

# SOCIAL SCIENCE STATISTICS

S U S A N E . W R I G H T



---

# Social Science Statistics

---

*Susan E. Wright*

*Drake University*

---

***Allyn and Bacon, Inc.***

***Boston London Sydney Toronto***

## *To Dean and Ehren*

---



Copyright © 1986 by Allyn and Bacon, Inc.  
7 Wells Avenue, Newton, Massachusetts 02159

All rights reserved. No part of the material protected by this copyright notice 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 written permission from the copyright owner.

Series Editor: Judy Shaw  
Production Coordinator: Susan Freese  
Editorial/Production Services: Total Concept Associates  
Text Designer: Judith Ashkenaz  
Cover Coordinator: Linda K. Dickinson

### **Library of Congress Cataloging-in-Publication Data**

Wright, Susan E.  
Social science statistics.  
Bibliography: p.  
Includes index.  
1. Statistics. 2. Social sciences—Statistical methods. I. Title.  
QA276.12.W75 1986 519.5'024301 85-19192  
ISBN 0-205-08671-3

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1 90 89 88 87 86

# **Social Science Statistics**

---

# Preface

---

As statistics become a more widespread and important form of communication in business and government, statistical literacy is becoming a mandatory characteristic for job applicants in many fields. Among the skills in demand is the ability to bridge the communication gap between the computer operator/professional statistician and the traditional manager who has little or no computer/statistical literacy. It is assumed that many students who use this text will need the skills described here in their future research or occupations. They may be asked to determine the appropriate statistics for analysis of data, to interpret the findings of statistical analysis, and to transmit their understandings to others who may or may not be sophisticated in statistical and research terminology. In other words, the premise is not that students using this text will become professional statisticians but that they will apply statistics in their graduate research or their future occupations.

## **PURPOSE**

---

Given the stated assumptions, this text seeks to achieve several goals. It presents the traditional basics of statistics, including the language, concepts, and skills needed for statistical literacy. In addition, it seeks to generate sufficient understanding and appreciation of the role of statistics in the research process to permit the student to make decisions about when statistical analysis is needed, to evaluate theoretical or programmatic problems and determine which statistics are most appropriate, to interpret the statistical results, and finally, to communicate the statistical results and interpretations in terms that are meaningful for further research or for applied decision making.

## **AUDIENCE AND PEDAGOGY**

---

This text is designed for use in an introductory statistics course that emphasizes social science methods and problems. Students of sociology, political science, history, public administration, business management, social psychology, speech communication, and many other areas in which statistics are applied will find the text appropriate. It is assumed that

the typical students in the course will have some math background but will not be mathematically sophisticated. It is also assumed that students will have no familiarity with the specialized language of statistics.

Although it is necessary for the professional statistician to understand formula derivation and other aspects of theoretical statistics, a mathematically sophisticated presentation of statistics often intimidates the typical social science student (particularly if math skills are rusty or anxiety is high) and makes attainment of a lasting, practical understanding of the use of statistics less likely. At the same time a simple how-to-calculate approach is inadequate. Students must be encouraged to understand the reasons behind the uses of the statistics and the concepts of research and statistical analysis that allow them to make informed decisions about the uses and interpretations of statistics.

Given these assumptions about the probable audience for the text, and the perception that a knowledge of statistical basics often is mistakenly taken for granted, causing needless frustration for the student, every effort has been made to explain concepts, terminology, and procedures thoroughly, using a vocabulary that gradually increases in sophistication. Throughout the book, basic information for each statistic is provided in a statistical summary box; then formulas and calculation examples are presented and the results of each statistical analysis are interpreted. For the first examples, the interpretations are simple verbalizations, but they become more complex as additional concepts and statistics are presented.

Because most students can grasp the complexities of a statistic more easily *after* they have some concrete calculative experience, additional discussion of characteristics, uses, and other important issues related to each statistic is delayed until after an initial calculation example has been worked through. Practice problems (with complete answers) are provided after every statistic or group of statistics has been presented to allow students to test and reinforce their skills and their understanding of meanings and procedures. At the end of each chapter, a summary discussion, a summary of the formulas, a self-test, and exercises are presented. The self-test problems are designed to help students determine whether they have grasped the concepts and procedures presented in the chapter; therefore, complete and detailed answers to the self-test problems are given at the end of the chapters. The exercises provide an opportunity for instructors to test the students' understanding of the concepts in each chapter; answers to the exercise problems are provided in the instructor's manual.

