STATISTICS
IN PRACTICE

1 .25 2 .40 3 .20 4 10

# STATISTICS IN PRACTICE

Ernest A. Blaisdell

Elizabethtown College Elizabethtown, Pennsylvania



Saunders College Publishing
Harcourt Brace Jovanovich College Publishers
Fort Worth Philadelphia San Diego New York Orlando Austin
San Antonio Toronto Montreal London Sydney Tokyo

#### Copyright © 1993 by Saunders College Publishing

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

Requests for permission to make copies of any part of the work should be mailed to Permissions Department, Harcourt Brace Jovanovich, Publishers, 8th Floor, Orlando, Florida 32887.

Text Typeface: Times Roman

Compositor: Progressive Typographers, Inc.

Acquisitions Editor: Robert Stern

Developmental Editor: Donald Gecewicz

Managing Editor: Carol Field Project Editor: Kimberly A. LoDico

Copy Editor: Zanae Rodrigo

Manager of Art and Design: Carol Bleistine

Art Director: Christine Schueler Art Assistant: Caroline McGowan Text Designer: Alan Wendt

Cover Designer: Lawrence R. Didona

Text Artwork: Grafacon, Inc.

Photo Researcher: Dena Digilio-Betz Director of EDP: Tim Frelick Production Manager: Bob Butler Marketing Manager: Monica Wilson

Cover Credit: © 1992 Bishop/Phototake, Inc. NYC

Printed in the United States of America

Statistics in Practice

ISBN: 0-03-032229-4

Library of Congress Catalog Card Number: 92-050412

2345 49 987654321

Book printed on **recycled** paper containing a minimum of 10% post-consumer waste and 50% pre-consumer waste. Cover printed on **recycled** paper containing a minimum of 10% post-consumer waste and 40% pre-consumer waste.

### To my loving Mother, Thelma

With fondness, Alf and I recall those many trips to and fro, especially when you multiplied by n, more or less.

Now the final journey home remains.

# **Preface**

"The time may not be very remote when it will be understood that for complete initiation as an efficient citizen of one of the great complex world states that are now developing, it is necessary to be able to compute, to think in averages and maxima and minima, as it is now to be able to read and to write."

Though written more than 60 years ago, H. G. Wells' passage from *Mankind in the Making* seems particularly relevant in today's electronic age of global communication. Understanding the uses of statistics and its role in assimilating information contained in reports, scientific journals, political coverage, or even the daily newspaper is a necessary part of modern education. I have attempted to bring my training as a statistician and my years of teaching experience to the shaping of a clear and concisely written text. In addition, there is always the difficulty of persuading students, many of whom have been conditioned into math anxiety, that they might actually enjoy statistics, that statistical concepts are worth learning, and that statistics derives from real problems in the real world. My goal has been to provide a presentation that is pedagogically and mathematically sound, yet sufficiently gentle to minimize math anxiety.

## Content Features

- To make data as vivid as possible, traditional methods of summarizing data are blended in Chapter 2 with more recently developed **data analysis techniques** such as dotplots, stem-and-leaf displays, 5-number summaries, and boxplots.
- Regression analysis, frequently used in many disciplines, can be introduced
  much earlier than it usually is in the traditional course. Chapter 3, therefore, is a
  concise introduction to the descriptive aspects of correlation and regression. It
  is written so that an instructor can vary its placement within the course syllabus.
  Furthermore, if the instructor wishes, this coverage can be complemented later

- with a detailed discussion in Chapter 14 of inferential methods in regression analysis.
- Included in Chapter 4, which introduces probability, is a concise section on elementary counting techniques. Students often find this topic difficult because they tend to approach each problem as either "a permutation or a combination." I believe a greater understanding can be achieved by de-emphasizing permutation formulas and stressing the versatility of the multiplication rule.
- The critical topics of **confidence intervals** and **hypothesis testing** merit separate chapters. They are introduced in Chapters 9 and 10, respectively, and in Chapter 11 they are jointly used to discuss two-sample inferences.
- **P-values** are prominent in the research literature of virtually all disciplines. Consequently, after the introduction of hypothesis testing in Chapter 10, the reader is frequently exposed to the use of **P-values** throughout the remainder of the book. This is done, however, only after the student has had adequate opportunity to comprehend the basic concepts of hypothesis testing and rejection regions.

