STATISTICS FOR BUSINESS & ECONOMICS

MARK L. BERENSON—DAVID M. LEVINE

Statistics for Business and Economics

Mark L. Berenson

Department of Statistics and Computer Information Systems
and Department of Health Care Administration
Baruch College, City University of New York

David M. Levine

Department of Statistics and Computer Information Systems
Baruch College, City University of New York

Prentice Hall
Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging-in-Publication Data

Berenson, Mark L.,
Statistics for business and economics / Mark L.
Berenson, David M. Levine.
p. cm.
Abridged ed. of: Basic business statistics. 4th ed. ©1989.
ISBN 0–13–319096–X
1. Commercial statistics. 2. Statistics. I. Levine, David M.,
II. Berenson, Mark L. Basic business statistics.
III. Title.
HF1017.B38 1990
519.5—dc20 89–23084
CIP

Editorial/production supervision: Rachel J. Witty, Letter Perfect, Inc.
Interior design: Andy Zutis
Cover design: Photo Plus Art, Celine Brandes
Manufacturing buyer: Peter Havens

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

ISBN 0-13-319096-X

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

To our wives,
Rhoda Berenson and **Marilyn Levine**
and to our children,
Kathy Berenson Balto, Lori Berenson, and **Sharyn Levine**

Preface

This text is an abridged version of the Fourth Edition of *Basic Business Statistics: Concepts and Applications*. It is intended for a one-semester course in business or economic statistics. When planning any textbook, the authors must decide how the text will differ from those already available and what contribution it will make to the field of study. These issues are resolved in several ways in this text.

Survey Database

A set of data based upon the results of a survey is included. Specifically, the data set constitutes the characteristics of 233 sampled single-family houses. This survey, used throughout the text, serves as a means of integrating the various topics, permitting a cohesive study of descriptive statistics, probability, statistical inference, and regression analysis. The use of an actual survey, examined from beginning to end, will aid the student in conducting basic research in future courses, theses, or occupational situations.

Problems

This text contains numerous problems at the end of sections as well as at the end of chapters. Many of these problems apply to realistic situations (using real data whenever possible) in various fields including accounting, economics, management, marketing, and public administration. The Answers to Selected Problems (indicated by the symbol

●) appear at the end of the text. In addition, detailed case studies are included at the end of each of seven chapters. Moreover, at the end of most chapters, a series of database exercises pertaining to the survey is presented and indicated by the symbol □.

ACTION (➠) and Lightning (🗲) Problems

Statistics is a living, breathing subject. It is not mere numbers crunching! Action problems enhance literacy by asking the student to write letters, memos and reports, and prepare talks. Lightning (🗲) problems are particularly thought-provoking or have no "exact" answer. Together, action and lightning problems force the reader to think and better enable the reader to understand the utility of statistical analysis as an aid to the solution of real problems in an organizational setting.

Computer Packages

A major feature of this text is the discussion of the use of such computer packages as SAS, SPSS^X, Minitab, STATGRAPHICS, and MYSTAT. Not only is output from each of these packages illustrated throughout the text, but the use of the computer as a tool for assisting in the decision-making process is interwoven in the various chapters. Extensive coverage is given to plotting data, interpreting computer output, and evaluating the assumptions of the particular statistical techniques, thereby simulating the approach used by a statistician in conducting an actual statistical analysis.

Applications to Quality and Productivity Management

It is widely recognized that the field of business has entered a new economic age, one in which competition from all over the world must be faced. This text provides detailed coverage (particularly in Chapter 16) of the management philosophy of W. Edwards Deming, the person who deserves a share of the credit for helping Japan to become known as a high-quality producer, and contains in-depth coverage of statistical process control, a topic of increasing importance in both the service and manufacturing industries.

Modern Methods

A sixth important feature of this text is the inclusion of recently developed methods in the field of statistics. As examples, exploratory data analysis techniques, dot charts, Pareto diagrams, digidot plots, and supertables are discussed (Chapters 3 and 4); a *p*-value approach to hypothesis testing is used (Chapters 11–13) and residual analysis in regression is covered (Chapter 14).