A unique feature of the text is that each section is summarized in a decision-making framework in which the characteristics of the statistics and criteria for their use are examined in discussion and illustrated through sample decision-making problems. Additional decision-making problems are offered in the form of practice and exercise problems for

---

students to test their own understandings. The final chapter integrates decision-making techniques for all statistics presented in the text.

## ORGANIZATION

---

The chapters in the text are organized according to several criteria, including the analytical goals of description versus inference, the level of measurement, and the number of variables to be analyzed. The introductory chapter concentrates on the basic concepts and procedures used in statistics. The remainder of the text is divided into several parts, each devoted to statistics that are designed for particular tasks.

Part I familiarizes the student with basic statistical language and procedures through examination of the first category of statistics—those designed for examination of single variables. Students are gradually introduced to such basic skills as formula reading and are encouraged to think about levels of measurement, shapes of distributions, and other factors that may affect the selection of a statistic for description. Chapter 2 examines techniques for organizing data, chapter 3 offers a very brief discussion of graphic presentations of data, and chapter 4 introduces basic statistical notation and procedures with which the student must be familiar before actually examining the first statistical formulas.

Chapters 5 and 6 present several univariate statistics that summarize the characteristics of a distribution in terms of a single value. These chapters are designed to provide an understanding of measures of central tendency and dispersion, but they also provide the first formulas, the first statistical calculation experience, and the first interpretive efforts. For the first statistical formulas, instructions are presented in detail. Some students may find this information repetitive of previously covered concepts, whereas others may find it necessary to review mathematical procedures and concepts before mastering the formulas and calculations (see the math review in appendix A). When students have completed part I, they should have a complete grasp of basic formula reading and computation procedures. Chapter 7 is devoted to a more intensive examination of the meaning and application of one measure—the standard deviation—through analysis of the concepts of normal curve and z-scores. Finally, chapter 8 brings all of these statistics together in a decision-making framework.

Part II of the text presents statistics for describing two-variable relationships, including correlation and prediction measures appropriate for each measurement level. Chapter 9 presents basic concepts, discusses the appropriate data distributions for bivariate analysis, and introduces the interpretations that are used for bivariate statistical analysis. Chapters 10, 11, and 12 present bivariate statistics appropriate for nominal, ordinal, and interval level variables, respectively. For every

statistic presented, calculation and interpretation are discussed, as are the issues that affect the selection of each statistic as the appropriate tool for analysis. Chapter 13 summarizes the bivariate statistics and presents a series of problems in which the student must select the correct statistic for analysis.

Part III examines statistical inference procedures, including estimation and hypothesis testing. In the discussion of inference and the principles that legitimize it, every effort has been made to cut through the inherent complexity of such concepts as probability to offer a clear, yet thorough discussion of the process of statistical inference. Chapter 14 presents the vocabulary of statistical inference and discusses such concepts as probability and sampling distributions. Chapter 15 is devoted to point and interval estimation procedures, and chapter 16 introduces hypothesis testing, using the binomial z-test for illustration. A single hypothesis-testing format is then used in chapters 17 through 19 in the presentation of various difference-of-means tests, analysis of variance, and chi-square tests. Chapter 19 covers inferential tests for ordinal data and inferential application of some of the statistics covered in part II. As in part II, part III provides extensive discussion of the uses, interpretation, and related issues for each statistic. Practice problems, self-test problems, and exercises again permit students to check their understanding and skills.

Part IV (chapter 20) summarizes the entire text and presents two decision-making discussions. The first set of decision-making problems covers only the inferential statistics presented in part III. The second decision-making discussion integrates the univariate and bivariate descriptive statistics with the inferential statistics to help the student become aware of the full array of statistical options and the criteria that must be considered when selecting a statistic.

