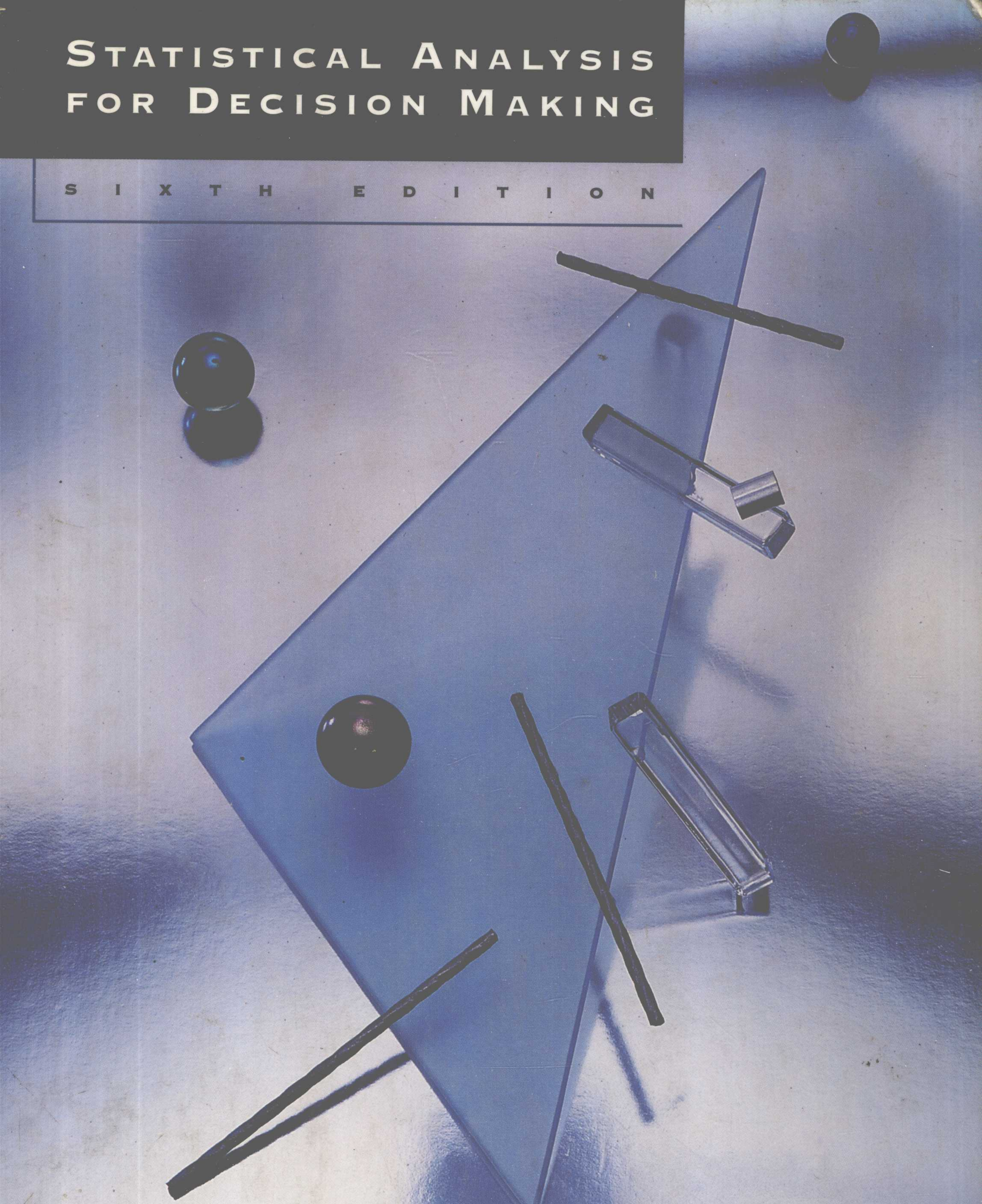
# STATISTICAL ANALYSIS FOR DECISION MAKING

