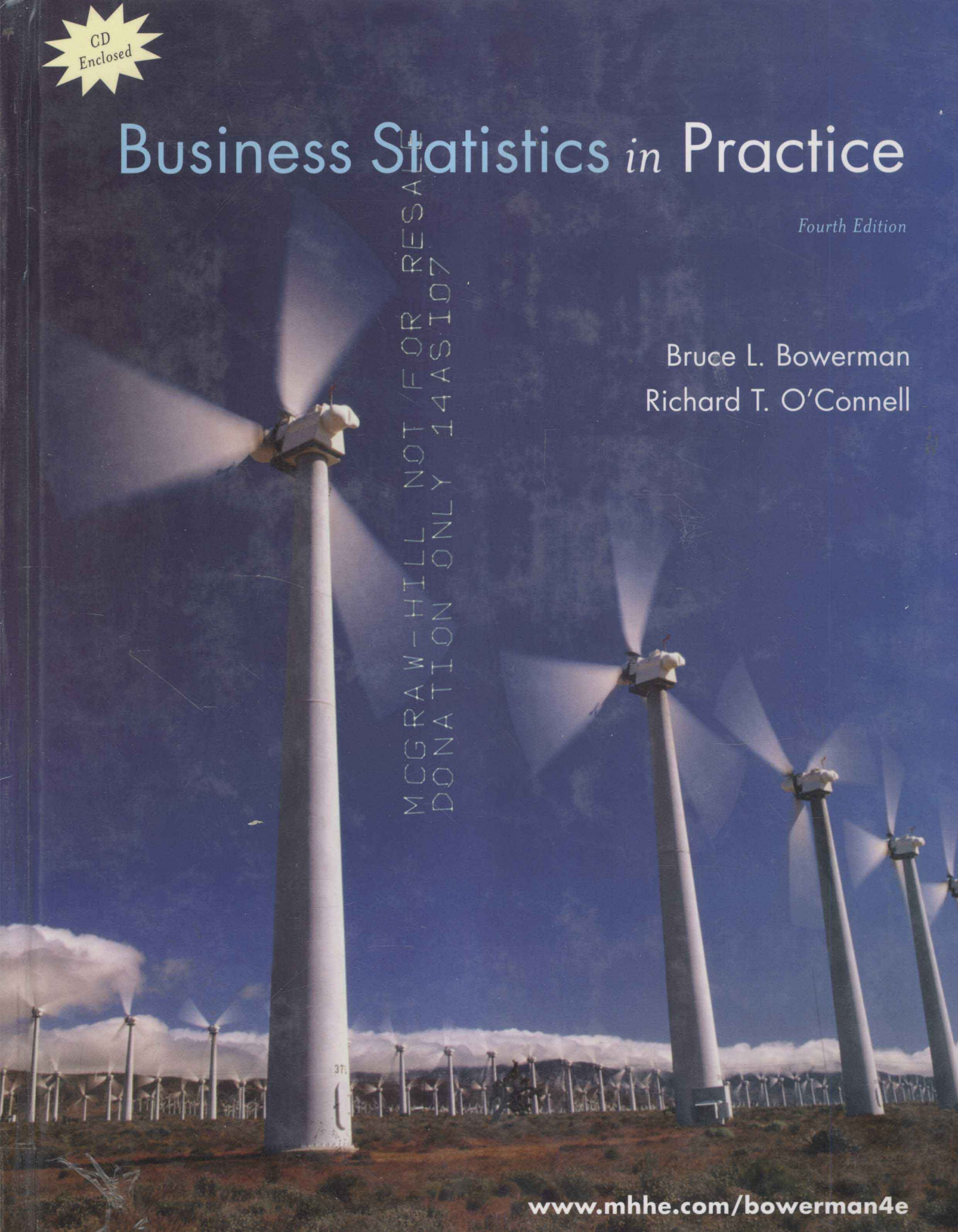
CD Enclosed

Business Statistics *in* Practice

Fourth Edition

Bruce L. Bowerman
Richard T. O'Connell



www.mhhe.com/bowerman4e

Bruce L. Bowerman

Richard T. O'Connell

Miami University

Business Statistics in Practice

FOURTH EDITION

with additional examples and exercises and selected appendices by
Steven C. Huchendorf

University of Minnesota

with MegaStat software and other contributions by
J. Burdene Orris

Butler University

McGraw-Hill Irwin

Boston Burr Ridge, IL Dubuque, IA Madison, WI New York San Francisco St. Louis
Bangkok Bogotá Caracas Kuala Lumpur Lisbon London Madrid Mexico City
Milan Montreal New Delhi Santiago Seoul Singapore Sydney Taipei Toronto

BUSINESS STATISTICS IN PRACTICE

This book is printed on acid-free paper.

2 3 4 5 6 7 8 9 0 WCK/WCK 0 9 8 7 6

ISBN-13: 978-0-07-297747-9
ISBN-10: 0-07-297747-7

Editorial director: *Brent Gordon*
Executive editor: *Scott Isenberg*
Senior developmental editor: *Wanda J. Zeman*
Senior marketing manager: *Douglas Reiner*
Senior media producer: *Victor Chiu*
Project manager: *Laura Griffin*
Production supervisor: *Debra R. Sylvester*
Senior designer: *Mary E. Kazak*
Senior photo research coordinator: *Jeremy Cheshareck*
Photo researcher: *Julie Tesser*
Media project manager: *Matthew Perry*
Senior supplement producer: *Carol Loreth*
Cover and interior design: *Jenny El-Shamy*
Cover image: © *Glen Allison/Gettyimages*
Typeface: *10/12 Times Roman*
Compositor: *Interactive Composition Corporation*
Printer: *Quebecor World Versailles Inc.*

Library of Congress Cataloging-in-Publication Data

Bowerman, Bruce L.
Business statistics in practice / Bruce L. Bowerman, Richard T. O'Connell ; with MegaStat and other contributions by J. Burdeane Orris.— 4th ed.
p. cm.
Includes bibliographical references and indexes.
ISBN-13: 978-0-07-297747-9
ISBN-10: 0-07-297747-7 (alk. paper)
1. Commercial statistics. 2. Statistics. I. O'Connell, Richard T. II. Orris, J. B. III. Title.
HF1017.B654 2007
519.5'024'65—dc22

2005053422

www.mhhe.com

Bruce L. Bowerman
To my wife, children, sister, and other family members:
Drena
Michael, Jinda, Benjamin, and Lex
Asa and Nicole
Susan
Fiona and Radeesa
Daphne, Chloe, and Edgar
Gwyneth and Tony

Richard T. O'Connell
To my wife and children:
Jean
Christopher and Bradley

About the Authors

Bruce L. Bowerman Bruce L. Bowerman is professor of decision sciences at Miami University in Oxford, Ohio. He received his Ph.D. degree in statistics from Iowa State University in 1974, and he has over 37 years of experience teaching basic statistics, regression analysis, time series forecasting, survey sampling, and design of experiments to both undergraduate and graduate students. In 1987 Professor Bowerman received an Outstanding Teaching award from the Miami University senior class, and in 1992 he received an Effective Educator award from the Richard T. Farmer School of Business Administration. Together with Richard T. O'Connell, Professor Bowerman has written 11 textbooks. These include *Forecasting and Time Series: An Applied Approach; Forecasting, Time Series, and Regression: An Applied Approach* (also coauthored with Anne B. Koehler); and *Linear Statistical Models: An Applied Approach.* The first edition of *Forecasting and Time Series* earned an Outstanding Academic Book award from *Choice* magazine. Professor Bowerman has also published a number of articles in applied stochastic processes, time series forecasting, and statistical education. In his spare time, Professor Bowerman enjoys watching movies and sports, playing tennis, and designing houses.