It is recognized that time may not permit coverage of all of these statistics in a single course. Thus, the text is designed to allow the instructor maximum flexibility in coverage. For instance, chapters 1 through 4 are written at a level that students should find relatively easy to grasp with a minimum of assistance; also, if preferred, only selected statistics from chapters 5 and 6 may be given class attention. Similarly, since the basic concepts on which bivariate description is based are discussed in chapter 9, the instructor may select among the statistics covered in chapters 10 through 13 for class discussion.

---

## ACKNOWLEDGMENTS

I am grateful to the Literary Executor of the late Sir Ronald Fisher, F.R.S., to Dr. Frank Yates, F.R.S., and to Longman Group Ltd., London, for permission to reprint Tables III and IV from their book *Sta-*



---

*tistical Tables for Biological, Agricultural and Medical Research* (6th Edition, 1974).

Many people have contributed to the development of this text, both directly and indirectly. First, I would like to thank my editors, Al Levitt and Judy Shaw, for their patience and support throughout. Also deserving thanks are Jerry Bode (Ball State University), Craig Little (SUNY-Cortland), Howard A. Parker (University of Hartford), Ted Tannenbaum (Glassboro State College), Kent Van Liere (University of Tennessee), Theodore C. Wagenaar (Miami University), Paul L. Wienir (Western Michigan University), James Williams (New Mexico State University), and Ron Wohlstein (Eastern Illinois University), who reviewed the chapters at each stage and made many constructive suggestions. My thanks also go to Connie Westbrook, who proofread the text, and to the many students who, while using early versions of the manuscript as a text, made many helpful suggestions. Finally, I want to thank my family for their support and encouragement, without which the manuscript could not have been completed.

---

---

# Contents

---

---

**Preface** xiii

**1 Basic Concepts in Applied Statistics: An Introduction** 1

Statistics as Tools—Trite but True 1  
The Research Process 2  
Communication with Statistics 3  
From Concepts to Data: Issues of Measurement 4  
Levels of Measurement 7  
Continuous and Discrete Measurement 12  
A Note on Mathematical Skills 12  
Basic Statistical Tasks: Description and Inference 13  
Summary 14  
Self-Test 15  
Exercises 16  
Self-Test Answers 17

***I Univariate Analysis: Description of Single Variables*** 19

**2 Organizing and Presenting Univariate Data** 20

Introduction 20  
Raw Data Tables 20  
Raw Data Distributions 22  
Ordered Distributions 23  
Simple Frequency Distributions 24  
Grouped Frequency Distributions 25  
Practice Problems: Frequency Distributions 32  
Expanding the Frequency Distribution 34  
Practice Problem: Cumulative Distributions 37  
Summary 38  
Self-Test 40  
Exercises 41  
Self-Test Answers 42

**3 Univariate Graphic Presentation** 46

Introduction 46  
Principles of Graphing 46

Graphs for Nominal and Ordinal Variables	47
Graphs for Interval Data	50
Distortion in Graphing	54
Summary	56
Self-Test	57
Exercises	59
Self-Test Answers	60

#### **4 Summation and Other Statistical Procedures 66**

Introduction	66
Data for Demonstration of Procedures	66
The Summation Symbol	67
Summation with Other Operations	69
Summary	73
Self-Test	74
Exercises	76
Self-Test Answers	77

#### **5 Central Tendency and Other Measures of Location 79**

Introduction	79
The Data for Analysis	80
Measures of Central Tendency: An Introduction	82
Central Tendency Measures for Raw and Ordered Distributions	83
Practice Problems: Central Tendency for Raw Distributions	88
Central Tendency Measures for Simple Frequency Distributions	90
Practice Problems: Simple Frequency Distributions	95
Central Tendency Measures for Grouped Frequency Distributions	98
Practice Problems: Grouped Frequency Distributions	101
Weighted Means for Calculation across Groups	104
Practice Problem: Weighted Mean	105
Comparative Characteristics and Relationships Among the Central Tendency Measures	106
Other Measures of Location	114
Practice Problem: Other Location Measures	116
Summary of the Statistical Analysis of the Special Program	117
Summary	118
Self-Test	119
Exercises	121
Self-Test Answers	122

