

A P P L I E D

STATISTICS

F O U R T H E D I T I O N

NETER WASSERMAN WHITMORE



FOURTH EDITION



Applied Statistics

JOHN NETER **University of Georgia**
WILLIAM WASSERMAN **Syracuse University**
G. A. WHITMORE **McGill University**

ALLYN AND BACON

Boston London Sydney Toronto

Editor-in-chief, Business: Rich Wohl
Editorial assistant: Dominique Vachon
Editorial-production service: Technical Texts, Inc.
Text designer: Sylvia Dovner
Cover designer: Susan Paradise
Cover administrator: Linda Dickinson
Composition and manufacturing buyer: Louise Richardson



Copyright © 1993, 1988, 1982, 1978
A Division of Simon & Schuster, Inc.
160 Gould Street, Needham Heights, MA 01294

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 of the copyright owner.

In regard to the software disk accompanying this text, the publisher assumes no responsibility for damages, errors, or omissions, without any limitation, which may result from the use of the program or the text.

Library of Congress Cataloging-in-Publication Data
Neter, John.

Applied statistics / John Neter, William Wasserman, G.A. Whitmore.
— 4th ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-205-13478-5

1. Social sciences—Statistical methods. 2. Management—
Statistical methods. 3. Economics—Statistical methods.

4. Statistics. I. Wasserman, William. II. Whitmore, G. A.

III. Title.

HA29.N437 1992

92-17067

519.5—dc20

CIP

ISBN 0-205-13478-5

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1 97 96 95 94 93 92

To Dottie, Cathy, and Lonnie

QUANTITATIVE METHODS AND APPLIED STATISTICS SERIES

ALLYN AND BACON

Barry Render, Consulting Editor

Roy E. Crummer Graduate School of Business, Rollins College

*Readings in Production and Operations Management:
A Productivity Approach*
Ahmadian, Afifi, and Chandler

*Introduction to Operations Research Techniques,
Second Edition*
Daellenbach, George, and McNickle

Managing Production: The Adventure
Fulmer

*POMS: Production and Operations Management
Software*
Gupta, Zanakis, and Mandakovic

Business Forecasting, Fourth Edition
Hanke and Reitsch

*Production and Operations Management: Strategies
and Tactics, Second Edition*
Heizer and Render

Statistical Analysis for Decision Making
Jarrett and Kraft

Management Science, Third Edition
Lee, Moore, and Taylor

*Micro Management Science: Microcomputer
Applications of Management Science, Second Edition*
Lee and Shim

Micro Manager 2.0 Software—IBM
Lee and Shim

Service Operations Management
Murdick, Render, and Russell

Applied Statistics, Fourth Edition
Neter, Wasserman, and Whitmore

*Microcomputer Software for Management Science and
Operations Management, Second Edition*
Render, Stair, Attaran, and Foeller

Introduction to Management Science
Render and Stair

*Quantitative Analysis for Management, Fourth
Edition*
Render and Stair

*Cases and Readings in Management Science, Second
Edition*
Render, Stair, and Greenberg

Topics in Just-In-Time Management
Schniederjans

Statistical Reasoning, Third Edition
Smith

*Production and Operations Management: A Self-
Correcting Approach, Third Edition*
Stair and Render

Introduction to Management Science, Third Edition
Taylor

Brief Business Statistics
Watson, Billingsley, Croft, and Huntsberger

*Statistics for Management and Economics, Fourth
Edition*
Watson, Billingsley, Huntsberger, and Croft

*PC-POM: Software for Production and Operations
Management*
Weiss

*Production and Operations Management, Second
Edition*
Weiss and Gershon

The Application of Regression Analysis
Wittink

Preface

To Our Readers:

Applied Statistics, Fourth Edition, is written for students in basic statistics courses in business, economics, and other social sciences, as well as for people already engaged in these fields who desire an introduction to statistical methods and their application. Our aim is to offer you a balanced presentation of fundamental statistical concepts and methods, along with practical advice on their effective application to real-world problems. The conceptual foundation of each subject is developed carefully up to that level needed for prudent and beneficial use of statistical methods in practice.