Flexibility

A final important feature is flexibility. This book is written for students taking either a one-semester or one-quarter basic statistics course. There are numerous ways in which the instructor could adapt material to meet specific needs. For example, a traditional course covering elements of both descriptive and inferential statistics might consist of Chapters 1–5 (except Section 4.3), 6 (Sections 6.1 to 6.8), 7, 8 (8.1–8.3), 9, 10 (except Section 10.7), 11, 14 (Sections 14.1–14.7) and selections from Chapters 12, 13 and 16. However, the material is organized so that instructors who do not wish to devote time to the questionnaire could skim or omit the data-collection phase of the survey (Chapter 2) and begin the course with Chapter 3 merely by using the results of the survey. If such a course were to emphasize inferential statistics, it could skim Chapters 1 and 2 and really begin with sections of Chapters 3, 4, 5, and 6 before covering Chapters 7 through 14. On the other hand, if the course were to primarily emphasize descriptive statistics, probability, and quality and productivity, then Chapters 1 through 6, 8 (Sections 8.1–8.3), 14 (Sections 14.1–14.7), 15, and 16 (Sections 16.1–16.7) would be included. Regardless of which topics are stressed by each instructor, the primary emphasis is on the concepts of basic statistical methods and their applications to business and economics.

It is our hope and anticipation that the unique approaches taken in this textbook will make the study of basic statistics more meaningful, rewarding, and comprehensible for all readers.

We are extremely grateful to the many organizations and companies that generously allowed us to use their actual data for developing problems and examples throughout our text. In particular, we would like to cite Time Inc. (publisher of *Fortune*), CBS Inc. (publisher of *Road & Track*), Dun & Bradstreet Publications Corporation (publisher of *Dun's Business Month*), Crain Communications (publisher of *Crain's New York Business*), The Condé Nast Publications, Inc. (publisher of *Street and Smith's*), Consumer's Union (publisher of *Consumer Reports*), Standard & Poor's Corporation (publisher of *Standard N.Y.S.E. Stock Reports*), Brian Joiner (Joiner Associates, Inc.), M.I.T. Center for Advanced Engineering Study, CEEPress Books, The American Association of University Professors (publisher of *Academe*), OECD (publisher of *Observer*), Gale Research, Inc., and The College and University Personnel Association (CUPA). Moreover, we would like to thank the *Biometrika* Trustees, The Rand Corporation, and the American Society for Testing and Materials for their kind permission to publish various tables in Appendix E. Furthermore, we would like to acknowledge the SAS Institute, SPSS Inc., Minitab Inc., STSC Inc., and SYSTAT Inc. for their permission to present computer output throughout the text.

We wish to express our thanks to Jeanne Hoeting, Valerie Ashton, Dennis Hogan, and Kate Moore of the editorial staff at Prentice-Hall and Rachel J. Witty of Letter Perfect, Inc., for their continued encouragement. We also wish to thank Professors Lou Bianco, Southeastern Massachusetts University, John McKenzie, Babson College, and David Fluharty, PEP Industries, for their constructive comments during the writing of this textbook. Finally, we would like to thank our wives and children for their patience, understanding, love, and assistance in making this book a reality. It is to them that we dedicate this book.

MARK L. BERENSON
DAVID M. LEVINE

Contents

1 Introduction

1.1 WHAT IS MODERN STATISTICS?

A century ago H. G. Wells commented that "statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write." Each day of our lives we are exposed to a wide assortment of numerical information pertaining to such phenomena as stock market activity, unemployment rates, medical research findings, opinion poll results, weather forecasts, and sports data. Frequently, such information has a profound impact on our lives.

> The subject of **modern statistics** encompasses the collection, presentation, and characterization of information to assist in both data analysis and the decision-making process.

1.2 THE GROWTH AND DEVELOPMENT OF MODERN STATISTICS

Historically, the growth and development of modern statistics can be traced to two separate phenomena—the needs of government to collect information on its citizenry (see References 4, 5, 13, and 14) and the development of the mathematics of probability theory.

Data have been collected throughout recorded history. During the Egyptian, Greek, and Roman civilizations information was obtained primarily for the purposes of taxation and military conscription. In the Middle Ages, church institutions often kept records concerning births, deaths, and marriages. In America, various records were kept during colonial times (see Reference 14), and beginning in 1790 the Federal Constitution required the taking of a census every ten years. Today these data are used for many purposes, including congressional apportionment and the allocation of federal funds.

1.2.1 Descriptive Statistics

These and other needs for data on a nationwide basis were closely intertwined with the development of descriptive statistics.