Richard T. O'Connell Richard T. O'Connell is associate professor of decision sciences at Miami University in Oxford, Ohio. He has more than 32 years of experience teaching basic statistics, statistical quality control and process improvement, regression analysis, time series forecasting, and design of experiments to both undergraduate and graduate business students. He also has extensive consulting experience and has taught workshops dealing with statistical process control and process improvement for a variety of companies in the Midwest. In 2000 Professor O'Connell received an Effective Educator award from the Richard T. Farmer School of Business Administration. Together with Bruce L. Bowerman, he has written 11 textbooks. These include *Forecasting and Time Series: An Applied Approach; Forecasting, Time Series, and Regression: An Applied Approach* (also coauthored with Anne B. Koehler); and *Linear Statistical Models: An Applied Approach.* Professor O'Connell has published a number of articles in the area of innovative statistical education. He is one of the first college instructors in the United States to integrate statistical process control and process improvement methodology into his basic business statistics course. He (with Professor Bowerman) has written several articles advocating this approach. He has also given presentations on this subject at meetings such as the Joint Statistical Meetings of the American Statistical Association and the Workshop on Total Quality Management: Developing Curricula and Research Agendas (sponsored by the Production and Operations Management Society). Professor O'Connell received an M.S. degree in decision sciences from Northwestern University in 1973, and he is currently a member of both the Decision Sciences Institute and the American Statistical Association. In his spare time, Professor O'Connell enjoys fishing, collecting 1950s and 1960s rock music, and following the Green Bay Packers and Purdue University sports.

Preface

In *Business Statistics in Practice, Fourth Edition,* we provide a modern, practical, and unique framework for teaching the first course in business statistics. This framework features case study and example driven discussions of all basic business statistics topics. In addition, we have endeavored to make this book the most clearly written, motivating, and easy to use business statistics text available. We have taken great pains to explain concepts simply from first principles. Therefore, the only prerequisite for this book is high school algebra.

Business Statistics in Practice has five attributes that make it an effective learning tool:

- A consistent theme of business improvement through statistical analysis.
- A unique use of "continuing" case studies that integrates different statistical areas.
- A real emphasis on the study of variation that stresses that the analysis of individual population observations is as important as the analysis of population means.
- A flexible topic flow that facilitates different topic choices and encourages different teaching approaches. In particular, since many courses give different emphases to probability, hypothesis testing, regression and statistical modeling, nonparametric statistics, and quality control, this book provides great flexibility with respect to how, when, and whether to cover these topics.
- A modern use of the statistical capabilities of the software packages MINITAB, Excel, and MegaStat (an Excel add-in package included on the text's student CD-ROM) that stresses statistical interpretation and reflects the use of these packages in the real world.

New to the fourth edition are

- *The cell phone case,* which is the first continuing case in Chapter 1 and discusses how a bank uses a random sample to estimate its cell phone costs. Using this estimate, the bank decides whether to outsource management of its wireless resources. This case should be particularly motivating to students because it addresses a real problem faced by both students and businesses—unpleasantly high cell phone bills.
- *Continuing cases with no need to refer back to previously given computer outputs.* Each time a continuing case is revisited, any needed computer output is included with the current case discussion. In addition, whenever possible the background information needed to understand the current analysis is provided, so the student does not need to refer back to previous material.
- *Business improvement icons* BI—placed in the page margins—that identify when an important business conclusion has been reached using statistical analysis. Each conclusion is also highlighted for additional emphasis.
- *Confidence intervals for and hypothesis tests about a population mean presented by using the σ known/σ unknown approach.* This approach simplifies the choice of z or t-based procedures and is consistent with computerized procedures provided by MINITAB, Excel, and MegaStat. A t distribution table with up to 100 degrees of freedom is given in Table A.4 of Appendix A. Confidence intervals for and hypothesis tests about the difference between two population means are also presented using the σ known/σ unknown approach.
- *Completely updated end of chapter computer appendices* that clearly show how to perform statistical analysis using MINITAB (Version 14), Microsoft Excel 2003, and the latest version of MegaStat.
- *Expanded coverage of sampling in Chapter 1.* We now discuss using both a random number table and computer generated random numbers to select a random sample. We also have added an optional section that introduces stratified, cluster, and systematic sampling and discusses the problems of undercoverage, nonresponse, and response bias.
- *A substantial number of new, real world data sets in the exercises,* particularly in the exercises of Chapter 1 (An Introduction to Business Statistics) and Chapter 2 (Descriptive Statistics).

- *An optional appendix on covariance and correlation.* This end of book appendix (Appendix B) can be covered either after covering scatter plots in Chapter 2 or before covering simple linear regression analysis in Chapter 11. Or, it can be omitted entirely without loss of continuity.
- *An optional appendix on normal probability plots.* This end of book appendix (Appendix D, Part 1) supplements the normal distribution discussion in Chapter 5.
- *A simpler and easier to understand example introducing sampling distributions.* This stock return example motivates the discussion of the sampling distribution of the sample mean in Chapter 6.
- *Increased emphasis on the concept of the margin of error* to better motivate the discussion of confidence intervals in Chapter 7.
- *A step-by-step hypothesis testing approach* that is used in almost all hypothesis testing examples in Chapter 8 (Hypothesis Testing) and Chapter 9 (Statistical Inferences Based on Two Samples). This approach consists of a seven-step procedure that is designed to break hypothesis testing down into small, easy to understand steps and to also clearly show how to use the book's hypothesis testing summary boxes. Although the seven-step procedure is not formally used after Chapter 9, the students' familiarity with the steps and summary boxes should enable them to successfully carry out hypothesis tests in later chapters.
- *Increased emphasis in Chapter 9 on the "unequal variances" t-based procedure for comparing two population means.* This procedure is becoming increasingly popular because it is available in most statistical software packages and is a very accurate approximation that does not require assuming equal population variances.
- *A simplified and improved discussion of simple and multiple regression analysis.* In simple regression (Chapter 11), we give more concise explanations of the simple linear regression model, least squares, and confidence and prediction intervals. In addition to using improved graphics, the chapter also provides the flexibility to cover simple coefficients of determination and correlation (Section 11.6) early or later in the chapter. In multiple regression (Chapter 12), we have refined the innovative, modular organization of the third edition. This will make it easier to selectively cover whatever multiple regression (and model building) topics are desired. We have also simplified the presentation of dummy variables and added a short section on *logistic regression.* In both the simple and multiple regression chapters, we have improved our explanations and use of MINITAB, Excel, and MegaStat regression outputs. Key outputs are more clearly annotated to help the beginner find needed regression quantities.
- *An optional appendix on Holt–Winters' exponential smoothing models.* This appendix (Appendix E) is now included in the book (and on the book's CD-ROM).

In addition, as in the third edition, there is an optional section in Chapter 5 that covers use of the cumulative normal table. Although (because of reviewer input) we use the standard normal table to explain confidence intervals and hypothesis tests based on the normal distribution, we have explicitly designed the figures illustrating normal curve areas so that the intervals and tests can also be explained using the cumulative normal table.