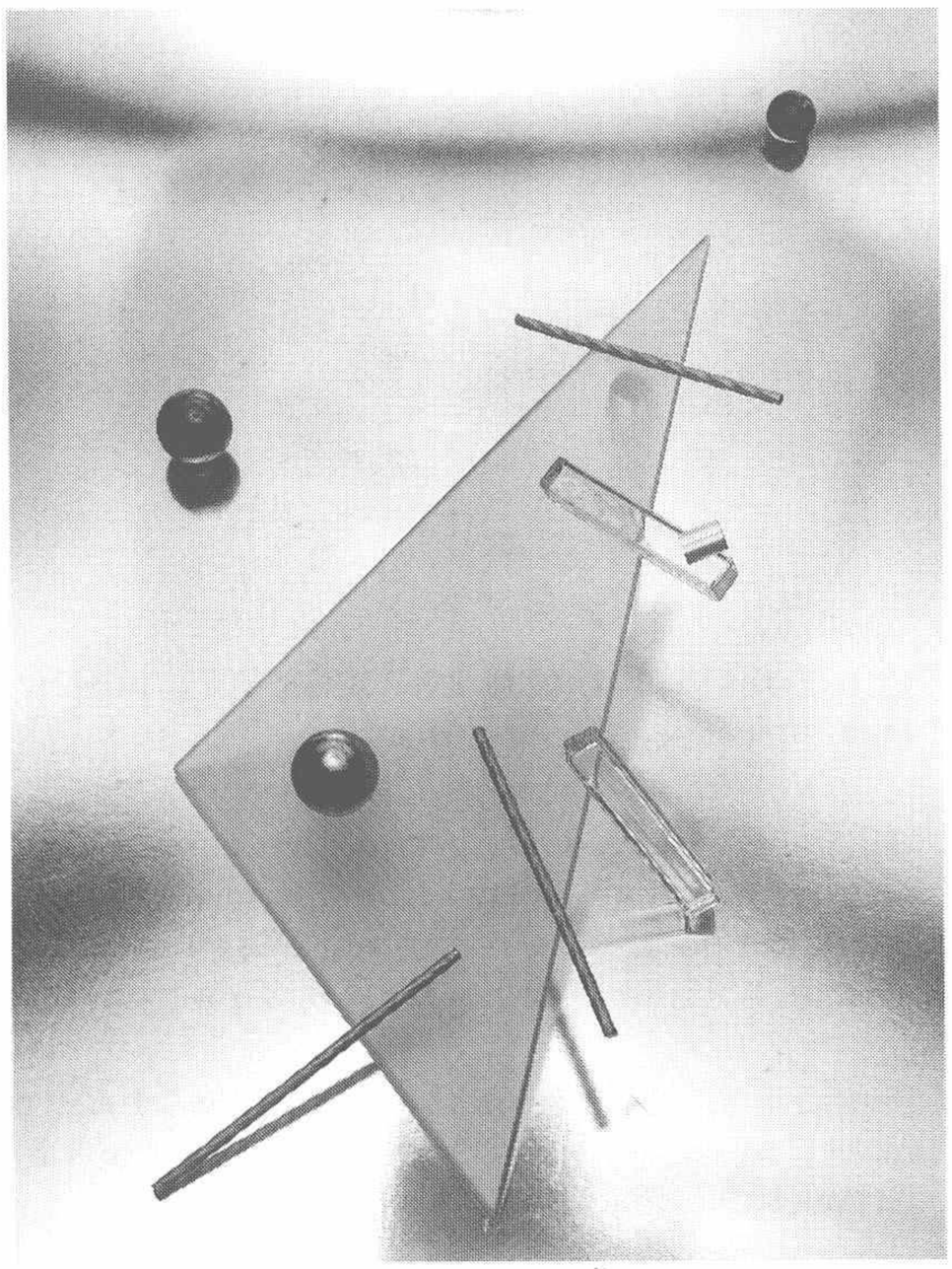
SIXTH EDITION



MORRIS HAMBURG PEG YOUNG

# Statistical Analysi for Decision Makin

SIXTH EDITION



MORRIS HAMBURG
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLANIA

PEG YOUNG
DEPARTMENT OF VETERANS AFFAIRS

THE DRYDEN PRESS
HARCOURT BRACE COLLEGE PUBLISHERS

FORT WORTH PHILADELPHIA SAN DIEGO NEW YORK ORLANDO AUSTIN SAN ANTONIO
TORONTO MONTREAL LONDON SYDNEY TOKYO

**Publisher**
Liz Widdicombe

