are presented before the procedures that employ ordinal or interval and ratio data. The resulting presentation is, by design, somewhat repetitious, but it allows for positive transfer among the commonly used procedures that are reviewed. For instance, important statistical concepts, including significance, effect size, and use of post hoc tests, are introduced with the review of the chi-square test of independence. Students see these concepts again when the procedures that employ interval and ratio data are discussed. The goal is for students to experience important concepts repeatedly and to understand how the different statistical procedures are related.

A final note for instructors:

I am fully aware that this book presents a different approach to teaching introductory statistics. And a new approach can be challenging, particularly for those of us who have always viewed the discipline of statistics from a particular perspective. All I can say is that, for me, this system works. My students are learning more material. They are learning more easily. And they are better prepared and more confident going into my department's subsequent methodology course. I know that no single approach to teaching statistics will work for everyone, but I encourage you to consider the difficulties that many students currently have with mastering this subject and whether they would find the approach taken by this text beneficial.

# ACKNOWLEDGMENTS

A list of everyone whose contribution to this book could appropriately be acknowledged would be quite long. However, the contributions of three individuals were so significant that I would be particularly remiss if they were not recognized:

I want to especially thank my wife, Dr. Susan E. Mason. The book was greatly strengthened by her insights on teaching statistics and her amazing talents as an editor, and it might not have come about without her encouragement and support. I will always be grateful to her.

I also want to thank Dr. Robert L. Collins, a very close friend, who first introduced me to the field of statistics and who has taught me so much over many years.

And I would like to thank Ms. Kelly Boos, an extraordinary secretary, who contributed in many ways to creating this finished product.

I would also like to acknowledge the constant support and advice of everyone at Sage, particularly Vicki Knight, Rajasree Ghosh, and Stephanie Palermini. Finally, I would like to acknowledge the numerous contributions of the Sage reviewers. They provided invaluable suggestions for the improvement of this text:

Michael A. Bock, American International College

C. A. Brooks, Bowie State University

Ngoc H. Bui, University of La Verne

Christy Miller Cowan, Lincoln Memorial University

Marilyn Dantico, Arizona State University

Johnathan D. Forbey, Ball State University

Henry Gorman, Austin College

Michele S. Hirsch, St. Francis College

Matthew Jerram, Suffolk University

Joseph G. Johnson, Miami University

Elizabeth Kudadjie-Gyamfi, Long Island University, Brooklyn

Veena S. Kulkarni, Arkansas State University

Sean Laraway, San Jose State University

Kristi Lemm, Western Washington University

Michael A Mangan, University of New Hampshire, Durham

Nanette Silverman, Dowling College

May Takeuchi, University of North Alabama

Thomas Uttaro, College of Staten Island, City University of New York

Amy Wagler, The University of Texas at El Paso

# BRIEF CONTENTS

About the Author	xvii
Preface	xix
Acknowledgments	xxi
Chapter 1. Introduction: Why You Need to Understand Statistical Reasoning	1
<b>PART I. DESCRIPTIVE STATISTICS: SEEING CLEARLY WITH DATA</b>	<b>13</b>
Chapter 2. Describing Nominal and Ordinal Data: The Descriptive Statistics Used With Nominal and Ordinal Data	15
Chapter 3. Describing Interval and Ratio Data—I: An Introduction to the Descriptive Statistics Used With Interval or Ratio Data	31
Chapter 4. Describing Interval and Ratio Data—II: Further Descriptive Statistics Used With Interval or Ratio Data	55
<b>PART II: COMPUTER ASSISTED STATISTICAL ANALYSIS</b>	<b>81</b>
Chapter 5. Using SPSS	83
<b>PART III: INFERENTIAL STATISTICS: THINKING CLEARLY ABOUT DATA</b>	<b>105</b>
Chapter 6. The Logic of Inferential Statistics: The Distinction Between Difference and Association Questions	107
Chapter 7. Finding Differences With Nominal Data—I: The Goodness-of-Fit Chi-Square	125
Chapter 8. Finding Differences With Nominal Data—II: The Chi-Square Test of Independence	161
Chapter 9. Finding Differences With Interval and Ratio Data—I: The One-Sample $z$ Test and the One-Sample $t$ Test	203
Chapter 10. Finding Differences With Interval and Ratio Data—II: The Independent Samples $t$ Test and the Dependent Samples $t$ Test	239