We now discuss in more detail the attributes that make *Business Statistics in Practice* an effective learning tool.

Business improvement through statistical analysis The ultimate goal of statistical analysis in business is business improvement. This theme is the foundation for the case studies and examples in this text, many of which are based on actual, real world situations. For example, consider the following synopses of three case studies.

- **The Cheese Spread Case:** The marketer of a soft cheese spread wishes to replace the spout on its plastic dispenser with a less expensive spout. The company uses confidence intervals to conclude that demand for the spread will remain sufficiently high when the change is made to make replacing the spout profitable.
- **The Trash Bag Case:** A leading producer of trash bags uses hypothesis testing to convince the standards and practices division of a major television network that advertising claims about its newest trash bag are valid.
- **The Fuel Consumption Case:** A natural gas company uses regression analysis to predict its city's natural gas needs accurately enough to avoid paying fines to a pipeline transmission system.

In each of these cases, statistical analysis leads to an informed action (replace the spout, advertise the claim, use the regression prediction procedure) that results in business improvement. Furthermore, we continue this theme throughout the presentation of all statistical techniques in this book. For instance, we use descriptive and inferential statistics to compare the risk and return characteristics of different investment choices in order to improve the way we manage an investment portfolio; we use statistical process control to improve manufacturing and service processes; and we use design of experiments to study the effects of several different advertising campaigns in order to improve how a product is marketed.

A unique continuity of presentation and use of case studies *Business Statistics in Practice* features a unique continuity of presentation that integrates different statistical areas. This integration is achieved by an early emphasis (in Chapters 1 and 2) on the difference between the population and the sample and by a continuing use of practical, realistic case studies that span not only individual chapters but also groups of chapters. Specifically, Chapter 1 shows how to select random (or approximately random) samples from populations and processes by introducing four case studies as examples and by presenting additional case studies as exercises. Then in Chapter 2 we show how to use descriptive statistics to estimate the important aspects of these populations and processes. We continue to employ these case studies through the probability and sampling distribution chapters until we use confidence intervals and hypothesis testing to make statistical inferences. Furthermore, we introduce new case studies in each and every chapter. For example, we introduce several case studies in our presentation of simple linear regression and then extend these case studies when we discuss multiple regression and model building to show how regression is used in the description, prediction, and control of business variables.

A real emphasis on the importance of variation *Business Statistics in Practice* emphasizes that since businesses must satisfy individual customers, the analysis of individual population observations—which is achieved by analyzing population variation—is as important as analyzing the population mean. Our discussion of variation begins in Chapter 1, where we intuitively examine the variation of sample data and use simple runs plots to evaluate statistical control. This discussion continues in Chapter 2, where we use the empirical rule to estimate tolerance intervals containing different percentages of population observations. For example, we use the empirical rule in the

- **Payment Time Case** to describe the variation of individual bill payment times around the estimated mean bill payment time for a new electronic billing system.
- **Marketing Research Case** to describe the variation of individual customer ratings of a new soft drink bottle design around the estimated mean rating of the new design.
- **Car Mileage Case** to describe the variation of individual gas mileages around the estimated mean mileage obtained by a new midsize car.

In addition, in the **coffee temperature case** we introduce the idea of process capability—determining whether almost all process observations fall within customer requirements—and in other case studies we consider the problems involved with describing the variation of highly skewed populations.

Our emphasis on variation continues throughout the book. For example, in Chapter 7 we clearly distinguish between a confidence interval for a population mean and a tolerance interval for a given percentage of individual population measurements. In Chapter 8 we discuss the effect of variation on the interpretation of a hypothesis test about the population mean. In Chapters 11 through 13 we show how prediction intervals can be used to evaluate the predictive capabilities of different regression and time series forecasting models. In addition, we demonstrate how prediction intervals are used to assess whether any individual population observations are "unusual" enough to suggest the need for process improvement. Finally, in Chapter 14 we present a complete discussion of statistical process control and improvement (including the six sigma philosophy adopted by Motorola, Inc., and a number of other prominent U.S. companies). Furthermore, in all of these chapters we use practical case studies to illustrate the ideas being presented.

A flexible topic flow Although the table of contents of this book reveals a rather standard topic organization, the book utilizes a flexible topic flow that facilitates different topic choices and encourages different teaching approaches. In particular, since different courses place different amounts of emphasis on probability, hypothesis testing, regression and statistical modeling, nonparametric statistics, and quality control, this book provides great flexibility with respect to

how, when, and whether to cover these topics. Furthermore, in optional sections, appendices, and self-learning exercises, the book gives the student the opportunity to study more advanced topics in a concise and practical way. Thus, as we now discuss, courses with a wide variety of topic coverages and emphases can be taught using this book.

Probability The most minimal approach to probability would cover Section 3.1 (the concept of probability), Section 4.1 (random variables), Section 5.1 (continuous probability distributions), and Section 5.3 (the normal distribution, including an intuitive example of the addition rule for mutually exclusive events). These sections are the only prerequisites for Chapters 6 through 14 (sampling distributions, confidence intervals, hypothesis testing, experimental design, regression, time series forecasting, and quality control).

Instructors who wish to also cover discrete probability distributions (Chapter 4) have the option of doing this either with a fairly minimal probability background or with a complete probability background. The fairly minimal probability background consists of Section 3.1 (the concept of probability) and Section 3.2 (using sample spaces to find probabilities). Note that this background is sufficient because, since Example 4.2 of Chapter 4 intuitively illustrates the multiplication rule for independent events and the addition rule for mutually exclusive events in the context of finding a discrete probability distribution, it is not necessary to cover the complete discussion of probability rules given in Sections 3.3 and 3.4. Of course, this complete discussion is necessary background for covering chi-square tests of independence (Chapter 16) and Bayes' Theorem and decision theory (Chapter 17). Also, the complete discussion features the **AccuRatings Case,** which is a very motivating data driven application of the probability rules.

Hypothesis testing In the fourth edition we have used a seven-step procedure to break hypothesis testing down into small, easy-to-understand steps and to clearly show how to use the book's hypothesis testing summary boxes. In addition, we have fully and concisely integrated the discussion of using rejection points and p-values. The seven-step procedure shows how to use both approaches, and the hypothesis testing boxes summarize both rejection points and p-values for each test. We have also motivated the link between the approaches by considering how major television networks sometimes use different α values when evaluating advertising claims. We are aware of several courses that introduce hypothesis testing in the context of using p-values to test the significance of regression coefficients. This can be done in our book by totally skipping Chapter 8 and by noting that every section throughout the rest of the book includes self-contained summary boxes (and examples) that fully cover any needed confidence intervals and hypothesis tests. Also, Chapter 6 (sampling distributions) intuitively illustrates the use of p-values in the context of evaluating a claim about a population mean and in the context of evaluating a claim about a population proportion. Therefore, Chapter 6 can be used as an extremely short, intuitive introduction to p-values.