> **Descriptive statistics** can be defined as those methods involving the collection, presentation, and characterization of a set of data in order to properly describe the various features of that set of data.

Although descriptive statistical methods are important for presenting and characterizing information (see Chapters 2 through 5), it has been the development of inferential statistical methods as an outgrowth of probability theory that has led to the wide application of statistics in all fields of research today.

1.2.2 Inferential Statistics

The initial impetus for formulation of the mathematics of probability theory came from the investigation of games of chance during the Renaissance. The foundations of the subject of probability can be traced back to the middle of the seventeenth century in the correspondence between the mathematician Pascal and the gambler Chevalier de Mere (see References 5, 8, and 9). These and other developments by such mathematicians as Bernoulli, DeMoivre, and Gauss were the forerunners of the subject of inferential statistics. However, it has only been since the turn of this century that statisticians such as Pearson, Fisher, Gosset, Neyman, Wald, and Tukey pioneered in the development of the methods of inferential statistics that are so widely applied in so many fields today.

> **Inferential statistics** can be defined as those methods that make possible the estimation of a characteristic of a population or the making of a decision concerning a population based only on sample results.

To clarify this, a few more definitions are necessary.

> A **population** (or **universe**) is the totality of items or things under consideration.
>
> A **sample** is the portion of the population that is selected for analysis.
>
> A **parameter** is a summary measure that is computed to describe a characteristic of an entire population.
>
> A **statistic** is a summary measure that is computed to describe a characteristic from only a sample of the population.

Thus, one major aspect of inferential statistics is the process of using sample statistics to draw conclusions about the true population parameters.

The need for inferential statistical methods derives from the need for sampling. As a population becomes large, it is usually too costly, too time consuming, and too cumbersome to obtain our information from the entire population. Decisions pertaining to the population's characteristics have to be based on the information contained in a sample of that population. Probability theory provides the link by ascertaining the likelihood that the results from the sample reflect the results from the population.

These ideas can be clearly seen in the example of a political poll. If the pollster wishes to estimate the percentage of the votes a candidate will receive in a particular

election, he or she will not interview each of the thousands (or even millions) of voters. Instead, a sample of voters will be selected. Based on the outcome from the sample, conclusions will be drawn concerning the entire population of voters. Appended to these conclusions will be a probability statement specifying the likelihood or confidence that the results from the sample reflect the true voting behavior in the population.

1.3 ENUMERATIVE VERSUS ANALYTICAL STUDIES

The examples that we have just provided concerning inferential statistical methods allow us to make an important distinction between two types of statistical studies that are undertaken: **enumerative** studies and **analytical** studies.

> **Enumerative** studies involve decision-making regarding a population and/or its characteristics.

The political poll is an example of an enumerative study since its objectives are to provide estimates of population characteristics and to take some action on that population. The listing of all the units (such as the registered voters) which belong to the population is called the **frame** (see Section 2.7) and provides the basis for the selection of the sample. Thus the focus of the enumerative study is on the counting (or measuring) of outcomes obtained from the frame.

> **Analytical** studies involve taking some action on a process to improve performance in the future.

The study of the outcomes of a manufacturing process taken over time is an example of an analytical study. The focus of an analytical study is on predictions of future process behavior and on understanding and improving a process. In an analytical study there is no identifiable universe, as in an enumerative study, and therefore there is also no frame. Perhaps we can highlight the distinction between enumerative and analytical studies by referring to Figure 1.1.

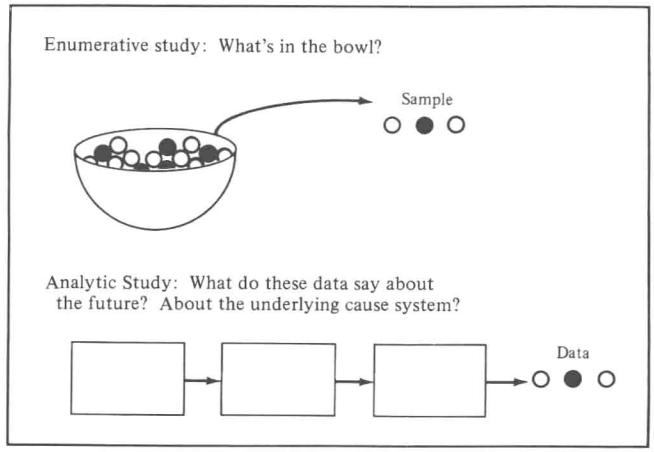

FIGURE 1.1
Enumerative vs. Analytic Study
SOURCE: Brian L. Joiner, "Transformation of the American Style of Teaching Statistics," (Madison, WI: Joiner Associates, 1986) Copyright © by Joiner Associates Inc., 1986 All Rights Reserved.