**Acquisitions Editor**
Scott Isenberg

**Developmental Editor**
Millicent Treloar

**Project Editor**
Jim Patterson

**Production Manager**
Kelly Cordes

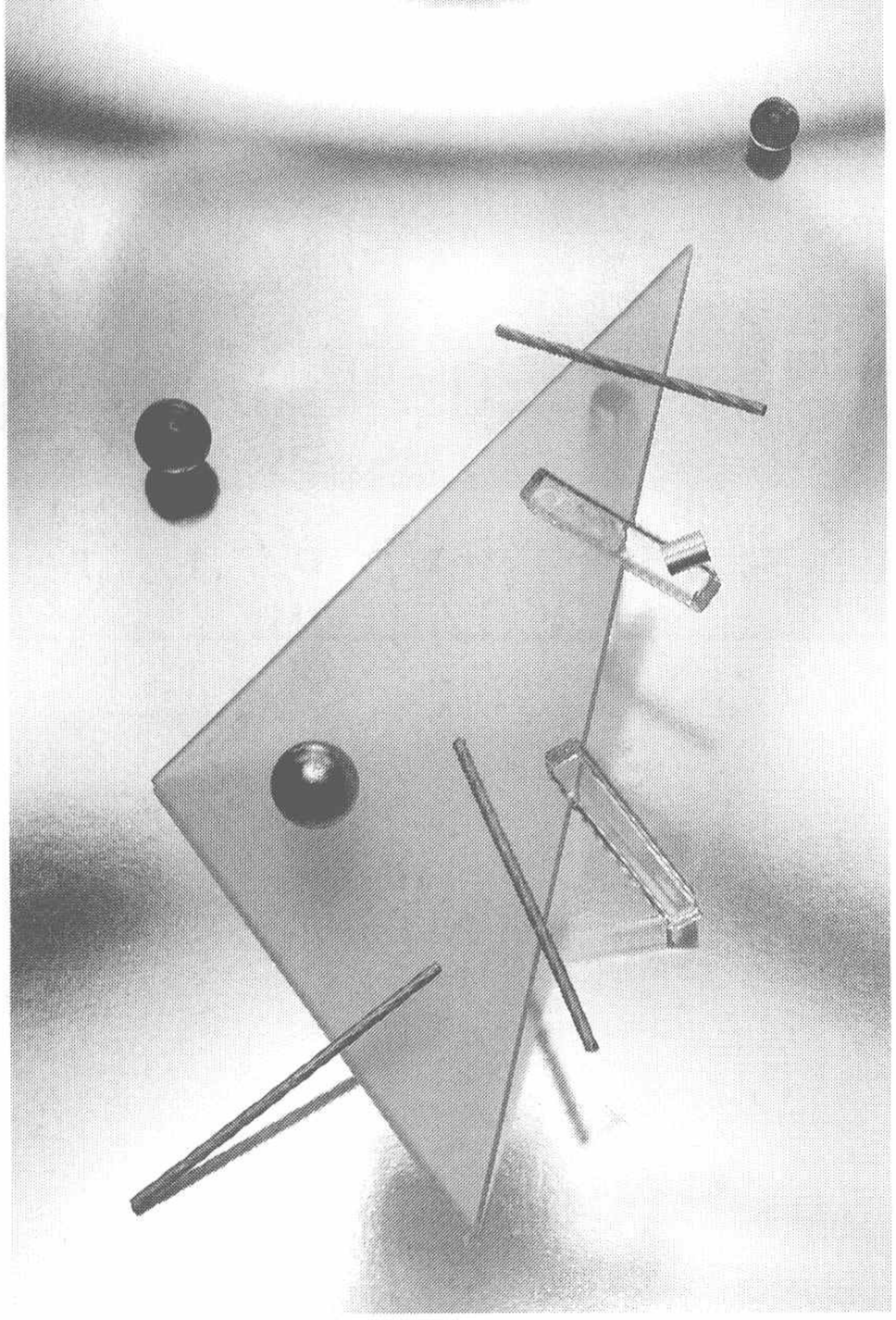
**Designer**
Linda Miller

**Developmental Assistant**
Van Strength

**Compositor**
Beacon Graphics

**Text Type**
Palatino

**Cover Image**
© 1993 Stock Yard





*Address for Editorial Correspondence*
The Dryden Press, 301 Commerce Street, Suite 3700, Fort Worth, TX 76102