Features of Applied Statistics

Clear Explanations. The book is written so that the explanations of important principles and concepts are clear and always contain illustrations by one or more examples drawn from real life. In addition, many case applications are presented so that you can understand statistical concepts and methods in the context of actual use.

Broad Coverage. The topical coverage of the text is broad, but it is in no sense a miscellany of statistical tools. Topics have been selected on the basis of two criteria: (1) their importance in actual applications of statistics, and (2) their contribution to the development and understanding of material presented subsequently in the book.

Clear Structure. Units, chapters, and sections have been organized and sequenced in a way that always keeps the main track of the subject clear to you. Technical notes and secondary observations are presented in *Comments* sections. Topics that are not essential to the main development of statistical ideas are presented in *Optional Topic* sections at the ends of chapters and can be omitted without loss of continuity. Important definitions and formulas are set out in a distinctive manner to aid you in learning and to facilitate ready reference.

Wide Use of Computer Output and Plots. Statistical packages play a major supporting role in applications of statistics and will be encountered by everyone, not just by statistical specialists. We have therefore integrated statistical computer output and plots into the text presentation in every subject area in a routine and natural manner, employing them just as they would be used in practice.

No Advanced Mathematics. We believe that fundamental statistical ideas can be conveyed with only a modest use of mathematics and that this approach in an introductory text leads to a fuller appreciation of statistics than a more mathematical approach. Use of this

book therefore requires only knowledge of college-entrance algebra but not calculus. Mathematical demonstrations are included where they make a significant contribution to your understanding of the subject, but all such demonstrations are segregated so that they do not interfere with the main presentation. Occasionally, a mathematical demonstration that requires calculus is presented in an optional section. These sections are marked “calculus needed” and may be omitted without loss of continuity.

Numerous Problems, Exercises, and Studies. Large numbers of questions and problems are given at the ends of chapters to assist your understanding of concepts and to enable you to obtain experience in applying statistical techniques in practical situations and in interpreting results of statistical investigations. The *Problems* sections contain basic problems and drill questions. The *Exercises* sections present questions dealing with more technical concepts, as well as extensions of ideas developed in the chapters. Finally the *Studies* sections contain major comprehensive problems and case studies. Numerical answers for selected problems (identified by an asterisk in the margin) are given at the end of the book to facilitate immediate checking by you.

A computer will facilitate the working out of many of the problems, exercises, and studies. Those problems for which use of a computer is essential are marked by the logo of a computer. Most of the larger data sets in problems, tables, and figures are contained in files on a computer disk that is available with the text. These data sets are identified by the logo of a floppy disk when first introduced in problems, tables, and figures.

Student Supplements

The following student supplements, specifically keyed to the Fourth Edition of *Applied Statistics*, are available to you.

1. *Study Guide* (prepared by William Sanders, Clarion University). The *Study Guide* is a self-learning study manual to help you in understanding the concepts and techniques presented in the text. It incorporates principles of programmed learning so that correct understanding is immediately reinforced as you work through self-contained units at your own speed. The *Study Guide* emphasizes understanding and calculations are kept to a minimum. Full verbal answers and step-by-step numerical answers are given to enable you to better understand concepts and techniques and to identify causes of any mistakes. The *Study Guide* also contains detailed answers for the starred problems in the text.

2. *Software Workbooks for Statistical Packages*. Software workbooks have been prepared for the Minitab and MYSTAT statistical software systems for use with this text. Each manual introduces you to the software and assists you in using the system for statistical computations and analysis with this text. The two manuals are:

Minitab for Applied Statistics (prepared by Warren Dixon, State University of New York at Plattsburg)

Business MYSTAT with QCTOOLS for Applied Statistics (prepared by Kaye A. de Ruiz, Air Force Academy)

Appendix E of this textbook contains a brief technical manual for Business MYSTAT and QCTOOLS that is a handy reference for these systems.