Chapter 11. Finding Differences With Interval and Ratio Data—III: The One-Way Between-Subjects ANOVA	287
Chapter 12. Finding Differences With Interval and Ratio Data—IV: The One-Way Within-Subjects ANOVA	323
Chapter 13. Finding Differences With Interval and Ratio Data—V: The Two-Way Between-Subjects ANOVA	359
Chapter 14. Identifying Associations With Nominal and Interval or Ratio Data: The Phi Correlation, the Pearson $r$ Correlation, and the Point Biserial Correlation	401
Chapter 15. Identifying Associations With Interval and Ratio Data: Linear Regression	451
Chapter 16. Congratulations: The Big Picture and Next Steps: Recapitulation and Final Considerations	473
Appendixes	489
Appendix A: Kruskal–Wallis $H$ Test: Difference Design With One Independent Variable, Two or More Independent Samples, and Ordinal Data	491
Appendix B: Spearman $r$ Correlation: Association Design With Ordinal Data	507
Appendix C: Multiple Regression	517
Appendix D: An Introduction to Power Analysis: Minimum Appropriate Sample Sizes	529
Appendix E: Statistical Symbols Used in This Text	539
Appendix F: Definitional Equations and, Where Appropriate, Their Computational Equation Equivalents	545
Appendix G: Inferential Statistical Procedures and Their Measures of Effect Size	557
Appendix H: Glossary of Terms	559
Appendix I: Answers to Chapter Problems	567
Appendix J: Statistical Tables	575
Appendix K: Overview of Statistical Procedures Covered in This Book	591
References	593
Index	595

# DETAILED CONTENTS

<b>About the Author</b>	<b>xvii</b>
<b>Preface</b>	<b>xvii</b>
<b>Acknowledgments</b>	<b>xxi</b>
<b>Chapter 1. Introduction: Why You Need to Understand Statistical Reasoning</b>	<b>1</b>
A Little History	1
The Two Uses of Statistics	2
Seeing Clearly Is Harder Than It Appears	3
Thinking Clearly Is Also Harder Than It Appears	3
Knowledge of Statistics Can Help You See and Think More Accurately	3
Plan of This Book	4
Goal of This Book	4
Taking a Step at a Time	5
A Brief Review of Mathematics	5
The Need for Math to Follow Rules	6
It Is Time to Learn Some Greek	7
Conclusion	9
Chapter Resources	9
Glossary of Terms	9
Questions: Chapter 1	9
<b>PART I. DESCRIPTIVE STATISTICS: SEEING CLEARLY WITH DATA</b>	<b>13</b>
<b>Chapter 2. Describing Nominal and Ordinal Data: The Descriptive Statistics Used With Nominal and Ordinal Data</b>	<b>15</b>
All Numbers Are Not Equal	15
Sometimes a Number Is Just a Name or Category	16
Sometimes a Number Tells Us the Order of Events	16
Sometimes We Can Add and Subtract Numbers	17
Some Numbers Can Be Multiplied and Divided	17
The Big Picture	18
Seeing Clearly With Nominal Data: An Introduction to Descriptive Statistics	19

It's Time to Begin Learning to See Better	19
Using Graphs and Charts to See Nominal Data More Clearly	20
Summary	23
Seeing Clearly With Ordinal Data: Continuing With Descriptive Statistics	24
First, Second, Third, I See a Pattern	24
Conclusion	26
Chapter Resources	27
Glossary of Terms	27
Questions: Chapter 2	28

<b>Chapter 3. Describing Interval and Ratio Data—I: An Introduction to the Descriptive Statistics Used With Interval or Ratio Data</b>	<b>31</b>
Summary to This Point	38
Measurement of the Variability of Interval and Ratio Data	39
When the Data Are Skewed	40
When the Data Are Normally Distributed	40
Population and Sample: Statisticians' Way of Saying "All" and "Some"	41
Reporting the Calculated Values of the Mean and Standard Deviation	46
Conclusion	49
Chapter Resources	49
Glossary of Terms	49
Questions: Chapter 3	50

