Social Statistics

A Text Using MicroCase®

William Fox

Fourth Edition

WADSWORTH MICROCASE
STUDENT VERSION
SOFTWARE
WIN 95

Social Statistics

Fourth Edition

William Fox
Skidmore College

THOMSON
WADSWORTH

Australia • Canada • Mexico • Singapore • Spain • United Kingdom • United States

Sociology Editor: *Bob Jucha*
Assistant Editor: *Stephanie Monzon*
Editorial Assistant: *Melissa Walter*

Marketing Manager: *Matthew Wright*
Production Service: *Jodi Gleason*
Copy Editor: *Margaret Moore*
Printer: *Transcontinental Printing Inc.*

COPYRIGHT © 2003 Wadsworth Group.
Wadsworth, is an imprint of the Wadsworth Group, a division of Thomson Learning, Inc.
Thomson Learning™ is a trademark used herein under license.

ALL RIGHTS RESERVED. No part of this work covered by the copyright hereon may be reproduced or used in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, Web distribution, or information storage and retrieval systems—without the prior written permission of the publisher.

Printed in Canada.
1 2 3 4 5 6 7 05 04 03 02 01

For permission to use material from this text, contact us by **Web**: http://www.thomsonrights.com
Fax: 1-800-730-2215 **Phone:** 1-800-730-2214

ISBN 0-534-52741-8

For more information, contact
Wadsworth/Thomson Learning
10 Davis Drive
Belmont, CA 94002-3098
USA

For more information about our products, contact us:
Thomson Learning Academic Resource Center
1-800-423-0563
http://www.wadsworth.com

International Headquarters
Thomson Learning
International Division
290 Harbor Drive, 2nd Floor
Stamford, CT 06902-7477
USA

UK/Europe/Middle East/South Africa
Thomson Learning
Berkshire House
168-173 High Holborn
London WC1V 7AA
United Kingdom

Asia
Thomson Learning
60 Albert Complex, #15-01
Singapore 189969

Canada
Nelson Thomson Learning
1120 Birchmount Road
Toronto, Ontario M1K 5G4

To Collette, for all that matters.

About the Author

William Fox has a B.A. in Political Science from the University of Michigan and a Ph.D. in Sociology from Indiana University. He has taught at the University of Iowa, Florida Atlantic University, and Richmond College in London, and is now Professor of Sociology at Skidmore College in upstate New York. His current research interests are in urban folklore and social history. Bill lives with his wife, Collette, in an empty Victorian nest in Saratoga Springs, New York, where he enjoys biographies, Mozart, and roots music.

此为试读,需要完整PDF请访问: www.ertongbook.com

Preface

This fourth edition of *Social Statistics* is prompted by the generous reception given previous editions and is guided by their users' comments and suggestions. Students and faculty alike have forwarded their ideas for conveying statistical ideas more effectively and efficiently. I have tried my best to incorporate the sense and substance of their suggestions in this edition. This edition uses updated and expanded data sets and incorporates a new version of the MicroCase software. I have again integrated data management techniques into statistical analysis and offered examples of statistical write-ups. Students welcomed the supportive, interactive style of earlier editions, and I have continued the same in this edition.

MicroCase

This text uses a student version of MicroCase, a straightforward, no-nonsense piece of statistics software. MicroCase does data analysis and gets the job done efficiently, competently, and intelligently. You will learn statistics, and substantive social science too, with minimal distraction from the computer itself. No inundation with esoteric options. Just learning statistics (or whatever). Just doing data analysis. I have written this text in the same pragmatic spirit as MicroCase.

You will like MicroCase a lot. It can do just about any sort of statistical analysis that you are ever likely to want to carry out. Here are some of the things that MicroCase can do:

- Compute percentages.
- Produce statistical tables.
- Produce pie charts and bar graphs.
- Draw outline maps showing geographic distributions of variables.
- Compute means and medians.

- Compute standard deviations and variances.
- Compute cross-tabulations, chi squares, and measures of association.
- Do *t* tests and analyses of variance.
- Compute regression and correlation coefficients.
- Compute multiple regression and correlation coefficients.
- Create new variables from old ones and manage data in useful ways.