3. *Data Disk*. Each software workbook contains a data disk for your use. Files of the larger data sets in the tables, figures, and problems, as well as the four large data sets in Appendix D, have been assembled in ASCII format on the disk. You can import these data files readily into Minitab, MYSTAT, and most other statistical packages.

A Final Word

You are about to embark on a journey that will show you how statistical thinking plays an important role in everyday life and how statistical concepts and methods are used in business, economics, and the other social sciences. Read the Introduction first and see how the CEO Compensation case and the Space Shuttle Challenger case illustrate the importance of statistical thinking. Then proceed into the chapters that follow. We are confident that you will enjoy and benefit from the journey.

Acknowledgments

We are greatly indebted to many individuals and organizations who have helped us in the preparation of this book. Our sincere thanks go to all who have provided us with case materials and illustrations that demonstrate the usefulness of statistical methods in business, economics, and the other social sciences. Many persons, including colleagues and reviewers, have made helpful suggestions and otherwise assisted us in this and previous editions, for which we are most grateful. We particularly wish to thank for their help Robert F. Berner, Murray J. Cote, R. V. Erickson, John E. Floyd, Diwakar Gupta, Derek Hart, Edgar Hickman, Shu-Ping Hodgson, Burt S. Holland, Oswald Honkalehto, H. K. Hsieh, Allan Humphrey, Raj Jaganathan, William Meeker, Robert Norland, Jr., Eddy Patuwo, Thomas Pray, Thomas Rothrock, Barbara Ruffle, J. Michael Ryan, Al Schainblatt, Kenneth C. Schneider, Randolph Shen, Brian Smith, Ehsan S. Soofi, Erland Sorensen, Albert Teitlebaum, Stephen Vardeman, Dean Wichern, and Morty Yalovsky.

We are also indebted to our respective universities for providing an environment in which the undertaking could be brought to fruition. A number of people were most helpful in working with us in preparing and checking the manuscript. We wish to express our appreciation to Karine Benzacar, Sandra June-Hatfield, Linda Keith, Karen Robertson, and Erika Whitmore for their valuable assistance. The staffs of Technical Texts and Allyn and Bacon helped us in many ways. Of course, we thank our families for their patience and encouragement while we completed this book.

Contents

PREFACE xv

INTRODUCTION I

Statistics in Daily Life	1
What is Statistics?	3
Role of Statistics	4
Statistical Computer Packages	4
Cited References	5

UNIT ONE

Data 7

I DATA ACQUISITION AND MANAGEMENT 9

1.1	Data Sets	9
1.2	Data Sources	11
1.3	Experimental and Observational Studies	13
1.4	Data Acquisition	15
1.5	Deficiencies in Data Sets	19
1.6	Computerized Data Management	21
	Cited Reference	23
	Problems	23
	Studies	27

2 DATA PATTERNS 29

2.1	Simple Displays of Quantitative Data	29
2.2	Frequency Distributions	33
2.3	Displays of Qualitative Data	43
2.4	Displays of Bivariate Data	46
2.5	Displays of Multivariate Data	51
2.6	Data Models and Residual Analysis	52
2.7	Misleading Graphic Presentations	54
	Cited References	56
	Problems	56
	Exercise	66
	Studies	66

3 DATA SUMMARY MEASURES 69

- 3.1 Descriptive Statistics 69
- 3.2 Measures of Position 71
- 3.3 Measures of Variability 79
- 3.4 Measures of Skewness 85
- 3.5 Standardized Observations 88
- 3.6 Box Plots 90
- 3.7 Optional Topic—Calculation of Summary Measures from Frequency Distributions 92
- 3.8 Optional Topic—Geometric and Harmonic Means 96
 - Problems 98
 - Exercises 104
 - Studies 105

UNIT TWO

Probability 107

4 BASIC PROBABILITY CONCEPTS 109

- 4.1 Random Trials, Sample Spaces, and Events 109
- 4.2 Probability 115
- 4.3 Probability Distributions 119
- 4.4 Basic Probability Theorems 124
- 4.5 Relationships Between Variables 129
 - Problems 133
 - Exercises 140
 - Studies 141