Regression and statistical modeling The fourth edition features an innovative organization of regression analysis that simplifies the flow of the overall discussion and makes it very easy to cover whatever regression topics are desired. As in the third edition, we have included an optional section on residual analysis at the end of the simple linear regression chapter (Chapter 11). In Chapter 12: Multiple Regression and Model Building, we have refined the modular organization of the third edition and have made it easier to cover whatever portions of multiple regression and model building are desired. As shown in a diagram on its opening page, Chapter 12 consists of four parts. Part 1: Basic Multiple Regression discusses the basic descriptive and inferential techniques of multiple regression analysis and would be a sufficient introduction to this topic for many introductory business statistics courses. After completing Part 1, the reader can study optional Part 2: Using Squared and Interaction Terms, optional Part 3: Dummy Variables and Advanced Statistical Inferences, and any section of optional Part 4: Model Building and Model Diagnostics. These optional parts can be covered in any order and without loss of continuity (note that Part 4 consists of four self-contained sections: model building and the effects of multicollinearity; residual analysis in multiple regression; diagnostics for detecting outlying and influential observations; and logistic regression). Furthermore, optional material covering model diagnostics and topics in some of the supplementary exercises tie key portions of the four parts together. This approach allows instructors to easily cover what they consider most important in courses with limited time devoted to regression analysis. Similarly, since many business statistics courses do not have substantial time to devote to experimental design (Chapter 10) and time series

forecasting (Chapter 13), we have put great effort into making our presentation of these topics both complete and easy to get through.

Nonparametric statistics We have placed all of the nonparametric techniques covered in the book in Chapter 15. Furthermore, at the end of the discussion of each parametric technique in Chapters 8 through 11 we refer readers to the section in Chapter 15 that discusses the nonparametric technique that would be used if the assumptions for the parametric technique fail to hold. Therefore, the instructor has the option of integrating the discussion of nonparametric statistics into the main flow of Chapters 8 through 11.

Quality control Process improvement through control charts is discussed in Chapter 14. Thus, this topic is placed outside of the main flow of what might be regarded as classical statistics. However, since Chapter 14 has as its only prerequisite Chapter 6 on sampling distributions, the instructor has the option to cover Chapter 14 at any point after Chapter 6.

Optional Advanced Topics In optional sections, appendices, and self-learning exercises, the book gives the student the opportunity to study more advanced topics in a concise and practical way. Examination of the table of contents reveals that many of the more advanced topics—for example, counting rules (Appendix C, Part 1), the hypergeometric distribution (Appendix C, Part 2), covariance and correlation (Appendix B), normal probability plots (Appendix D, Part 1), the Poisson and exponential distributions (Sections 4.4 and 5.5), calculating the probability of a Type II error (Section 8.6), and statistical inferences for a population variance (Section 8.8)—are included in many other business statistics books. However, some of the more advanced topics, while not unique to this book, are less frequently covered in other basic statistics texts. These topics (the most advanced of which are discussed in CD-ROM Appendices F through L) are as follows:

- Properties of the Mean and Variance of a Random Variable, and the Covariance Between Two Random Variables (Appendix D, Part 2).
- Derivations of the Mean and the Variance of the Sample Mean and of the Mean and the Variance of the Sample Proportion (Appendix D, Part 3).
- Confidence Intervals for Parameters of Finite Populations (Section 7.5), including sample size determination (Exercise 7.57).
- An Introduction to Survey Sampling (Section 1.5); estimation formulas, optimal allocation, and sample size determination in stratified random sampling (Appendix F, Part 1); and estimation formulas in one- and two-stage cluster sampling and ratio estimation (Appendix F, Part 2).
- A Comparison of Confidence Intervals and Tolerance Intervals (Section 7.6).
- Using Matrix Algebra to Perform Regression Calculations (Appendix G).
- The regression approach to one-way analysis of variance (Exercise 12.45), and the regression approach to two-way analysis of variance (Appendix H).
- Advanced Model Diagnostics (Exercises 12.76 and 12.77) and Model Building with Squared and Interaction Terms (Exercise 12.74).
- Logistic Regression (Section 12.15) and Discriminant Analysis (Exercise 12.78).
- Factor Analysis, Cluster Analysis, and Multidimensional Scaling (Appendix I).
- The Box–Jenkins methodology, a fairly complete discussion featuring nonseasonal and seasonal modeling, using autocorrelated error term models in regression analysis, intervention analysis, and transfer function models (Appendix J).
- Individuals charts and c charts (Appendix L).

Furthermore, we have put great effort into making the discussion of all of the more advanced topics clear, concise, and easy to get through. This gives the instructor considerable flexibility in designing different business statistics courses. For example, a professor teaching a second course in business statistics can opt to either cover a variety of intermediate topics or present a more in-depth treatment of regression analysis and forecasting.

MINITAB, Excel, and MegaStat *Business Statistics in Practice, Fourth Edition,* features a modern use of the statistical capabilities of the software packages MINITAB, Excel, and the Excel add-in MegaStat. Throughout the book we provide an abundant number of outputs from all three packages in both examples and exercises that allow students to concentrate on statistical

interpretations. This use of outputs is particularly prominent in statistical areas where hand calculations are impossible or impractical and where having students run their own programs (while theoretically optimal) would, because of time constraints, not allow them to see a wide variety of applications. These areas include descriptive statistics, ANOVA, regression, and time series forecasting. In addition, appendices at the end of each chapter show in detail how to use MINITAB, Excel, and MegaStat to implement the statistical techniques discussed in the chapter. For the fourth edition, the developer of MegaStat, Professor J. B. Orris of Butler University, has worked closely with us. We believe that MegaStat is the most comprehensive, accurate, and easy to use Excel add-in package in existence. In addition to remedying most of the computational problems associated with Excel Data Analysis Tools, MegaStat is also specifically designed to enhance the use of *Business Statistics in Practice.* For example,

- In addition to giving the usual descriptive statistics, frequency distributions, and histograms, MegaStat provides stem-and-leaf displays, box plots, dot plots, runs plots, normal plots, and output for the Empirical Rule (as well as tolerance intervals estimated to contain any specified percentage of individual observations). MegaStat also gives the option to calculate tolerance intervals and confidence intervals using the same dialog box. Therefore, students can better understand the crucial difference between these two types of intervals (as illustrated on pages 293 and 294).
- The MegaStat dialog box for every one and two sample hypothesis testing procedure for means and proportions allows the user to calculate a confidence interval for the population parameter being tested. Therefore, the student is encouraged to evaluate both statistical significance and practical importance. Such evaluation is a consistent theme of *Business Statistics in Practice* (in particular, see Chapters 8 and 9).
- MegaStat's one-way ANOVA, randomized block, and two-factor ANOVA procedures provide graphical output helping students to better analyze experimental data. In addition, each procedure provides easy to understand pairwise comparisons of population means using both Tukey procedures and individual t-tests. Such graphical analysis and pairwise comparisons are emphasized in Chapter 10.
- In addition to providing confidence intervals and prediction intervals in simple and multiple regression, MegaStat gives a full range of residual plots, normal plots, and outlying and influential observation diagnostics, as well as the variance inflation factors for the independent variables in a regression model. In addition, MegaStat provides an all possible regressions output that summarizes all well known model selection criteria, as well as the p-values for the independent variables. MegaStat also gives a stepwise selection procedure that provides more information than given by classical stepwise regression or backward elimination. MegaStat's regression capabilities are designed to enhance the regression coverage in Chapters 11 and 12. Furthermore, all of MegaStat's regression capabilities can be accessed in one very easy to use dialog box, allowing the student to carry out a wide range of regression procedures in a correct, informative, and simple way.

In addition, MegaStat is fully capable of performing analysis related to discrete and continuous probability distributions, time series forecasting, nonparametric statistics, chi-square tests, and statistical quality control charts—virtually all topics covered by *Business Statistics in Practice.* MegaStat is provided on the student CD-ROM.