There's no need to worry if you don't know what some of these things are. You will know all of them after finishing this text and accompanying workbook.

MicroCase is available for Windows-compatible computers like IBM, Dell, and Compaq computers. (Sorry, fellow Apple fans, but we have to use Windows computers for MicroCase.) You already own MicroCase. A version called Student MicroCase came on a CD-ROM disc with this text's workbook. Student MicroCase is a slimmed down but still powerful version of the full MicroCase.[1]

Don't worry if you don't know much about computers or even if you have never used one. You don't need to know much about computers themselves to make good, intelligent use of MicroCase to do statistics. You don't have to write computer programs or design computer systems—nothing like that. The accompanying workbook will give you all the information you need.

What to Expect in this Text

Let me explain how I have organized this text. We will learn statistics by beginning with the simplest situations and then moving through increasingly complex analyses. We will start with methods for analyzing only one variable at a time. Then we will take up methods for analyzing a relationship between two variables. That's a little more complex. And finally we will consider techniques for studying relationships among three or more variables at a time. That's really complex. Sensibly enough (at least if you know some Latin), these three situations involve ***univariate, bivariate,*** and ***multivariate analyses,*** respectively. That is how we will proceed in this text. The workbook has exercises that encourage you to apply your knowledge of statistics and carry out computer analyses of real data.

---

[1] Your school may have the full MicroCase if it participates in the MicroCase Curriculum Plan and has a site license. You can find out from your instructor if the full MicroCase is available at your school.

This book covers basic statistical methods. Although you will not be a full-fledged statistician after finishing *Social Statistics*, you will know enough statistics to critically evaluate others' uses of basic statistical techniques and to analyze data yourself. In other words, you will be both an informed consumer and a responsible producer of statistics. You will also be well prepared to study more advanced statistics should you choose to do so. In short, you won't learn everything about statistics from this text, but you will learn a lot.

I have written this text for students in the social sciences and related disciplines such as social work, criminal justice, public administration, and education. I assume that you need to carry out real analyses using real data, so I do not avoid the problems you and I find in the "real world"—missing information, skewed data, outliers, variables that need to be collapsed, and so on. In this text we will face the messiness that makes the social world so challenging to study and so much fun to live in. This text relies heavily on real data from sources like the U.S. Census, *Uniform Crime Reports*, and the General Social Survey—data that full-fledged social scientists analyze in their own research. Thus, we can't avoid facing the real issues and problems that come with real data. Still, there are times when messiness of data can get in the way of learning statistics, so I have not hesitated to make up data to introduce certain statistical techniques when "clean" examples foster learning. We'll use whatever data—real or imaginary—best helps us learn statistics.

I also assume that you need to present your statistical analyses to others (your instructor now, employers or a larger "public" later in your life), so I include guidelines for writing up statistical analyses, formatting tables, and constructing graphs. You will also find examples of analysis write-ups following several chapters. You will probably find these write-ups useful as models for describing your own research results. You don't need to be a professional statistician to communicate statistical findings effectively to others. You will be able to do so after finishing this text.

So this text does not skimp on nitty-gritty details about doing real statistics in the real world. But I am also concerned that you understand statistical principles and reasoning. You need this understanding to know what statistical procedures are right for given situations and to interpret statistical analyses sensibly. Therefore, I will go beyond formulas and procedures to explain why we are doing what we do. You will find almost no computational formulas in this book to facilitate calculations. Computers render such formulas unnecessary. There is, of course, a place for computational formulas. They mostly belong in a Museum of Statistical Antiquities, maybe in a glass case next to slide rules, to remind us how unpleasant and tedious statistics used to be.

But this text does not shy away entirely from formulas and calculations. To the contrary, we will learn to do calculations "by hand" before using the friendly computer to do this mostly routine work for us. Initially while learning data analysis, calculating statistics with our human brains is essential for understanding why statistics work the way they do and what the computer's faster brain does for us. However, we will use definitional rather than computational formulas. Although less convenient and more tedious, definitional formulas describe what statistics are really about and thus are the best way to understand how statistics works. The *Doing Statistics Using MicroCase* workbook accompanying this text invites you to do paper-and-pencil calculations before carrying out computer analyses.