*Address for Orders*
The Dryden Press, 6277 Sea Harbor Drive, Orlando, FL 32887
1-800-782-4479, or 1-800-433-0001 (in Florida)

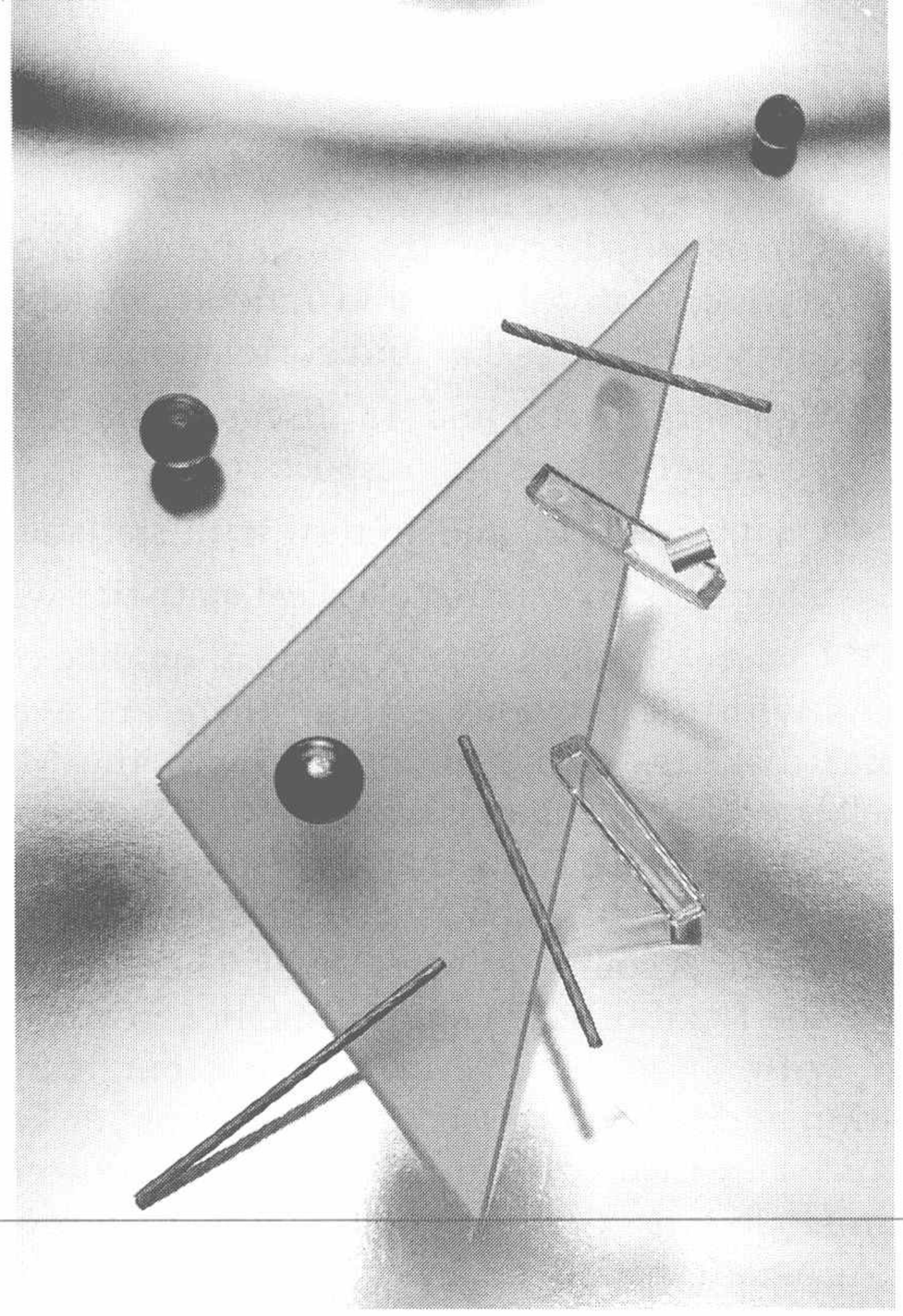
ISBN: 0-03-096914-X

Library of Congress Catalogue Number: 93-71261

Printed in the United States of America

4 5 6 7 8 9 0 1 2 3 069 9 8 7 6 5 4 3 2 1

The Dryden Press
Harcourt Brace College Publishers



# PREFACE

Although numerous changes and additions have been made in this sixth edition of *Statistical Analysis for Decision Making,* the basic objective of the book remains the same—to present the fundamental concepts of statistics in a clear and straightforward way, thus assisting the student to develop critical judgment and decision-making ability using quantitative tools. We try to convey the idea that statistics is an exciting field that deals with a scientific method of obtaining, analyzing, and using numerical data to draw inferences and to make decisions.