Further Features The book's CD-ROM, in addition to containing the previously discussed advanced topic appendices and MegaStat, also features Excel templates, data files, tutorials, web links, self graded quizzes, PowerPoint presentations, and Visual Statistics 2.0 by Doane, Mathieson, and Tracy. Visual Statistics is a Windows software program that helps students learn statistics through interactive experimentation and visualization. Visual statistics icons in the text identify concepts that are further explained by Visual Statistics. This edition also features Homework Manager. This is an online electronic tutor customized to the text and available as an option to students.

In addition, the book has the following supplements: an instructor's solutions manual developed by Patrick Schur, Miami University, and test bank developed by Denise Krallman, Miami University (included on the instructor's edition CD-ROM and available in print format); a student study guide developed by Sandra Strassar, Valparaiso University (available in print format); and PowerPoint transparency masters developed by Ronny Richardson, Southern Polytechnic State University, with contributions by Harvey Singer, George Mason University.

Acknowledgments We wish to thank the many people who have helped to make this book a reality. First, we wish to thank Drena Bowerman, who spent many hours cutting and taping and making trips to the copy shop, so that we could complete the manuscript on time. Second, we wish to thank Professor J. B. Orris of Butler University. Professor Orris's hard work and dedication have made MegaStat a truly excellent statistical software package. In addition, his feedback and many intellectual contributions have greatly improved this book. Third, we wish to thank Professor Steven Huchendorf of the University of Minnesota. Professor Huchendorf provided a substantial number of new exercises for the fourth edition and helped tremendously in the development and writing of Appendix B: Covariance and Correlation and Appendix D (Part 1): Normal Probability Plots. He did this in a very timely fashion, and we really appreciate his work. Fourth, we wish to thank Professor Michael L. Hand of Willamette University, who is a coauthor of the second edition of this book. Although Professor Hand did not work on the fourth edition, we thank him for his contributions to the second edition and regard him as a valued friend and colleague. Finally, we wish to thank Professor Anne Koehler of Miami University. Professor Koehler wrote the original versions of the MINITAB and Excel appendices included in the text. The final appendices are largely based on her original versions. We cannot thank Professor Koehler enough for her selfless work, which is a hallmark of her career. We also wish to thank the people at McGraw-Hill/Irwin for their dedication to this book. These people include Executive Editor Scott Isenberg, who has made many excellent suggestions for developing the second, third, and fourth editions of this book and who is, in general, a very positive and helpful resource to the authors; Executive Editor Dick Hercher, who persuaded us initially to publish this book with McGraw-Hill/Irwin and who continues to offer sound advice and support with each revision of the book; Senior Developmental Editor Wanda Zeman, who has shown great dedication through four editions of this book (Wanda's many excellent ideas and tireless attention to detail have been instrumental in the continual improvement of the book); and Project Manager Laura Griffin, who has capably and diligently guided this book through its production, and who consistently proves to be a tremendous help to the authors.

We thank the many contributors who reviewed text chapters and/or provided helpful feedback at one of our reviewer conferences. We are grateful to all of them.

Sung Ahn, Washington State University
Arsene Aka, Catholic University of America
Charles Apigian, Middle Tennessee State University
Philip Boudreaux, University of Louisiana–Lafayette
Robert Brookshire, James Madison University
Wen Chiang, University of Tulsa
Ali Choudry, Florida International University
Howard Clayton, Auburn University
Richard Cox, University of Arkansas–Little Rock
Bradford Crain, Portland State University
Nandita Das, Lehigh University
Cassandra DiRienzo, Elon University
Anne Marie Drougas, Dominican University
Joy Field, Boston College
Gary Franko, Siena College
David Friesen, Midwestern State University
Thomas Groleau, Carthage College
Cindy Hinz, Saint Bonaventure University
Johnny Ho, Columbus State University
Steven Huchendorf, University of Minnesota
Peter Ittig, University of Massachusetts
Chun Jin, Central Connecticut University
Mark Karseig, Central Missouri State University

Ron Klimberg, Saint Joseph's University
Jennifer Kohn, Montclair State University
Maria Lambert, Pitt Community College
Constantine Loucopoulos, Northeastern State University
Brad McDonald, Northern Illinois University
Constance McLaren, Indiana State University
Jackie Miller, Ohio State University
John M. Miller, Sam Houston State University
Patricia A. Mullins, University of Wisconsin–Madison
Tappan K. Nayak, George Washington University
Sufi M. Nazem, University of Nebraska–Omaha
Gary Nelson, Central Community College–Columbus
Ceyhun Ozgur, Valparaiso University
Jayprakash Patankar, University of Akron
Deborah Primm, Jacksonville State University
Harold Rahmlon, Saint Joseph's University
Said E Said, East Carolina University
Scott Seipel, Middle Tennessee State University
Sankara Sethuraman, Augusta State University
Daniel Shimsak, University of Massachusetts–Boston
Walter Simmons, John Carroll University
Harvey Singer, George Mason University
Robert Smidt, California Polytechnic University
Toni Somers, Wayne State University
Robert Stevens, University of Louisiana–Monroe
Faye Teer, James Madison University
Dharma Thiruvaiyaru, Augusta State University
James Thorson, Southern Connecticut State University
Bijesh Tolia, Chicago State University
Akinori Tomahara, Queens College
Fan T. Tseng, University of Alabama–Huntsville
Lee Van Scyoc, University of Wisconsin–Oshkosh
William Warde, Oklahoma State University
Allen Webster, Bradley University
Mark Witkowski, University of Arkansas–Little Rock
Louis Woods, University of North Florida
Jack Yurkiewicz, Pace University
Jay Zagorsky, Boston University
Zhen Zhu, University of Central Oklahoma
Zhiwei Zhu, University of Louisiana–Lafayette

We also wish to thank the error checker, Don R. Robinson, Illinois State University, who was very helpful. Most important, we wish to thank our families for their acceptance, unconditional love, and support.

Bruce L. Bowerman
Richard T. O'Connell

Guided Tour

Business Statistics in Practice, Fourth Edition, has been written with students' needs in mind. Its clear and understandable explanations and use of real world case studies and examples present content that business students can relate to. Because today's students learn in a visual and interactive way, the text is supplemented by a free student CD-ROM, containing a host of updated resources and helpful study aids. In addition, both students and instructors are provided with additional resources on the text website. Thus, students are given a number of statistical tools in a variety of ways and shown how these tools can be used to positively impact business and other organizations.

Chapter Introductions

Each chapter opens with a preview showing how the statistical topics to be discussed apply to real business problems. The continuing case examples that run throughout the book are briefly introduced along with the techniques that will be used to analyze them.

Visual Statistics 2.0

Visual Statistics, described later in the tour, helps students learn statistics through interactive experimentation and visualization. Concepts in the text that are treated in the Visual Statistics software program are identified by icon, with chapter reference, in the margin of the text next to the concept.

In Chapter 1 we saw that although we can sometimes take a census of an entire population, we often must randomly select a sample from a population. When we have taken a census or a sample, we typically wish to describe the observed data set. In particular, we describe a sample in order to make inferences about the sampled population.