## **Exercises and Examples**

- There is an abundant quantity of interesting exercises (1,633) and illustrations based on real-life situations and cited sources from a wide spectrum of disciplines. They are stated concisely, without burying the reader in verbiage.
- The exercises have been carefully selected and constructed to ensure that they meaningfully contribute to the learning process and enhance an appreciation of how statistics intermingles with our daily lives. The order of presentation progresses from mastering the basics to practical applications. Data used in previous exercise sets are always reproduced when used in subsequent applications. To serve the needs of instructors, odd-numbered problems are frequently paired with even-numbered problems. Answers are given in the book for all review exercises and for all odd-numbered end-of-section exercises.
- Worked examples are set up so that students can "walk through" them step by step. This approach helps the student understand the rationale of each statistical procedure. Procedure boxes that recap in a step-by-step manner what students should understand about a given process are liberally provided.

## Pedagogical Features

- Each chapter opens with a preview, "Looking Ahead." The opening photograph and accompanying caption set the theme for the chapter and give the students a foretaste of what they will explore in the pages that follow.
- To enhance the book's appearance and reader friendliness, liberal use is made of photos, marginal notes of interest, newspaper and magazine excerpts, and historical highlights of prominent mathematicians and statisticians.

- To help students master and retain the concepts in each chapter, the end-of-chapter material includes a summary, "Looking Back," a "Key Words" list, a "MINITAB Commands" list, and "Review Exercises," a set of comprehensive problems.
- Important formulas, definitions, procedures, tips, and computer commands are highlighted so that the reader can give them priority on first reading and, later, during review.
- A removable, detailed formula card and separate tables card have been bound into the book for possible use during examinations.
- For a ready reference, the normal, chi-square, and t tables are reproduced on the inside covers.

#### Use of MINITAB

This book is unique in its abundant use of the statistical computer package MINITAB. Instead of being used as a mere appendage to each chapter, MINITAB is woven throughout the text, with each command explained when it is first used in an application. This integration emphasizes the computer's role as a practical tool for relieving much of the drudgery associated with data sets, allowing the user more time to focus on other aspects of the analysis such as selecting a proper procedure, describing and interpreting data, and displaying the results. MINITAB's use also enhances the comprehension of many statistical concepts and techniques presented in a first course.

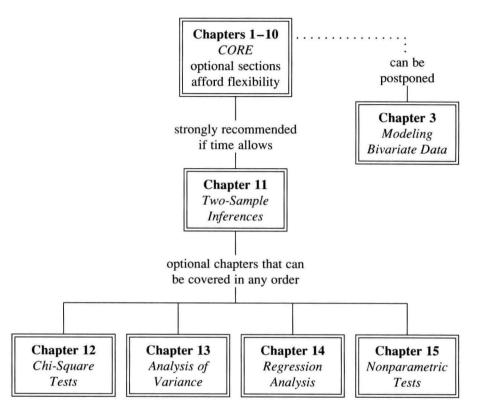
Although there are many statistical software packages currently available, I chose MINITAB because of its wide acceptance in educational instruction and its extensive use around the world in business and government. It is currently used at more than 2,000 colleges and universities and by 70% of Fortune 500 companies. Of equal importance, MINITAB can be learned quickly and easily, providing students with a powerful data analysis system that can be used in other courses and in their professional careers.

Because not all students have access to MINITAB, I have incorporated it into the text in a manner that affords an instructor considerable flexibility concerning its usage in the course. An instructor can choose any of the following options, each of which has been used by the author during the class-testing of this book.

- Active Computer Usage. Sufficient MINITAB instruction is provided so that students can write their own commands for statistical analyses. The book contains a total of 240 MINITAB assignments. They are flagged with the symbol and are placed at the end of an exercise set. MINITAB coverage is so extensive that the usual MINITAB supplement manual is unnecessary.
- Passive Computer Usage. Students can be instructed to examine only the output of the MINITAB exhibits and to just look over (or ignore) the commands used to generate the results.
- No Computer Usage. An instructor may prefer to have the class skip entirely
  the MINITAB exhibits. Implementation of this option is facilitated by the fact
  that all MINITAB output is prominently highlighted.