This textbook is designed primarily for a first course in statistics for students in business and economics, although the presentation of topics and development of methodology are appropriate for students in public administration, social sciences, and liberal arts as well.

Because most readers of this book will be consumers rather than producers of statistical analyses, emphasis is placed on the reasoning and logic underlying statistical concepts and techniques and on the interpretation and use of statistical results. The only mathematical background required is high school algebra. Emphasis is placed on showing the *power* of modern statistical reasoning and the *scope* and *versatility* of the methods, without bogging down the discussion in mathematical formalities. Mathematical derivations have not been included in the body of the text, but some are given in footnotes and in Appendix C. Points for which an explanation in calculus is practical and helpful also appear in footnotes.

The structure of the book gives the instructor considerable flexibility in designing a course. With appropriate selection of chapters, this book may be used for either a one- or a two-semester course.

- Chapters 1–10 cover the fundamentals of classical statistics—descriptive statistics, probability and random variables, sampling, statistical inference, and regression and correlation analysis.
- Chapters 11, 12, and 13 discuss time series, index numbers, and nonparametric statistics, respectively.
- Chapters 14, 15, and 16 deal with statistical decision analysis.
- Chapter 17 compares classical statistics with Bayesian decision analysis.
- Chapter 18 deals with statistical quality control.

As with the previous edition, there are available with this edition computer diskettes designed for use with the Minitab and EasyStat software packages. The Minitab diskete contains exercises that appear at the end of most chapters, data sets that correspond to the exercises, and solutions to the Minitab exercises. The EasyStat diskette also contains problems and data sets and has an accompanying manual. Therefore, instructors using the textbook have the flexibility of using Minitab, a powerful commercial software package with the capability of carrying out many statistical procedures both within and beyond the scope of this book, or EasyStat, a menu-driven educational statistical software package, less powerful and comprehensive than Minitab, but user-friendly and simple to apply.

Although the advantages of integrating the use of computers with a statistics book are evident, please note that most of the exercises in this book may be solved with pencil, paper, and hand calculators. Thus, the sixth edition should appeal to instructors who favor heavy use of computers in their courses, as well as to those who feel that one better understands the fundamental theory and methodology of statistical analysis by thinking through and carrying out calculations than by placing sole reliance on the computer.

A large number of changes and revisions have been made in this edition. Many of the exercises are new. New review exercises appear at the ends of the first thirteen chapters, except for Chapter 4. New "immediate feedback" exercises have been added in most chapters. These somewhat more elementary exercises give the student drill and confidence before proceeding to more difficult problems. Review exercises for Chapters 1 through 5, 6 through 8, and 9 and 10 generally require the student to use material from more than one section or chapter. New such exercises have been added, focusing on requiring the selection of appropriate techniques.

Cases have been placed at the ends of the chapters to give the students an added flavor of real-world applications.

In Chapter 1, fuller treatment has been given to the types of data, scales of measurement, and characteristics of data used in statistical analyses. Then comments in the beginnings of many of the later chapters connect these data classification ideas to the statistical procedures discussed in those chapters.

In Chapter 4, a section on survey sampling has been added that deals with stratified sampling, cluster sampling, and systematic sampling.

In keeping with a modern trend toward greater emphasis on approaches dealing with the design and implementation of quality control and quality improvement procedures, we have added a new Chapter 18 on Statistical Quality Control. Although numerous examples and exercises in other chapters deal with quality control applications, the added chapter provides an integrated introduction to some of the philosophical and statistical approaches to quality control and quality improvement systems.

As in the fifth edition, Key Terms and Key Formulas appear at the end of most chapters. The endpapers contain four of the most useful tables from the

text and Appendix A, repeated for easy reference. A glossary of symbols follows Appendix E, with each symbol keyed to the section in which that symbol is introduced.