But in the last analysis (literally), the computer exercises in the workbook are even more important than the "by hand" exercises. I have written the workbook for you to use actively with a computer. You should, in fact, read and do most of the workbook while using a computer.

You will find this textbook and workbook conversational in style. That's deliberate. It's the way I teach. My students and I interact in the classroom and, within the limits imposed by print, there is no reason why you and I should not interact via this text. So, I hope you will not take my use of the second-person "you" as an unwarranted familiarity, nor my frequent use of "we" as either an affectation or a ploy. You and I—*we*—are learning statistics together even though you are learning stat for the first time and I am learning it for the umpteenth time.

When you are finished with this book, I hope you will begin a lifetime using statistics effectively. I trust that you will be an intelligent critic of others' applications of statistics and will use statistics intelligently yourself. But beyond these practical goals, I hope you will appreciate the beauty, elegance, and grace of statistics. Like other magnificent achievements of human reason, statistics at its core is an aesthetic endeavor.

What You Need to Bring to This Text

You are probably interested in what you need to know *before* you begin to learn statistics. I assume you know basic arithmetic—adding, subtracting, multiplying, and dividing. You should also be reasonably comfortable with fractions and decimals, squares and square roots, and simple equations.

Whatever else you need to know you will learn along the way. All the better if you have had a course in research methods, but that is not essential. Likewise, I hope you have had several substantive

courses in the social sciences, but those too are not really necessary. Frankly, I assume more motivation and good will on your part than I do prior knowledge.

I trust that you have a good statistics teacher. I think good textbooks matter (that's why I have worked hard on this one), but reading even the best text is less important than learning from a knowledgeable and enthusiastic teacher.

Oh yes, you also need a calculator. An inexpensive one is fine as long as it finds square roots. Some graph paper will also be useful.

Acknowledgments

This fourth edition, like its predecessors, was written at Skidmore College. I am fortunate indeed to teach at a college so committed to the liberal arts, high-quality teaching, and support of its faculty . . . and so cognizant that these three commitments are one.

This text began with my former colleague Richard Rosenfeld, now at the University of Missouri at St. Louis. Rick extolled the virtues of MicroCase, gently nagged me to try it, suggested the need for an accompanying stat text, and then actively supported my project at critical points. Without Rick, I would not have published this book.

Many users of previous editions offered suggestions that were most helpful for revisions. Especially valuable for this fourth edition have been comments and ideas from these colleagues across the country:

Michelle Benson, University at Buffalo, SUNY
Judith G. Calvo, California State University, Fresno
Walter F. Carroll, Bridgewater State College
Douglas Forbes, University of Wisconsin, Marshfield
David J. Gotcher, Abilene Christian University
Daniel A. Powers, University of Texas at Austin
Zhenchao Qian, Arizona State University
David Shafer, University of South Florida
James V. Spickard, University of Redlands
John Tinker, California State University, Fresno
Yvonne Vissing, Salem State College
Susan Waller, University of Central Oklahoma

The Wadsworth folks—Julie Aguilar, Jodi Gleason, Margaret Moore, and Meredith Reitman—did everything they could to aid this book's progress. They variously caught errors, edited, and provided MicroCase displays, all with much care and patience. Most important of all, however, they offered patience. I thank them all for seeing this project through with an optimal blend of professionalism and empathy.

Any errors that remain are mine alone.

Thanks, too, to my students over the years, especially at Skidmore College. True, they couldn't stop me from trying different teaching strategies on them. But they need not have been so good-natured nor have taught me so much, and for that I am grateful. I received assistance from Skidmore College's computer specialists, especially Leo Geoffrion and Beth Dupont. These are the user-friendliest bunch of alpha geeks anywhere. Lauren Sweeney, my terrific assistant, worked on this textbook project with extraordinary diligence and moxie.

I want to thank Elton Jackson, Professor Emeritus of Sociology at Indiana University, for teaching me statistics. On the first day of class, Elton promised a new way of seeing the world. He delivered on that promise. I learned statistics and much more from Elton Jackson.