5 RANDOM VARIABLES 143

- 5.1 Basic Concepts 143
- 5.2 Discrete Random Variables 144
- 5.3 Expected Value and Variance 149
- 5.4 Standardized Random Variables 151
- 5.5 Continuous Random Variables 153
 - Problems 155
 - Exercises 160
 - Studies 160

6 ADDITIONAL TOPICS IN PROBABILITY 163

- 6.1 Bayes' Theorem 163
- 6.2 Functions of Random Variables 167
- 6.3 Covariance and Correlation 173
- 6.4 Optional Topic—Calculations for Continuous Random Variables (Calculus Needed) 178
 - Problems 180

Exercises	184
Studies	185

7 COMMON PROBABILITY DISTRIBUTIONS: DISCRETE RANDOM VARIABLES 187

7.1	Discrete Uniform Probability Distributions	187
7.2	Binomial Probability Distributions	189
7.3	Poisson Probability Distributions	196
7.4	Optional Topic—Hypergeometric Probability Distributions	199
	Problems	202
	Exercises	205
	Studies	206

8 COMMON PROBABILITY DISTRIBUTIONS: CONTINUOUS RANDOM VARIABLES 209

8.1	Continuous Uniform Probability Distributions	209
8.2	Normal Probability Distributions	211
8.3	Exponential Probability Distributions	221
8.4	Optional Topic—Relation Between Exponential and Poisson Distributions	225
	Problems	228
	Exercises	231
	Studies	231

UNIT THREE

Estimation and Testing I 233

9 STATISTICAL SAMPLING 235

9.1	Populations	235
9.2	Censuses and Samples	237
9.3	Simple Random Sampling from a Finite Population	241
9.4	Simple Random Sampling from an Infinite Population	245
9.5	Sample Statistics and Population Parameters	248
	Problems	251
	Exercises	257
	Studies	258

10 POINT ESTIMATION AND SAMPLING DISTRIBUTION OF \bar{X} 259

10.1	Point Estimation	259
10.2	Experimental Study of \bar{X}	261
10.3	Theoretical Results About \bar{X}	265
10.4	Using the Normal Approximation for \bar{X}	269
10.5	Exact Sampling Distribution of \bar{X}	271

10.6	Criteria for Choosing Point Estimator	272
10.7	Optional Topic—Maximum Likelihood Estimation	277
	Problems	281
	Exercises	285
	Studies	286

I I INTERVAL ESTIMATION OF POPULATION MEAN 289

11.1	Interval Estimation	289
11.2	Confidence Interval for μ —Large Sample	293
11.3	Confidence Interval for μ —Small Sample	298
11.4	Confidence Interval for Population Total	301
11.5	Planning of Sample Size	302
11.6	Optional Topic—One-Sided Confidence Intervals for μ	305
11.7	Optional Topic—Prediction Interval for New Observation	306
	Problems	309
	Exercises	313
	Studies	313

I 2 TESTS FOR POPULATION MEAN 315

12.1	Statistical Tests	315
12.2	Construction of Statistical Decision Rules	322
12.3	<i>P</i> -Values	330
12.4	Small-Sample Tests	335
12.5	Optional Topic—Relation Between Confidence Intervals and Tests for μ	338
12.6	Optional Topic—Power of Test	339
12.7	Optional Topic—Planning of Sample Size	348
	Problems	353
	Exercises	361
	Studies	361

I 3 INFERENCES FOR POPULATION PROPORTION 363

13.1	Population and Sample Proportions	363
13.2	Sampling Distributions of f and \bar{p}	364
13.3	Estimation of Population Proportion	371
13.4	Tests for Population Proportion	374
13.5	Optional Topic—Planning Sample Size for Test About p	376
13.6	Optional Topic—One-Sided Confidence Intervals for p	381
13.7	Optional Topic—Inferences for p with Small Sample	382
	Problems	384
	Exercises	390
	Studies	390