Several supplementary pedagogical aids for this sixth edition are available in computer diskettes or separate publications:

- A computer diskette for use with Minitab software contains computer exercises, solutions, and data sets. Many exercises are also shown in the text at the end of appropriate chapters; and as an aid to students, instructions for Minitab use immediately follow the Preface. Some of the data sets were derived from the Citibase Series and Slater-Hall Information Products and are larger and more realistic than the data in typical textbook exercises.
- A diskette for the EasyStat computer program operates within the framework of a spreadsheet and includes a data set for 1,000 urban families. Students can select random samples for carrying out exercises. The Minitab and EasyStat programs are available for students who have access to an IBM-PC. An EasyStat manual is available to students.
- The Study Guide provides supplementary exercises, as well as worked-out problems and explanations.
- Solutions to all of the text exercises and Minitab exercises appear in the Solutions Manual. (Only the even-numbered text solutions appear in the book.)
- Additional exercises (with solutions) suitable for homework or testing are available in a Testbook, which also has a set of multiple-choice questions for examinations.

As in earlier editions thanks are again extended to the many organizations and individuals that have made contributions to this book. Special mention goes to colleagues in the Statistics Department of the Wharton School of the University of Pennsylvania for constructive comments and recommendations based on their experience teaching from previous editions and from interacting with students. We also express our grateful appreciation to Richard W. Gideon, president of Dick Gideon Enterprises, for the television and other data and estimates that he made available for us in the chapter on time series.

We are very grateful to those who assisted in the development of earlier editions and special thanks go to Dinesh Bhoj (Rutgers University), Cheryl L. Eavey (Washington University), Frank C. Fuller (Jacksonville State University), Mark P. Karscig (Central Missouri State University), Jean-Claude Léon (The Catholic University of America), John Mark (Boston University-London), John P. Phillips (Boston University), Stephen Replogle (Arkansas State University), Fatollah Salimian (Salisbury State University), Kishor Thanawala (Villanova University), Ivan Weinel (University of Missouri), whose excellent constructive comments and suggestions were critical to the sixth edition.

The comments of teachers and students who have used previous editions have continued to be very helpful. My appreciation is especially extended to the following instructors:

Richard W. Andrews (University of Michigan), David Ashley (University of Missouri), Joe D. Berry (Marion Military Institute), Eileen C. Boardman (Colorado State University), James M. Cannon (Florida Institute of Technology), William D. Coffey (St. Edward's University), F. Damanpour (La Salle

College), Maynard M. Dolecheck (Northeast Louisiana University), Fred H. Dorner (Trinity University), Linda W. Dudycha (University of Wisconsin), David Eichelsdorfer (Gannon University), David Frew (Gannon University), Carl Gordon (California State University-Sacramento), James C. Goodwin (University of Richmond), Charles R. Gorman (La Salle College), I. Greenberg (George Mason University), John B. Guerard, Jr. (Lehigh University), Jack A. Holt (University of Virginia), Yutaka Horiba (Tulane University), Gary Kern (University of Virginia), Burton J. Kleinman (Widener University), Ming-Te Lu (St. Cloud State University), William F. Matlack (University of Pittsburgh), Patrick McKeown (University of Georgia), Rudy Moore (Rosemont College), Elias Alphonse Parent (George Mason University), Chander T. Rajaratnam (Rutgers University), Manuel G. Russon (University of South Alabama), S. R. Ruth (George Mason University), James R. Schaefer (University of Wisconsin), Stanley R. Schultz (Cleveland State University), Marion G. Sobol (Southern Methodist University), William L. Steglitz (Nichols College), William R. Stewart, Jr. (College of William and Mary), Kishor Thanawala (Villanova University), Chris A. Theodore (Boston University), Robert J. Thorton (Lehigh University), Ebenge Usip (Youngstown State University), Charles F. Warnock (Colorado State University), Roger L. Wright (University of Michigan), and Thomas A. Yancey (University of Illinois).

Again an especially warm expression of gratitude goes to Bryan Sayer and Zhongquan Zhou for their expert work in the preparation of diskettes and related material for the EasyStat and Minitab computer programs, respectively. Lastly, our thanks are extended to the fine professional staff at the Dryden Press: Scott Isenberg, Acquisitions Editor; Millicent Treloar, Developmental Editor; Jim Patterson, Project Editor; Linda Miller, Designer; Kelly Cordes, Production Manager; and Van Strength, Developmental Assistant.