My appreciation goes to several individuals and institutions who provided the data files upon which most of the workbook exercises are based. First, I obtained all my data sets from the MicroCase Data Archive, which saved me weeks of work. Special thanks goes to Tom W. Smith at the National Opinion Research Center for his continued direction and administration of the General Social Survey. Thanks also goes to the Inter-University Consortium for Political and Social Research (ICPSR) and to The Roper Center for their many years of distribution of the General Social Survey. The data archives maintained by the ICPSR and The Roper Center are an invaluable service to social science researchers.

I am grateful to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., Dr. Frank Yates, F.R.S., and the Longman Group Ltd, London, for permission to reprint parts of Tables III and IV from their book *Statistical Tables for Biological, Agricultural and Medical Research.*

But most of all, I want to thank Emmylou Harris and Collette Fox. Emmylou and all her acoustic pals make great music that goes so well with statistics and just about everything else in life. And Collette makes that life worthwhile.

Any Comments or Suggestions?

I want very much to hear from you if you have any helpful feedback concerning this text. Let me know what you like and don't like about it, what works well and what can be improved, what might be left out and what needs to be added. E-mail me at statprof@skidmore.edu. Or you can write me by snail mail at the following address:

William Fox
Department of Sociology, Anthropology,
and Social Work
Skidmore College
Saratoga Springs, NY 12866

此为试读,需要完整PDF请访问: www.ertongbook.com

One final word before we start learning statistics: I had a great time writing this book. I hope you enjoy it just as much.

Bill Fox
Skidmore College

Contents

About the Author xi
Preface xiii

Part I: Introduction and Univariate Analyses

Chapter 1: Statistics and Variables 3
1.1 Statistics and Data 4
1.2 Overview of Statistics 5
1.3 Samples and Populations 6
1.4 Variables 8
1.5 Levels of Measurement 9
1.6 Mutually Exclusive and Collectively Exhaustive 14
1.7 Continuous and Discrete Variables 15
1.8 What Cases, Variables, and Data Files Look Like 16
1.9 Aggregate Data 18
1.10 Ideas and Thinking 20
1.11 Playing with Data 22
1.12 Summing Up Chapter 1 22

Chapter 2: Frequency and Percentage Distributions 27
2.1 Frequency Distributions 28
2.2 Percentage Distributions 31
2.3 Cumulative Distributions 34
2.4 Creating Sensible and Well-Formatted Tables 36
2.5 Collapsing Variables 38
2.6 Excluding Missing Data 42
2.7 Selecting Subsets of Cases 45
2.8 Pie Charts and Bar Graphs 45
2.9 Outliers 48

2.10 Mapping Ecological Variables 49
2.11 Summing Up Chapter 2 51

Writing Statistics 1: Percentages Distributions, Graphs, and Maps 55

Chapter 3: Averages 61
3.1 Mode 62
3.2 Median 63
3.3 Mean 66
3.4 Properties of the Mean 68
3.5 The Mean for Dichotomous Variables 70
3.6 Which to Use—Mode, Median, or Mean? 72
3.7 Summing Up Chapter 3 75

Chapter 4: Measures of Variation 79
4.1 Variances and Standard Deviations 80
4.2 Shapes of Distributions 87
4.3 Standard Scores (Z-Scores) 89
4.4 Normal Distributions 91
4.5 Sampling Distributions 94
4.6 Confidence Intervals 98
4.7 Some Cautions Using Univariate Statistics 100
4.8 Summing Up Chapter 4 102

Writing Statistics 2: Averages and Standard Deviations 105

Part II: Bivariate Analyses

Chapter 5: Cross-tabulation 109
5.1 Bivariate Frequency Tables 110
5.2 Bivariate Percentage Tables 113
5.3 How to Read Percentage Tables 116
5.4 Positive, Negative, and Curvilinear Relationships 118
5.5 Format Conventions for Bivariate Tables 122
5.6 Stacked Bar Graphs for Bivariate Relationships 125
5.7 A Caution About Bivariate Tables Based on Small Ns 126
5.8 Association Does *Not* Imply Causation 127
5.9 Summing Up Chapter 5 129

Chapter 6: The Chi-Square Test of Statistical Significance 133
6.1 The Logic of Tests of Statistical Significance 134
6.2 The Chi-Square Test 137
6.3 Problems with Expected Frequencies Less Than 5 144
6.4 Statistical Significance Does Not Mean Substantive Significance 145