I 4 COMPARISONS OF TWO POPULATIONS AND OTHER INFERENCES 393

14.1	Comparative Studies	393
------	---------------------	-----

- 14.2 Inferences About $\mu_2 - \mu_1$ —Independent Samples and Completely Randomized Designs 397
- 14.3 Inferences About $\mu_2 - \mu_1$ —Matched Samples and Randomized Block Designs 404
- 14.4 Inferences for $p_2 - p_1$ —Independent Samples and Completely Randomized Designs 408
- 14.5 Optional Topic—Inferences About Population Variance and Ratio of Two Population Variances 413
 - Problems 421
 - Exercises 430
 - Studies 431

UNIT FOUR

Estimation and Testing II 433

15 NONPARAMETRIC PROCEDURES 435

- 15.1 Inferences About Population Median 435
- 15.2 Inferences About $\eta_2 - \eta_1$ —Independent Samples and Completely Randomized Designs 443
- 15.3 Inferences About η_D —Matched Samples and Randomized Block Designs 451
- 15.4 Optional Topic—Test for Randomness 459
- 15.5 Optional Topic—Jackknife Estimation 463
 - Problems 467
 - Exercises 475
 - Studies 475

16 GOODNESS OF FIT 477

- 16.1 Probability Plots 477
- 16.2 Chi-Square Test 482
- 16.3 Test Involving Discrete Population Distribution 488
- 16.4 Test Involving Continuous Population Distribution 489
- 16.5 Optional Topic—Confidence Band and Test for Cumulative Probability Function 492
 - Problems 496
 - Exercises 500
 - Studies 501

17 MULTINOMIAL POPULATIONS 503

- 17.1 Nature of Multinomial Populations 503
- 17.2 Multinomial Probability Distributions 504
- 17.3 Inferences Concerning the Parameters p_i 505
- 17.4 Bivariate Multinomial Populations 510
- 17.5 Comparisons of Several Multinomial Populations 517
 - Problems 521

Exercises	526
Studies	527

UNIT FIVE

Linear Statistical Models 529

18 SIMPLE LINEAR REGRESSION—I 531

18.1	Relation Between Dependent and Independent Variables	531
18.2	Simple Linear Regression Model	534
18.3	Point Estimation of β_0 and β_1	540
18.4	Point Estimation of Mean Response	544
18.5	Residuals	546
18.6	Analysis of Variance	548
18.7	Coefficients of Simple Determination and Correlation	555
18.8	Computer Output	558
	Problems	559
	Exercises	563
	Studies	564

19 SIMPLE LINEAR REGRESSION—II 565

19.1	Confidence Interval for Mean Response	565
19.2	Prediction Interval for New Response	569
19.3	Inferences Concerning β_1	574
19.4	Inferences Concerning β_0	577
19.5	F Test of $\beta_1 = 0$	579
19.6	Evaluation of Appropriateness of Regression Model	581
19.7	Remedial Actions When Regression Model Not Appropriate	586
19.8	Practical Considerations in Using Regression Analysis	587
	Problems	591
	Exercises	597
	Studies	597

20 MULTIPLE REGRESSION 599

20.1	Multiple Regression Models	599
20.2	Some Basic Results for Multiple Regression Models	601
20.3	Residual Analysis	608
20.4	F Test for Regression Relation	610
20.5	Inferences Concerning Individual Regression Coefficients	611
20.6	Confidence Interval for $E\{Y_h\}$	613
20.7	Prediction Interval for $Y_{h(\text{new})}$	614
20.8	Indicator Variables	614
20.9	Modeling Curvilinear Relationships	620
20.10	Modeling Interaction Effects	623

20.11	Optional Topic—Multicollinearity	627
20.12	Optional Topic—Selecting the Regression Model	632
20.13	Optional Topic— F Test for Several Regression Coefficients	634
	Problems	638
	Exercises	649
	Studies	649