We are grateful to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S., and to Longman Group, Ltd., London, for permission to reprint Tables III and IV from their book *Statistical Tables for Biological, Agricultural, and Medical Research* (Sixth Edition, 1974). My gratitude also goes to the other authors and publishers whose generous permission to reprint tables or excerpts from tables has been acknowledged at appropriate places.

As in previous editions, Morris Hamburg dedicates this book to his wife, June, and his children, Barbara and Neil. Peg Young dedicates this book to her husband, Michael Fries.

MORRIS HAMBURG
PEG YOUNG

*An Author's Note:*

After serving as sole author for the previous editions of this book, it is a privilege to have Peg Young join me as coauthor on this sixth edition. She is responsible for most of the changes and revisions that appear herein. In my opinion, this edition has benefitted greatly from Dr. Young's fresh approaches and her wise and balanced judgments. From the viewpoint of the student, this edition can undoubtedly be characterized as more "user friendly" than its predecessors.

MORRIS HAMBURG

# INSTRUCTIONS FOR

# MINITAB USE

If you are using a personal computer and have Minitab software installed in drive C, you can enter Minitab by typing **cd Minitab** and then **Minitab**.

Once you are in Minitab, you can familiarize yourself with the program by using the HELP command. When working problems, you will need to retrieve data sets from the disk that accompanies the Solutions Manual by typing:

**MTB > retrieve 'a:xxxx'**

where **xxxx** is the name of the data set specified in the problem.

To output your own solution, use Minitab command

**MTB > outfile 'yyyy.zzz'**

where **yyyy.zzz** is a filename of your choice. When you have finished a problem, type:

**MTB > nooutfile**

to close your own solution file. To modify and print your solution file, you have to exit Minitab. To exit Minitab, simply type:

**MTB > stop**

You can print and edit your solution file in many ways depending on the computer software with which you are working. We suggest that you use a work processing software, for example, WordPerfect, to edit your solution file first. You can also use the software to print the solution file. On the other hand, to print in DOS, simply type:

**C:\ > print 'yyyy.zzz'**

Listed below are filenames for solutions and data sets. The chapter numbers in the solutions correspond to the same chapter numbers in the text.

*Chapter Numbers and Filenames of Solutions*

| CHAPTER | FILENAME (SOLUTIONS) |
|---|---|
| 1 | **abort.ans** |
| 1 | **mstat.ans** |
| 1 | **bankroa.ans** |

*Chapter Numbers and Filenames of Solutions (continued)*

| CHAPTER | FILENAME (SOLUTIONS) |
|---|---|
| 3 | **discrete.ans** |
| 5 | **random.ans** |
| Review | **rev1-5.ans** |
| 6 | **roa.ans** |
| 7 | **testing.ans** |
| 8 | **table.ans** |
| Review | **rev6-8.ans** |
| 9 | **regress.ans** |
| 10 | **multiple.ans** |
| 11 | **series.ans** |
| 13 | **nonpar.ans** |