#### **6 Measures of Dispersion 126**

Introduction	126
The Measures of Dispersion	126

---

Measures of Dispersion for Raw and Ordered Distributions	129
Discussion of the Variance and the Standard Deviation	135
Practice Problem: Raw and Ordered Distributions	137
Measures of Dispersion for Simple Frequency Distributions	140
Practice Problem: Simple Frequency Distributions	143
Measures of Dispersion for Grouped Frequency Distributions	145
Practice Problems: Grouped Frequency Distributions	147
The Index of Qualitative Variation	151
Practice Problem: Index of Qualitative Variation	154
Summary of the Statistical Analysis of the Special Program	155
Summary	156
Self-Test	157
Exercises	160
Self-Test Answers	160

**7 The Normal Distribution and z-Scores: Use and Interpretation of the Standard Deviation 165**

Introduction	165
The Normal Distribution	165
z-Scores	170
Fitting a Real Distribution to the Normal Distribution Model	173
Transformed Scores	175
The Standard Normal Distribution	177
Not All Distributions Are Normal	183
Summary	183
Self-Test	184
Exercises	185
Self-Test Answers	186

**8 Decision Making for Univariate Statistics 189**

Introduction	189
Decision-Making Criteria	189
Decision-Making Examples	191
Practice Problems: Decision Making	193
Summary	195
Self-Test	195
Exercises	196
Self-Test Answers	197

***II Bivariate Descriptive Statistics 199***

**9 Bivariate Statistics: An Introduction 200**

Types of Relationships	200
Organizing and Presenting Data: Two-Variable Distributions	202

Practice Problem: Contingency Tables	209
Graphing to Demonstrate Bivariate Relationships	211
Interpretation of Relationship Measures	216
Summary	219
Self-Test	220
Exercises	222
Self-Test Answers	223

## **10 Nominal Level Measures of Relationship 226**

Introduction	226
Percentage Differences	227
Practice Problem: Percentage Differences	230
Lambda Statistics	232
Practice Problems: Symmetrical and Asymmetrical Lambda	237
Phi Statistics	238
Practice Problems: Phi for Nominal Variables	240
Summary	241
Self-Test	242
Exercises	243
Self-Test Answers	244

## **11 Bivariate Statistics for Ordinal Level Data 247**

Introduction	247
Ordinal Distributions	247
Fully Ordered Distributions: Rank Order Correlation	249
Practice Problem: Spearman's $r$	254
Measures of Concordance	256
Measures of Association for Partially Ordered Variables	260
Further Discussion of Tau- $b$ , Gamma, and Somer's $d_{yx}$	268
Practice Problems: Tau, Gamma, and Somer's $d_{yx}$	271
Measures for Dichotomous Ordinal Variables	275
Summary	277
Self-Test	279
Exercises	281
Self-Test Answers	283

## **12 Interval Level Measures of Relationship 287**

Introduction	287
Data for Computation	288
Pearson's Correlation Coefficient	288
Coefficient of Determination	293
Practice Problem: Pearson's $r$ and $r^2$	295
Regression Statistics	297
Correlation and Regression: Another Example	309
Practice Problem: Regression	313

---

Further Interpretations and Relationships of Correlation and Regression Statistics	315
The Correlation Ratio	317
Practice Problem: The Correlation Ratio	326
Summary	328
Self-Test	331
Exercises	332
Self-Test Answers	334
<b>13 Selecting Statistics: Decision Making for Bivariate Problems</b>	<b>338</b>
Introduction	338
Decision-Making Criteria	339
The Decision-Making Chart	339
Decision-Making Examples I	340
Practice Problems: Decision Making I	345
Decision-Making Examples II	348
Practice Problems: Decision Making II	350
Summary	352
Self-Test	353
Exercises	355
Self-Test Answers	358
<b>III Statistical Inference</b>	<b>361</b>
<b>14 An Introduction to Inferential Statistics</b>	<b>362</b>
Inference: Basic Concepts and Processes	362
Probability and Sampling Distributions in Statistical Inference	366
The Standard Normal Distribution	369
A Note on Sample Size	375
Types of Statistical Inference	376
Summary	376
Self-Test	378
Exercises	378
Self-Test Answers	379
<b>15 Statistical Estimation</b>	<b>381</b>
Introduction	381
Point Estimation	381
Error in Point Estimations	384
Practice Problem: Point Estimation	388
Interval Estimation	389
Interval Estimation with Small Samples	395
Practice Problems: Interval Estimation	398
Summary	399
Self-Test	401