## Flexibility in Topical Coverage

The book is designed so that an instructor has a great deal of flexibility in topical coverage. The diagram below displays several possibilities.



Chapters 1 through 10 form the core of the text, and Chapter 11 is strongly recommended if time allows. For instructors who want to spend less time covering these chapters, optional sections 6.4, The Hypergeometric Probability Distribution; 6.5, The Poisson Probability Distribution; and 11.3, Small-Sample Inferences for Two Means: Independent Samples and Unequal Variances, can be excluded. Additional time can be gained by also excluding sections 9.7, Chi-Square Probability Distributions; 9.8, Confidence Interval for a Variance; 10.5, Hypothesis Test for a Variance; 11.6, F Probability Distributions, and 11.7, Inferences for Two Variances.

# Supplements

The following supplements have been prepared to enhance the use of this book. They are available, free of charge, to instructors who adopt the text.

Student Solutions Manual. This supplement was prepared by Ronald L. Shubert of Elizabethtown College and is available to students for purchase. It con-

- tains detailed solutions for all review exercises and all odd-numbered, end-ofsection exercises.
- Instructor's Manual. Also prepared by Ronald L. Shubert, this manual contains detailed solutions to all exercises, and a sample course syllabus with helpful suggestions.
- ExaMaster<sup>TM</sup> Computerized Test Bank. Available for IBM-compatible and Macintosh computers, this test bank contains more than 1,200 questions, each written especially for this text. A virtually unlimited number of tests can be custom designed by an instructor. Tests can contain a mixture of multiple-choice and free-response questions sorted by several different categories. An instructor can also add and edit questions, and grading keys can be generated. Full documentation accompanies the test bank.
- Printed Test Bank. A printed version of the Computerized Test Bank is also available. This supplement was prepared by Amy F. Relyea of Miami University of Ohio.
- Professor's Grade Book. Computer software for managing student records accompanies the Computerized Test Bank.
- Data Disk. A computer diskette is available containing the data sets for 384 exercises, including all data sets for the MINITAB problems. The data sets are stored as ASCII files, and they may be freely duplicated for student use at adopting institutions. The data disk will be distributed to instructors free of charge upon adoption.

# Acknowledgments

My sincere thanks are extended to the many individuals who contributed to the development of this book. The following people served as reviewers.

Graydon Bell

Northern Arizona University

Patricia M. Buchanan

Pennsylvania State University

Chris Burditt

Napa Valley College

Darrell F. Clevidence

Carl Sandburg College

Pat Deamer

Skyline College

William D. Ergle

Roanoke College

Bryan V. Hearsey Lebanon Valley College

William E. Hinds

Midwestern State University

Kermit Hutcheson University of Georgia Marlene J. Kovaly

Florida Junior College at Jacksonville

Mike Orkin

California State University, Hayward

Larry Ringer

Texas A & M University

Gerald Rogers

New Mexico State University

Adele Shapiro

Palm Beach Community College

George Sturm

Grand Valley State University

Mary Sue Younger University of Tennessee

Douglas A. Zahn

Florida State University

Considerable effort has been devoted to make the text as accurate as possible. I am grateful to both Patricia Buchanan of the Pennsylvania State University and David Mathiason of the Rochester Institute of Technology for serving as accuracy reviewers of all examples and all answers at the back of the text. Any errors that might remain are, however, the sole responsibility of the author.

For their assistance in class-testing the manuscript at Elizabethtown College, I would like to thank Donald E. Koontz, John E. Koontz, Robert K. Morse, Larry Polin, and Laurie Showers. Special thanks are extended to Ronald Shubert, who also participated in the class-testing and wrote the Student Solutions Manual and the Instructor's Manual. I also want to thank Amy F. Relyea for preparing the Printed Test Bank.

I am grateful to Minitab, Inc. for providing a copy of the latest release of MINITAB available at the time of this writing—Release 8.2 for the MS-DOS standard version. Information about MINITAB can be obtained by contacting

Minitab, Inc. 3081 Enterprise Drive State College, PA 16801 Phone: (814) 238-3280

Fax: (814) 238-4383

MINITAB is a registered trademark of Minitab, Inc.