6.5 Significance Tests on Population Data 147
6.6 Summing Up Chapter 6 147

Chapter 7: Measures of Association for Cross-tabulations 151
7.1 Overview of Measures of Association 152
7.2 Chi-Square–Based Measures for Nominal Variables: V and ϕ 152
7.3 Lambda 155
7.4 Choosing a Nominal Measure of Association 158
7.5 Measures of Association for Ordinal Variables: Gamma 159
7.6 Somers' D_{YX} 164
7.7 Measures of Association: An Overview 168
7.8 Summing Up Chapter 7 169

Writing Statistics 3: Bivariate Cross-tabulations 172

Chapter 8: Comparison of Means and *t* Test 175
8.1 Box-and-Whiskers Diagrams/Differences Between Means 176
8.2 *t* Test for the Difference Between Means 179
8.3 Assumptions and Cautions Concerning *t* Test 186
8.4 One-Tailed and Two-Tailed Tests 188
8.5 Confidence Intervals for Differences Between Means 190
8.6 Summing Up Chapter 8 192

Writing Statistics 4: Comparison of Means and *t* Test 195

Chapter 9: Analysis of Variance 197
9.1 Box-and-Whiskers Diagrams/Differences Among Means 198
9.2 Purpose and Assumptions of Analysis of Variance 201
9.3 The Logic of Analysis of Variance 202
9.4 The ANOVA Table 211
9.5 The Correlation Ratio (E^2) 211
9.6 Two-Way Analysis of Variance (and Beyond) 213
9.7 Three Cautions About Statistically Significant *F* Ratios 214
9.8 Summing Up Chapter 9 215

Writing Statistics 5: Analysis of Variance 218

Chapter 10: Regression and Correlation 221
10.1 Scatterplots 222
10.2 Scatterplots and the Strength of Relationships 225
10.3 Some Limitations of Scatterplots 228
10.4 Regression and Least-Squares Lines 229
10.5 Calculating Regression Coefficients 233
10.6 Correlation Coefficient (r) 234
10.7 r^2 as Proportion of Variation Explained 239

10.8 Correlations Between Dichotomous Variables 241
10.9 Association *Still* Does Not Imply Causation 242
10.10 Linear and Nonlinear Relationships 242
10.11 Test of Significance for a Correlation Coefficient 243
10.12 Correlation Matrix 245
10.13 Summing Up Chapter 10 247

Writing Statistics 6: Regression and Correlation 251

Part III: Multivariate Analyses

Chapter 11: Multivariate Cross-tabulation 255
11.1 The Logic of Causal Relationships 256
11.2 Spurious Relationships 258
11.3 Some Terminology 261
11.4 Examples of Spurious Relationships 262
11.5 Replication 263
11.6 Somewhere Between Explanation and Replication 264
11.7 Specification 265
11.8 Suppressor Variables 266
11.9 Controlling for an Intervening Variable 269
11.10 Partial Gamma 271
11.11 An Overview of Elaboration 272
11.12 Elaboration and Problems of Small Ns 273
11.13 The Relationship of Multivariate Analysis to Experiment Design 274
11.14 Summing Up Chapter 11 276

Writing Statistics 7: Multivariate Cross-tabulation 279

Chapter 12: Multiple Regression and Correlation 281
12.1 Extending the Regression Model 281
12.2 Multiple Correlation Coefficient 287
12.3 Standardized Regression Coefficients (Beta Coefficients) 289
12.4 Significance Tests for Multiple Correlation Coefficients 291
12.5 Regression with Dichotomous and Dummy Variables 292
12.6 Summing Up Chapter 12 295

Writing Statistics 8: Multiple Regression and Correlation 299

Appendix: Statistical Tables

Table 1: The Chi-Square Distribution 303
Table 2: The *t* Distribution 304
Table 3A: The *F* Distribution: p = .05 305

Table 3B: The *F* Distribution: p = .01 306
Table 3C: The *F* Distribution: p = .001 307

Glossary 309

Bibliography 325

Index 329

此为试读,需要完整PDF请访问: www.ertongbook.com