In the enumerative study the bowl represents the population. The questions of interest revolve around the issue of "What's in the bowl?". In the analytical study the questions of interest revolve around how the data collected (often over a period of time) can be used for future predictions.

The distinction between enumerative and analytical studies is an important one since the methods that have been developed primarily for enumerative studies may be misleading or incorrect for analytical studies (see References 1 and 2). In this textbook we shall develop methods that are appropriate for these two different types of studies. Some of the methods are appropriate for either type of study (see Chapters 1–8 and parts of Chapter 14). Other methods are appropriate primarily for enumerative studies (see Chapters 9–13) or primarily for analytical studies (see Chapters 15 and 16).

1.4 THE ROLE OF COMPUTER PACKAGES IN STATISTICS

In the last decade revolutionary changes have taken place in our society owing to the rapidly expanding application of computer technology. In particular, the increasing availability of the microcomputer has brought these changes into our businesses, our homes, and our classrooms. Statisticians and researchers in various disciplines can now use the computer in the application of statistical methods, particularly when large databases or highly complex computational procedures are involved.

Over this past decade groups of standardized programs have been assembled as a collection or "package" by various software developers (see References 6–7, 10–12). Multiple users of a large mainframe computer have been able to share the use of these programs with different data. Moreover, in the last five years there has been a widespread development of these statistical packages for use on a microcomputer. Certain packages that were previously available only for mainframe and minicomputers (such as SAS, SPSS^x and Minitab) are now available in microcomputer versions, and many new packages have been specifically developed for particular brands of microcomputers.

1.5 WHY STUDY MODERN STATISTICS?

Along with the major advances in information processing, the use of statistical methods as an aid to data analysis and decision making has grown dramatically over the past decade and will continue to grow in the future. For example, the study and application of statistics are an integral part of the successful Japanese management approach to total quality control (see References 2 and 3). Some of the techniques, which are used at all organizational levels from company president to line worker, will be described in Chapter 16 of this text.

By studying modern statistics, we will obtain an appreciation for and an understanding of those techniques that are used on the numerical information we encounter in both our professional and nonprofessional lives. The concepts and methods described in this text provide a fundamental background in the subject of modern statistics and its applications in a wide variety of disciplines.

Problems

To answer the questions below you may wish to go to your library and use the following sources of reference:

Indexes

Business Periodical Index
New York Times Index
Wall Street Journal Index

Business Magazines

Business Week
Forbes
Fortune

General Magazines

Newsweek
Time
U.S. News & World Report

Newspapers

U.S.A. Today
Local newspapers

General Information

Statistical Abstract of the United States

For Problems 1.1 to 1.6, specify the general problem to be solved, the specific inference to be made, what the population is, and (if you are describing the results of an actual published study) what the weaknesses of the study might be. Where appropriate, tell what parameters are of primary interest and what statistics are used to arrrive at a conclusion.

1.1 Describe three applications of statistics to economics or finance.

1.2 Describe three applications of statistics to sports.

1.3 Describe three applications of statistics to political science or public administration.

1.4 Describe three applications of statistics to advertising.

1.5 Describe three applications of statistics to marketing research.

1.6 Describe three applications of statistics to medical research or health care administration.

References

1. DEMING, W. E. "On Probability as a Basis for Action" *American Statistician*, 29, 1975, pp. 146–152.
2. DEMING, W. E. *Out of the Crisis* (Cambridge, MA.: Massachusetts Institute of Technology Center for Advanced Engineering Study, 1986).
3. ISHIKAWA, K., *What is Total Quality Control? The Japanese Way* (Translated by D. J. Lu) (Englewood Cliffs, N.J.: Prentice-Hall, 1985).
4. KENDALL, M. G., AND R. L. PLACKETT, eds., *Studies in the History of Statistics and Probability*, Vol. II (London: Charles W. Griffin, 1977).