In this chapter we learn about **descriptive statistics,** which is the science of describing the important characteristics of a population or sample. Generally, we look at several important aspects of a set of measurements. One such aspect is the **central tendency,** or middle, of the data set. For instance, we might estimate a typical bottle design rating in the marketing research case or a typical city driving mileage in the car mileage case. Another important aspect of a data set is the **variability,** or spread, of the data. For example, we might estimate the largest and smallest gas mileage that would likely be obtained when a new midsize car is purchased. Or, in the marketing research case we might measure the spread of the bottle design ratings. If the ratings are clustered closely together, consumers' ratings are much the same (or are consistent). If the ratings are spread far apart, then consumers have widely varying opinions of the new bottle design. A third important aspect of a data set is the **shape** of the population or sample. Looking at a data set's shape tells us how the population or sample is *distributed* over various values (more about this later). Still another important aspect is whether **outliers** exist. For instance, if there are outlying bottle design ratings, then several consumers have opinions about the design that are very different from the opinions of most of the sampled consumers. Descriptive statistics also involves using **graphical methods** to depict data sets and to study relationships between different variables.

In this chapter we use a variety of methods to describe the cell phone usages, bottle design ratings, coffee temperatures, and car mileages introduced in the cases of Chapter 1. In addition, we introduce three new cases:

The Payment Time Case: A management consulting firm assesses how effectively a new electronic billing system reduces bill payment times.

The Electronic Article Surveillance Case: A survey is used to study the unintended effects on consumer attitudes of false electronic article surveillance alarms.

The Marketing Ethics Case: A survey is conducted to study marketing researchers' attitudes toward violating confidentiality in marketing research studies.

2.1 Describing the Shape of a Distribution

We begin looking at the characteristics of a population by describing the population's overall pattern of variation. That is, we describe the shape of the distribution of population measurements. We often employ a sample of measurements taken from a population in order to infer what the population looks like.

Several graphical methods—the **stem-and-leaf display,** the **histogram,** and the **dot plot**—are often used to portray shapes of distributions.

Stem-and-leaf displays We illustrate how to construct stem-and-leaf displays in the following examples.

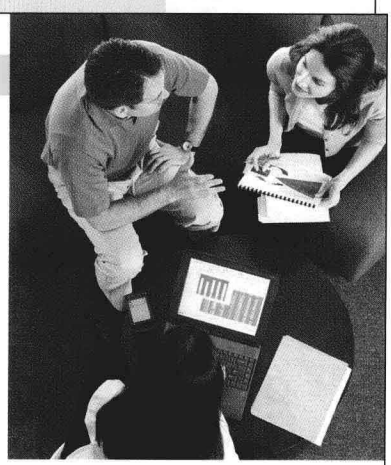
Example 2.1 The Car Mileage Case

Table 2.1 presents the sample of 49 gas mileages that have been obtained by the new midsize model in Example 1.4 (page 15). To graphically portray the pattern of variation in these mileages, we can construct a stem-and-leaf display. In order to do this, we first notice that the sample mileages range from 29.8 to 33.3. For this data we will (somewhat arbitrarily) construct a display having the first two digits of the mileages—the whole numbers 29, 30, 31, 32, and 33—as the **stems.** These are placed in a column on the left side of the display as follows:

29
30
31
32
33

CHAPTER 2

Descriptive Statistics



Chapter Outline

2.1 Describing the Shape of a Distribution
2.2 Describing Central Tendency
2.3 Measures of Variation
2.4 Percentiles, Quartiles, and Box-and-Whiskers Displays
2.5 Describing Qualitative Data
2.6 Using Scatter Plots to Study Relationships between Variables (Optional)
2.7 Misleading Graphs and Charts (Optional)
2.8 Weighted Means and Grouped Data (Optional)
2.9 The Geometric Mean (Optional)

Case Studies

The text provides a unique use of case studies that span individual chapters and groups of chapters. Cases are used to introduce the concepts, to demonstrate the methods, and to provide students with motivating exercises. These case studies help students see how statistics is used in business and can be used to improve processes.

Student Friendly Presentation

The authors have made improvements throughout the text to make learning easier for students. The following examples highlight some of these improvements.

Step-by-Step Hypothesis Testing Approach

This approach consists of a seven-step procedure designed to break hypothesis testing down into small, easy to understand steps. This procedure is used in almost all the examples in Chapters 8 and 9 and can be applied by students throughout the remainder of the text where hypothesis testing is done.

Testing a "greater than" alternative hypothesis by using a rejection point rule In Section 8.1 we explained how to set up appropriate null and alternative hypotheses. We also discussed how to specify a value for α, the probability of a Type I error (also called the **level of significance**) of the hypothesis test, and we introduced the idea of a test statistic. We can use these concepts to begin developing a seven step hypothesis testing procedure. We will introduce these steps in the context of the trash bag case and testing a "greater than" alternative hypothesis.

Step 1: State the null hypothesis H_0 and the alternative hypothesis H_a. In the trash bag case, we will test H_0: $\mu \leq 50$ versus H_a: $\mu > 50$. Here, μ is the mean breaking strength of the new trash bag.

Step 2: Specify the level of significance α. The television network will run the commercial stating that the new trash bag is stronger than the former bag if we can reject H_0: $\mu \leq 50$ in favor of H_a: $\mu > 50$ by setting α equal to .05.

Step 3: Select the test statistic. In order to test H_0: $\mu \leq 50$ versus H_a: $\mu > 50$, we will test the modified null hypothesis H_0: $\mu = 50$ versus H_a: $\mu > 50$. The idea here is that if there is sufficient evidence to reject the hypothesis that μ equals 50 in favor of $\mu > 50$, then there is certainly also sufficient evidence to reject the hypothesis that μ is less than or equal to 50. In order to test H_0: $\mu = 50$ versus H_a: $\mu > 50$, we will randomly select a sample of $n = 40$ new trash bags and calculate the mean $\bar{x}$ of the breaking strengths of these bags. We will then utilize the **test statistic**

$$z = \frac{\bar{x} - 50}{\sigma_{\bar{x}}} = \frac{\bar{x} - 50}{\sigma/\sqrt{n}}$$

A positive value of this test statistic results from an $\bar{x}$ that is greater than 50 and thus provides evidence against H_0: $\mu = 50$ and in favor of H_a: $\mu > 50$.

Step 4: Determine the rejection point rule for deciding whether to reject H_0. To decide how large the test statistic must be to reject H_0 in favor of H_a by setting the probability of a Type I error equal to α, we do the following:

- Place the probability of a Type I error, α, in the right-hand tail of the standard normal curve and use the normal table (see Table A.3, page 824) to find the normal point z_α. Here z_α, which we call a **rejection point** (or **critical point**), is the point on the horizontal axis under the standard normal curve that gives a right-hand tail area equal to α.
- **Reject H_0: $\mu = 50$ in favor of H_a: $\mu > 50$ if and only if the test statistic z is greater than the rejection point z_α.** (This is the **rejection point rule.**)

Greater Accessibility of Continuing Cases

Each time a continuing case is revisited, any needed computer output and, whenever possible, relevant background information is included with the current case discussion. Consequently, students seldom need to refer back to previously covered material in order to grasp the content included in a given case segment.

A 99 percent confidence interval for p is

$$\left[\hat{p} \pm z_{.005}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] = \left[.063 \pm 2.575\sqrt{\frac{(.063)(.937)}{1000}}\right]$$
$$= [.063 \pm .0198]$$
$$= [.0432, .0828]$$

The upper limits of both the 95 percent and 99 percent intervals are less than .10. Therefore, we have very strong evidence that the true proportion p of all current purchasers who would stop buying the cheese spread is less than .10. Based on this result, it seems reasonable to use the new spout.