I am greatly indebted to the staff at Saunders College Publishing. I wish to thank Robert B. Stern (Senior Acquisitions Editor), Kimberly A. LoDico (Project Editor), Donald J. Gecewicz (Developmental Editor), Karyn Valerius (Editorial Assistant), Christine Schueler (Art Director), and Dena Digilio-Betz (Photo Research Editor).

Finally, I wish to extend special thanks to my wife, Judith, for her support and encouragement during the preparation of this work.

E.A.B. Lebanon, Pennsylvania November, 1992

# Index of Applications

Acceptance rates, Pennsylvania colleges, 33, 44

AFDC monthly benefits, 25, 30, 31, 42, 69, 76, 81, 84

Aflatoxin, safety of country's corn supply, 289

Aggressiveness, boys and later criminal activities, 519

Aggressiveness, girls and later criminal activities, 520

Air conditioning, percentage of cars equipped with, 176, 232

Airline complaints, baggage, 282

Alarm systems, burglarized houses with, 353

American League, batting averages and wins, 127, 599

American League, earned run averages and wins, 97, 124, 126, 580, 585, 593

Apples, regional comparison of ascorbic acid, 550

Arrests, driving under the influence, 62

Asbestos, compliance costs with new regulations, 393

Asbestos, health risks in schools, 248

Assaults, college campuses, 44

Attendance, two-year colleges, 5

Bald eagles, nest sites in Maine, 294

Batteries, comparison of brands, 556, 628

Bicentennial contest, extemporaneous speaking, 243

Birthday, employees receiving the day off, 352

Birthday problem, group sizes and probabilities, 171

Births, number of female, 145

Births, percentage of male, 170

Blue moon, frequency of, 155

Boggs, Wade, on-base percentage, 181

Breakfast, percentage of children who skip, 403

Butter, percentage of saturated fat, 423

Cable and pay TV, monthly basic rates, 53

Caesarean deliveries, increases in, 282

Caesarean section, number of live births, 145

Cafeterias, subsidized by employers, 282

Calculators, use in math classes, 358, 405, 419

Camcorders, type of store and selling price, 553, 628

Cancellations, airline flights, 145

Cancer, colon, levamisole and fluorouracil treatment, 466

Cancer, poor diet and risk awareness, 198, 223, 237

Candy sales, market share by company, 52

Cats, owners who talk to their, 198

Cholesterol, plasma total levels in males, 82, 86

Cholesterol, risk levels, 50

Coffee manufacturers, decreasing package size, 270, 314

Coffee, steeping method of brewing, 15

Coffee, use of drip-filtered method, 251

College athletes, majoring in business, 278

Colleges, tuition charges and faculty salaries, 641

Compound interest, 19

Coupons, redemption savings, 224

Credit card transactions, processed by Royal Bank of Canada, 248

DDT, in trout from different lakes, 550, 651

Dental injuries, seasonal variation and gender, 509

Dental injuries, type and patient gender, 508

Depression, among outpatients at health centers, 353

Dioxin, cardboard milk containers, 435

Doctor visits, family practice, 172

Doctorates, granted by mathematical sciences departments, 505

Documentary fees, new car purchases, 484, 621

Dog food, comparison of brands, 549

Dow Jones Utility Average, stock price changes, 132

Drivers, aggressive, 466

Electricity, sources of, 93

Enrollments, women in higher education, 171

Entry systems, key card, 153

Exercise, life extension, 17, 119, 128, 571, 591

Facelifts, ages of those receiving, 423

Faculty, interest in research or teaching, 17

Family, public's attitude about status of, 145

Fish farm, comparison of diets, 625

Food labels, meat products, 489, 514

Forbes Four Hundred, ages of richest people in America, 57

Foreign language, requirement for bachelor of arts, 371