21 ANALYSIS OF VARIANCE 651

21.1	ANOVA Models	651
21.2	Single-Factor ANOVA Model	653
21.3	Test for Equality of Treatment Means	655
21.4	Analysis of Treatment Effects	664
21.5	Residual Analysis	668
21.6	Other Types of ANOVA Studies	671
21.7	Optional Topic—Correspondence Between ANOVA and Regression	673
	Problems	677
	Exercises	683
	Studies	683

UNIT SIX

Quality Improvement and Sample Surveys 685

22 QUALITY CONTROL AND QUALITY IMPROVEMENT 687

22.1	Quality and Role of Statistics	687
22.2	Product and Process Design	690
22.3	Assessing Conformance—Control Charts	696
22.4	Assessing Conformance—Acceptance Sampling	710
22.5	Assessing Performance—Reliability	714
	Cited References	718
	Problems	718
	Exercises	726
	Studies	727

23 SURVEY SAMPLING 729

23.1	Simple Random Sampling from Finite Populations	729
23.2	Stratified Sampling	733
23.3	Probability-Proportional-to-Size Sampling	739
23.4	Cluster Sampling	741
23.5	Systematic Sampling	743
	Problems	744
	Exercises	749
	Studies	750

UNIT SEVEN

Time Series Analysis and Index Numbers 751

24 TIME SERIES AND FORECASTING: DESCRIPTIVE METHODS 753

- 24.1 Time Series 753
- 24.2 Smoothing 754
- 24.3 Time Series Components 757
- 24.4 Trend Component 760
- 24.5 Cyclical Component 766
- 24.6 Seasonal Component 772
- 24.7 Additive Time Series Model 778
 - Problems 781
 - Exercises 789
 - Studies 790

25 TIME SERIES AND FORECASTING: REGRESSION METHODS AND EXPONENTIAL SMOOTHING 793

- 25.1 Regression Models with Independent Error Terms 793
- 25.2 Regression Models with Autocorrelated Error Terms 799
- 25.3 Exponential Smoothing for Stationary Time Series 807
- 25.4 Exponential Smoothing for Time Series with Trend Component 812
- 25.5 Exponential Smoothing for Time Series with Trend and Seasonal Components 816
 - Problems 820
 - Exercises 829
 - Studies 829

26 PRICE AND QUANTITY INDEXES 831

- 26.1 Percent Relatives and Link Relatives 831
- 26.2 Price Indexes by Method of Weighted Aggregates 833
- 26.3 Price Indexes by Method of Weighted Average of Relatives 835
- 26.4 Considerations in Compiling Index Series 837
- 26.5 Uses of Price Indexes 840
- 26.6 Quantity Indexes 842
 - Problems 843
 - Exercises 848
 - Studies 848

UNIT EIGHT

Bayesian Decision Making 851

27 BAYESIAN DECISION MAKING WITHOUT SAMPLE INFORMATION 853

- 27.1 Decision Problems 853

27.2	Decision Making Under Certainty	856
27.3	Decision Making Under Uncertainty	857
27.4	Bayesian Decision Making with Prior Information Only	861
27.5	Multistage Decision Problems	866
27.6	Optional Topic—Utility	868
	Problems	874
	Exercises	880
	Studies	880

28 BAYESIAN DECISION MAKING WITH SAMPLE INFORMATION 883

28.1	Nature of Decision Problem	883
28.2	Optimal Decision Rule for Given Sample Size	885
28.3	Determination of Optimal Sample Size	890
28.4	Determining Bayes Act by Revision of Prior Probabilities	894
	Problems	896
	Exercises	900
	Studies	900

APPENDIX A: MATHEMATICAL REVIEW 903

APPENDIX B: χ^2 , t , and F DISTRIBUTIONS 909

APPENDIX C: TABLES 921

APPENDIX D: DATA SETS 941

APPENDIX E: BUSINESS MYSTAT MANUAL 953

APPENDIX F: ANSWERS TO SELECTED PROBLEMS 971

INDEX 981