BI

In the cheese spread example, a sample of 1,000 purchasers gives us a 95 percent confidence interval for p—[.063 ± .0151]—with a reasonably small margin of error of .0151. Generally speaking, quite a large sample is needed in order to make the margin of error in a confidence interval for p reasonably small. The next two examples demonstrate that a sample size of 200, which most people would consider quite large, does not necessarily give a 95 percent confidence interval for p with a small margin of error.

Business Improvement

Business improvement applications are identified by icons and are highlighted—to indicate when important business improvement conclusions have been reached using statistical analysis.

Appendix 11.3 ■ Simple Linear Regression Analysis Using MegaStat

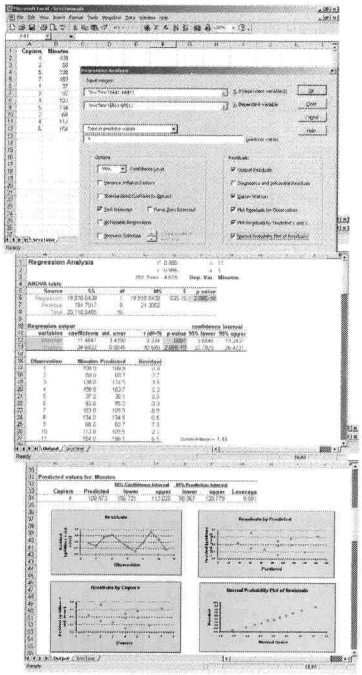
The instructions in this section begin by describing the entry of data into an Excel worksheet. Alternatively, the data may be loaded directly from the data disk included with the text. The appropriate data file name is given at the top of each instruction block. Please refer to Appendix 1.2 for further information about entering data and saving and printing results in Excel. Please refer to Appendix 1.3 for more information about using MegaStat.

Simple linear regression for the service time data in Figure 11.14 on page 476 (data file: SrvcTime.xls):

- Enter the service time data (page 456) with the numbers of copiers serviced in column A (with label Copiers) and the service times in column B (with label Minutes).
- Select **MegaStat : Correlation/Regression : Regression Analysis.**
- In the Regression Analysis dialog box, click in the "Independent variables" box and use the AutoExpand feature to enter the range A1.A12.
- Click in the "Dependent variable" box and use the AutoExpand feature to enter the range B1.B12.
- Check the appropriate Options and Residuals check boxes as follows:
  1. Check "Test Intercept" to include a y-intercept and to test its significance.
  2. Check "Output Residuals" to obtain a list of the model residuals.
  3. Check "Plot Residuals by Observation" and "Plot Residuals by Predicted Y and X" to obtain residual plots versus time, versus the predicted values of y, and versus the values of the independent variable.
  4. Check "Normal Probability Plot of Residuals" to obtain a normal plot.
  5. Check "Durbin-Watson" to obtain the Durbin–Watson statistic.

To obtain a **point prediction** of *y* when four computers will be serviced (as well as a confidence interval and prediction interval):

- Click on the drop-down menu above the Predictor Values box and select "Type in predictor values."
- Type the value of the independent variable for which a prediction is desired (here equal to 4) into the "predictor values" box.
- Select a desired level of confidence (here 95%) from the Confidence Level drop-down menu or type in a value.
- Click OK in the Regression Analysis dialog box.



Excel/MINITAB/MegaStat Tutorials

The end of chapter appendices contain helpful tutorials that teach students how to carry out statistical analysis using Excel, MINITAB, and MegaStat. These tutorials include step-by-step instructions for performing almost every type of statistical method presented in the book. For additional help, video tutorials for Excel, MINITAB, and MegaStat are provided on the Student CD-ROM.

Excel, MINITAB, and MegaStat Output

Throughout the text, Excel, MINITAB, and MegaStat outputs illustrate how statistical analysis is done electronically.

FIGURE 2.4 MegaStat Output of a Relative Frequency Histogram of the 49 Mileages

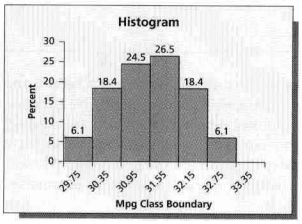


FIGURE 2.5 MINITAB Output of a Relative Frequency Histogram of the 100 Cell Phone Usages

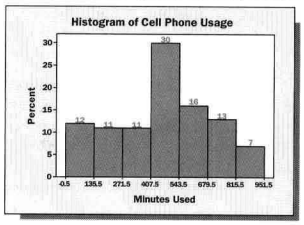
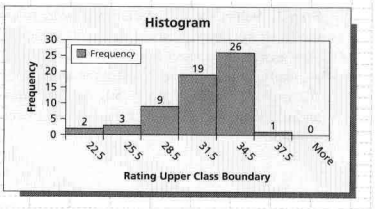
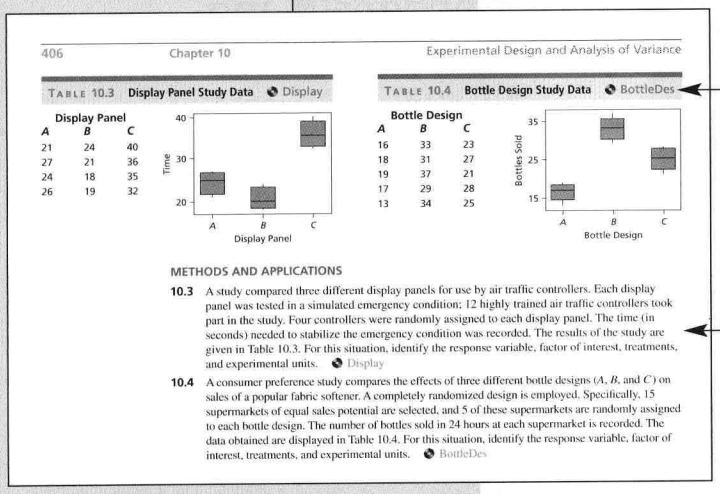


FIGURE 2.6 Excel Output of a Frequency Histogram of 60 Bottle Design Ratings

	A	B
1	Rating	Frequency
2	22.5	2
3	25.5	3
4	28.5	9
5	31.5	19
6	34.5	26
7	37.5	1
8	More	0





406 Chapter 10 Experimental Design and Analysis of Variance

TABLE 10.3 Display Panel Study Data Display

Display Panel		
A	B	C
21	24	40
27	21	36
24	18	35
26	19	32

TABLE 10.4 Bottle Design Study Data BottleDes

Bottle Design		
A	B	C
16	33	23
18	31	27
19	37	21
17	29	28
13	34	25

METHODS AND APPLICATIONS

10.3 A study compared three different display panels for use by air traffic controllers. Each display panel was tested in a simulated emergency condition; 12 highly trained air traffic controllers took part in the study. Four controllers were randomly assigned to each display panel. The time (in seconds) needed to stabilize the emergency condition was recorded. The results of the study are given in Table 10.3. For this situation, identify the response variable, factor of interest, treatments, and experimental units. Display

10.4 A consumer preference study compares the effects of three different bottle designs (*A*, *B*, and *C*) on sales of a popular fabric softener. A completely randomized design is employed. Specifically, 15 supermarkets of equal sales potential are selected, and 5 of these supermarkets are randomly assigned to each bottle design. The number of bottles sold in 24 hours at each supermarket is recorded. The data obtained are displayed in Table 10.4. For this situation, identify the response variable, factor of interest, treatments, and experimental units. BottleDes

Exercises...