Fruits, per capita consumption, 68

Gender pronouns, use of, 633

Geography, seriously deficient in knowledge of, 371

Government workers, educational level, 159, 232, 499

Grace, saying before holiday meals, 16

Grapes, Chilean, cyanide contamination, 325

Gum disease, probability of contracting, 286

Hair loss, drug to prevent, 352

Hand function, Jebsen-Taylor test, 132, 570, 579

Harvard freshmen, percentage staying up late, 173

Health care costs, fear of, 15

Health care costs, smoking, 483

HIV virus, NFL mandatory testing, 350

Hole-in-one, U.S. Open, Oak Hill Country Club, 177

Honey, colony yield of, 424

Hospital delivery, cost for normal pregnancy, 286, 317, 347

Households, average size, 358

Housing, amount of living space, 274

Housing costs, Habitat for Humanity, 274

Housing prices, Lancaster County, Pennsylvania, 286

Housing prices, metropolitan areas, 36, 73, 460

Hypertension, treatment and side effects, 466

Identical twins, probability in twin births, 180, 237

Insurance, automobile premiums by state, 94

Internal Revenue Service, taxpayer suits against, 283

Internal Revenue Service, TCMP program, 243, 358

Lake contamination, industrial wastes, 449

- Lead exposure, impaired neurological development, 182, 237, 286
- Left-handers, vulnerability to accidents, 290
- Libraries, state expenditures, 21, 36, 72
- Lincoln pennies, graded by date, 496
- Lincoln pennies, graded by date and condition, 501
- Living alone, number of persons by age and sex, 169
- Lobsters, prices received by fishermen for, 270
- Lotteries, number of states with, 240
- Lottery, Pennsylvania Super 7, 219, 240
- Lottery, testing randomness of daily numbers, 651
- Lottery, Tri-State MEGABUCKS, 242
- Marriage, effect on health, 316
- Marriage, first, average age of men, 320, 398
- Medical insurance plans, cost to the employer, 335, 356
- Medical interns, length of work week, 439
- Metcalf, Eric, extra foot bone, 247
- Metropolitan areas, changes in population rank, 644
- Microwave ovens, percentage of houses with, 215, 223, 237
- Miles-per-gallon, pickup trucks, 45
- Milk, comparing costs in geographical regions, 555, 629
- Mortgage rates, eastern Pennsylvania, 336, 370, 411
- Myocardial infarction, influence of gender on survival, 517
- NFL Super Bowl, rushing performance leaders, 51
- Nintendo game pak, request for, 182
- Oat bran, in doughnuts, 275
- Oat bran, in soft drinks and candy, 352
- Occupational therapy treatment and severity of disability, 509

- Off-price retailers, shopping time in stores, 317
- Olympic gold medals, 35, 44, 47
- Paper, bond writing, quality and sources of pulp, 557, 630
- Pascal's famous wager, 201
- Paychecks, expenditures by category, 51
- Peanut butter, relating quality and price, 638
- Persian Gulf War, salt content of readyto-eat meal, 49
- Personal computers, access times of hard disk drives, 508
- Personal computers, market share by company, 52
- Personal computers, operating speed of microprocessor, 318, 343, 385
- Personal computers, percentage of households with, 222
- Personal computers, software, technical assistance, 499
- Personal computers, survey on laptops, 251, 269
- Personal computers, word processing training programs, 555, 628
- Ph.D.s, earned by women, 23
- Ph.D.s, salary offers and fields of study, 24
- Pizza companies, percentage of total sales, 93
- Pizza, frozen, price comparison of, 652
- Population, percentage change by state, 88
- Population, United States, 1, 116
- Populations, 10 most populous states, 2
- Porsche 911, major maintenance service times, 369, 399, 418
- Potato barrel, eliminating the wooden, 274
- Poverty, among first graders, 251
- Power windows, percentage of cars equipped with, 168, 232