<b>Chapter 4. Describing Interval and Ratio Data—II: Further Descriptive Statistics Used With Interval or Ratio Data</b>	<b>55</b>
What You Always Wanted to Know About Intelligence Quotient But No One Told You	58
Descriptive Statistics: The $z$ Score	62
A Little History and a Very Impressive Equation	62
The Standard, Very Important, $z$ Score	63
Who Says You Can't Compare Apples and Oranges?	64
The $z$ Table: Who Would Have Thought That a Few Numbers Could Be So Much Fun?	68
Conclusion	76
Chapter Resources	77
Glossary of Terms	77
Questions: Chapter 4	77

## **PART II. COMPUTER ASSISTED STATISTICAL ANALYSIS** **81**

<b>Chapter 5. Using SPSS</b>	<b>83</b>
Our First Example Using SPSS	84
To Begin SPSS	84
To Find Frequencies	89

To Create a Bar Graph	91
To Create a Pie Chart	92
A Second Example Using SPSS	94
To Begin SPSS Proceed as in the Previous Example	94
To Find Frequencies	97
To Create a Histogram	97
Conclusion	100
Chapter Resources	101
Glossary of Terms	101
Questions: Chapter 5	101
SPSS Problems: Chapter 5	102

### **PART III. INFERENCEAL STATISTICS: THINKING CLEARLY ABOUT DATA** **105**

<b>Chapter 6. The Logic of Inferential Statistics: The Distinction Between Difference and Association Questions</b>	<b>107</b>
Where We Have Been	107
Where We Are Going	107
Decisions, Decisions, Decisions	108
A Little History	110
How Science Works	111
Making an Experimentally Based Decision	113
Whew, That Was a Lot to Remember	114
Probability, Error, and Power	114
Real-World Limitations	117
Conclusion	119
Chapter Resources	119
Glossary of Terms	119
Questions: Chapter 6	120
SPSS Problem: Chapter 6	124

<b>Chapter 7. Finding Differences With Nominal Data—I: The Goodness-of-Fit Chi-Square</b>	<b>125</b>
Goodness-of-Fit Chi-Square	126
Another Example	134
A Final Example of the Goodness-of-Fit Chi-Square	137
Reporting the Results of the Goodness-of-Fit Chi-Square	138
Purpose and Limitations of Using the Goodness-of-Fit Chi-Square Test	139
Assumptions of the Chi-Square Test	140
Conclusion	140
Chapter Resources	140



Glossary of Terms	140
Questions: Chapter 7	141
SPSS	146
Using SPSS With the Goodness-of-Fit Chi-Square	146
To Begin SPSS	146
To Enter Data in SPSS	149
To Conduct a Goodness-of-Fit Chi-Square With SPSS	149
Our Second Example Using SPSS With the Goodness-of-Fit Chi-Square	153
To Begin SPSS	153
To Enter Data in SPSS	154
To Conduct a Goodness-of-Fit Chi-Square With SPSS	154
SPSS Problems: Chapter 7	159
<b>Chapter 8. Finding Differences With Nominal Data—II:</b>	
<b>The Chi-Square Test of Independence</b>	<b>161</b>
Analyzing a Difference Design With Two Variables, Each With at Least Two Outcomes	161
Another Example	167
Reporting the Results of a $2 \times 2$ Chi-Square Test of Independence	171
Reporting the Results of a Chi-Square Test of Independence Larger Than $2 \times 2$	179
A Final Observation	180
Purpose and Limitations of the Chi-Square Test of Independence	180
Conclusion	181
Chapter Resources	182
Glossary of Terms	182
Questions: Chapter 8	182
SPSS	185
Our First Example Using SPSS With the Chi-Square Test of Independence	185
To Begin SPSS	185
To Enter Data in SPSS	188
To Conduct a Chi-Square Test of Independence	188
Our Second Example Using SPSS With the Chi-Square Test of Independence	193
To Begin SPSS	193
To Enter Data in SPSS	196
To Conduct a Chi-Square Test of Independence	196
SPSS Problems: Chapter 8	201
<b>Chapter 9. Finding Differences With Interval and Ratio Data—I:</b>	
<b>The One-Sample <math>z</math> Test and the One-Sample <math>t</math> Test</b>	<b>203</b>
One-Sample $z$ Test	204
Reporting the Results of the One-Sample $z$ Test	216
Confidence Interval for $z$	217