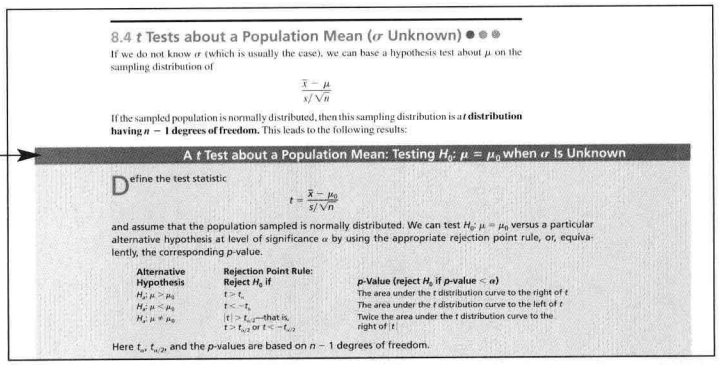
There are over 1000 exercises in the text. Many use real data from the current business literature. Data sets on the Student CD-ROM are identified by icon in the text. Within each chapter, exercises are broken into two parts—"Concepts" and "Methods and Applications." The methods and applications exercises vary in rigor from routine calculations to fairly sophisticated case study analysis. In addition, there are Internet exercises to help students make use of the Internet for gathering and using real data and supplementary exercises at the ends of chapters.

...And More Exercises

are found on the student text's website.

Boxed Equations, Formulas, and Definitions

Each chapter contains easy-to-find boxes that will help students identify and understand the key ideas in the chapter.



8.4 *t* Tests about a Population Mean (σ Unknown)

If we do not know σ (which is usually the case), we can base a hypothesis test about μ on the sampling distribution of

$$\frac{\bar{x} - \mu}{s/\sqrt{n}}$$

If the sampled population is normally distributed, then this sampling distribution is a ***t* distribution having $n - 1$ degrees of freedom.** This leads to the following results:

A *t* Test about a Population Mean: Testing H_0: $\mu = \mu_0$ when σ Is Unknown

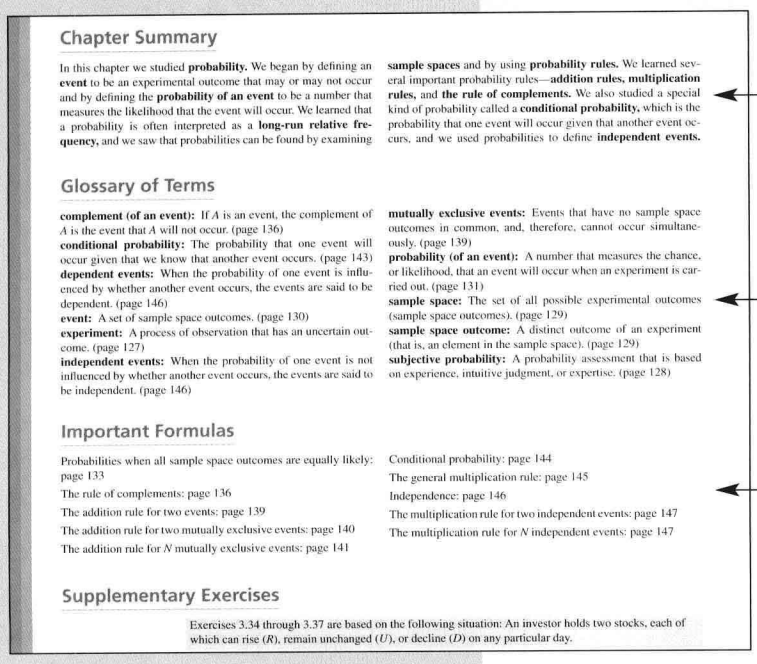
Define the test statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

and assume that the population sampled is normally distributed. We can test H_0: $\mu = \mu_0$ versus a particular alternative hypothesis at level of significance α by using the appropriate rejection point rule, or, equivalently, the corresponding *p*-value.

Alternative Hypothesis	Rejection Point Rule: Reject H_0 if	*p*-Value (reject H_0 if *p*-value $< \alpha$)
H_a: $\mu > \mu_0$	$t > t_\alpha$	The area under the *t* distribution curve to the right of *t*
H_a: $\mu < \mu_0$	$t < -t_\alpha$	The area under the *t* distribution curve to the left of *t*
H_a: $\mu \neq \mu_0$	$\lvert t \rvert > t_{\alpha/2}$—that is, $t > t_{\alpha/2}$ or $t < -t_{\alpha/2}$	Twice the area under the *t* distribution curve to the right of $\lvert t \rvert$

Here t_α, $t_{\alpha/2}$, and the *p*-values are based on $n - 1$ degrees of freedom.



Chapter Summary

In this chapter we studied **probability.** We began by defining an **event** to be an experimental outcome that may or may not occur and by defining the **probability of an event** to be a number that measures the likelihood that the event will occur. We learned that a probability is often interpreted as a **long-run relative frequency,** and we saw that probabilities can be found by examining **sample spaces** and by using **probability rules.** We learned several important probability rules—**addition rules, multiplication rules,** and **the rule of complements.** We also studied a special kind of probability called a **conditional probability,** which is the probability that one event will occur given that another event occurs, and we used probabilities to define **independent events.**

Glossary of Terms

complement (of an event): If *A* is an event, the complement of *A* is the event that *A* will not occur. (page 136)
conditional probability: The probability that one event will occur given that we know that another event occurs. (page 143)
dependent events: When the probability of one event is influenced by whether another event occurs, the events are said to be dependent. (page 146)
event: A set of sample space outcomes. (page 130)
experiment: A process of observation that has an uncertain outcome. (page 127)
independent events: When the probability of one event is not influenced by whether another event occurs, the events are said to be independent. (page 146)
mutually exclusive events: Events that have no sample space outcomes in common, and, therefore, cannot occur simultaneously. (page 139)
probability (of an event): A number that measures the chance, or likelihood, that an event will occur when an experiment is carried out. (page 131)
sample space: The set of all possible experimental outcomes (sample space outcomes). (page 129)
sample space outcome: A distinct outcome of an experiment (that is, an element in the sample space). (page 129)
subjective probability: A probability assessment that is based on experience, intuitive judgment, or expertise. (page 128)

Important Formulas

Probabilities when all sample space outcomes are equally likely: page 133
The rule of complements: page 136
The addition rule for two events: page 139
The addition rule for two mutually exclusive events: page 140
The addition rule for *N* mutually exclusive events: page 141
Conditional probability: page 144
The general multiplication rule: page 145
Independence: page 146
The multiplication rule for two independent events: page 147
The multiplication rule for *N* independent events: page 147

Supplementary Exercises

Exercises 3.34 through 3.37 are based on the following situation: An investor holds two stocks, each of which can rise (*R*), remain unchanged (*U*), or decline (*D*) on any particular day.

Chapter Ending Material

The end of each chapter includes a chapter summary, a comprehensive glossary of terms, and important formula references. The examples shown here are from Chapter 3, Probability.