Prices, gold and silver by year, 118, 127, 570, 580, 586, 592

Psyllium, in breakfast cereal, 366

Radial keratotomy, correction of vision problems, 352

Radial tires, percentage of new cars equipped with, 282

Radon, comparing schools for levels of, 549, 629

Radon, dangerous levels in homes, 193

Radon, expected number of deaths from, 252

Recreational land developers, prizes given by, 185, 201

Resource-Based Relative Value Scale (RBRVS), Medicare, 98, 559

Schools, graded by students, 252

Seat belt users, percentage by state, 63, 73

Sirloin beef, grams of fat, 274

Six Sigma strategy, quality control, 74, 255

Smokers, in Utah, 286

Smokers, state with highest percentage of, 231

Smoking, Surgeon General's report, 511

Space orbiting debris, risk to space vehicles, 180

Space shuttle flight, probability of catastrophic failure, 177, 233

Spaghetti sauce, amount of beef in, 538

Sports cars, different versions of same model, 494

Stains, comparing drying times of brands, 551, 652

Stock market newsletter, monitoring interest rates, 387

Stock prices, January effect, 119, 599

Stocks, most actively traded, 18

Students-per-crime ratios, college campuses, 44

Sun, survey about, 404

Tax Reform Act of 1986, 333

Tax returns, fees charged and errors made by preparers, 594, 603, 646

Teachers, average work week, 347, 358, 399, 418

Teachers, moonlighting during school year, 282, 463

Telephone surveys, using 900 numbers, 298

Temperature, converting from Fahrenheit to Celsius, 18

Thyroid stimulating hormone (TSH), analyzing low levels of, 367

Tree stands, related hunting injuries, 392, 404

Truck driver deaths, testing positive for alcohol, 352

Tuberculosis cases, Mountain states, 15

Tuition, Pennsylvania medical schools, 86

Unemployment rates, state, 35, 79, 87

Valves, automobile engines, 393

Vehicles, by country, 70

Video cassette recorder, households with, 231

Video display terminals, health risks, 427, 465

Vitamins, percentage of adults who take regularly, 282, 405

Wages, Asian countries, 70

Wages, European countries, 71

Water, hardness, 479, 528

Wedding ring, percentage of women who wear, 222

Winter blues (SAD), full-fledged, Montgomery County, Maryland, 232

Winter blues (SAD), mild form, Montgomery County, Maryland, 500

World Series, managers with at least 20 wins, 61

Writing readiness skills, sensorimotor treatment program, 457

Yogurt, per capita consumption, 285

# Contents

Chapter 1

A First Look at Statistics and MINITAB 1	
1.1 1.2 1.3	Why Study Statistics? 1 What Is Statistics? 3 An Introduction to MINITAB 7 Looking Back 14 Key Words 14 MINITAB Commands 14 Review Exercises 15
Chap	ter 2
Describing Data: Graphical and Numerical Methods 22	
2.1	Graphical Methods for Quantitative Data: Frequency Distributions and Histograms 24 Frequency Distributions 24 Relative Frequency Distributions 28 Histograms 29
2.2	Graphical Methods for Quantitative Data: Dotplots and Stem-and-Leaf Displays 37  Dotplots 38  Stem-and-Leaf Displays 39
2.3	Graphical Methods for Qualitative Data: Bar and Pie Charts 46 Bar Charts 47 Pie Charts 47
2.4	Numerical Measures of Central Location 53 The Mean 54 The Weighted Mean 56 The Median 58 The Mode 60

2.5	Numerical Measures of Variability 64
	The Range 64
	The Variance 65
	The Standard Deviation 68
2.6	Understanding the Significance of Standard
	Deviation 74
	The Empirical Rule 75
	Chebyshev's Theorem 77
2.7	Measures of Relative Position 80
	Z-Scores 80
	Percentiles 82
	5-Number Summaries 83
	Boxplots 84
	Looking Back 88
	Key Words 90
	MINITAB Commands 91
	Review Exercises 91
Char	oter 3
-	
	eling Bivariate Data: Describing the Relationship
Betw	veen Two Variables 96
3.1	Straight Lines and the SS( ) Notation 99
	Straight Lines 100
	The SS( ) Notation 102
3.2	Scatter Diagrams 105
3.3	Linear Regression and the Method of Least Squares 110
3.4	Correlation: Measuring the Usefulness of the Model 120
	Looking Back 129
	Key Words 130
	MINITAB Commands 131
	Review Exercises 131

# Chapter 4

Probability: Measuring Uncertainty 134

**4.1** Probability, Sample Spaces, and Events 136

The Relative Frequency Definition of Probability 139