Exercises 403  
Self-Test Answers 404

**16 Hypothesis Testing Concept and Procedures: The Binomial Test Example 407**

Introduction 407  
Hypotheses: Establishing the Focus of the Hypothesis Testing Procedure 407  
A Hypothesis Testing Format 408  
The Binomial Test 415  
Hypothesis Format and Computation Examples: Binomial Tests 415  
Additional Examples of the Binomial Hypothesis Test 424  
Further Discussion of Binomial Hypothesis Testing 427  
Comment on Interpretations 430  
Summary 431  
Self-Test 433  
Exercises 433  
Self-Test Answers 434

**17 Testing Hypotheses about Means: One-Sample and Two-Sample Tests 436**

Introduction 436  
One-Sample Tests for the Mean 436  
One-Sample Difference of Means  $z$ -Test 437  
Statements of Attained Significance 444  
One-Sample Difference of Means  $t$ -Test 445  
Further Discussion of the  $z$ -Test Versus the  $t$ -Test of Differences between Means 449  
Practice Problems: One-Sample Difference of Means 450  
Two-Sample Difference of Means Tests 451  
Two-Sample Difference of Means for Independent Samples 452  
Practice Problems: Difference of Means for Two Independent Samples 459  
Difference of Means for Matched Samples 462  
Practice Problems: Difference of Means for Matched Samples 467  
Summary 470  
Self-Test 472  
Exercises 473  
Self-Test Answers 475

**18 Analysis of Variance: Comparing the Means of Three or More Samples 480**

Introduction 480  
One-Way Analysis of Variance 481

---

The <i>F</i> -Test for Variances	482
ANOVA as a Test of Variances	483
Hypothesis Testing Format for ANOVA	485
Discussion of ANOVA	490
Practice Problems: ANOVA	493
ANOVA Posttests: Determining Which Means Differ Significantly	497
Practice Problems: ANOVA Posttests	504
Summary	506
Self-Test	508
Exercises	509
Self-Test Answers	510

## **19 Inference with Ordinal and Nominal Variables 514**

Introduction	514
Chi-Square	514
Practice Problems: Chi-Square	525
Testing Differences between Medians and Other Ordinal Tests	528
Practice Problems: Inferential Tests for Ordinal Variables	534
Inferential Applications of Descriptive Statistics	536
Summary	540
Self-Test	542
Exercises	543
Self-Test Answers	545

## ***IV Summary 549***

## **20 Summary and Decision-Making Framework 550**

Summary of Statistical Concepts	550
Decision Making for Inferential Statistics	552
Practice Problems: Inferential Decision Making	557
Summary of Decision Making for Descriptive and Inferential Statistics	558
Self-Test	563
Exercises	565
Self-Test Answers	567

## **Appendix A: Mathematics Review 569**

## **Appendix B: Tables 578**

## **Appendix C: Glossary 587**

## **Index 599**



# 1

---

---

## Basic Concepts in Applied Statistics *An Introduction*

### STATISTICS AS TOOLS—TRITE BUT TRUE

---

- ■ Statistics! The word alone is enough to strike fear in the hearts of many students who find themselves in a program that requires a course in the subject. Yet most of us deal with statistics daily: We hear reports that the *average* cost of food has risen. The nightly news reports on price *indices* of many types. Students want to know the *mean* when an exam is returned and often ask whether an instructor grades on a *curve*. Americans in general are concerned about the crime *rate*. People who attend horse races are well informed on the *odds*. Despite these common uses of statistics, however, most people remain convinced that they can neither understand nor use statistics.

In reality, statistics are simply tools. As is true of any tool, once we know its purpose, the appropriate techniques for using it, and any necessary safety precautions, we can use a statistic easily and effectively. This understanding of statistics as tools is becoming more vital because as statistics become more common in public communication, they will also become basic tools in business and government. The age of the computer has increased the capability of businesses to accumulate and store massive amounts of information. It also has made possible the quick and accurate analysis of such information through statistical techniques.

The onset of the age of communication, based on computer and statistical analysis, has created a major personnel gap for many organizations. People who are trained to operate computers are readily available, as are professional statisticians; and many people are well