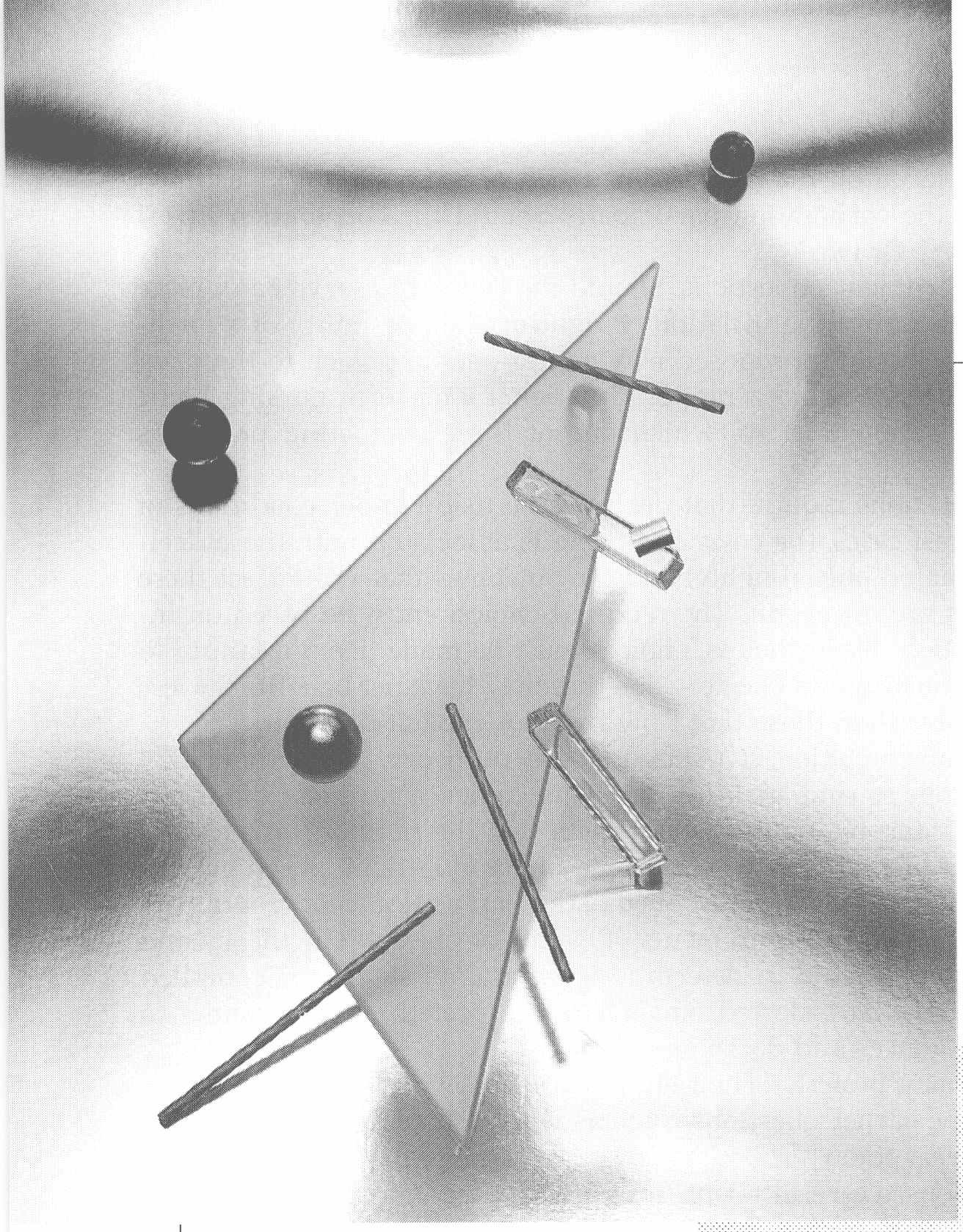
## A LIST OF FILENAMES (DATA)

| | | |
|---|---|---|
| **ability.mtw** | **abort.mtw** | **bankroa.mtw** |
| **butter.mtw** | **customer.mtw** | **databank.mtw** |
| **dmotor.mtw** | **dow.mtw** | **dummy.mtw** |
| **gnp.mtw** | **invest.mtw** | **lab.mtw** |
| **lifetime.mtw** | **mstat.mtw** | **profit.mtw** |
| **project.mtw** | **rating.mtw** | **roa.mtw** |
| **salary.mtw** | **sales.mtw** | **vote.mtw** |

## REMARKS

- These instructions, problems, and solutions were developed with PC Minitab, Release 6.1.1. If you use an earlier or later release, there will be some very minor differences in the appearance of some output.
- Some sets of the data were taken from the Citibase Series and Slater-Hall Information Products. Citibase Series is a databank containing a large number of mainly U.S. macroeconomic time-series data. Slater-Hall Information Products records both economic and demographic statistics at the following four levels: national, state, county, and city.
- Although some instructions for the use of many Minitab commands are given in the computer problems, a Minitab handbook may be useful for problem solving. Also, you can always type **HELP** in Minitab to get information when you are uncertain. Two Minitab handbooks are listed below for your reference:

1. Ryan, Barbara F., Brian L. Joiner, and Thomas A. Ryan. *Minitab Handbook,* 2d ed. Boston: Duxbury Press, 1985.
2. Miller, Robert B. *Minitab Handbook for Business and Economics,* 2d ed. Boston: PWS-Kent Publishing Company, 1988.



# INTRODUCTION

*What do the following questions have in common? Such questions occur in every aspect of our lives. Personal questions: Whom should I marry? What field should I choose as my life's work? Which automobile should I buy? What clothes should I wear today?*

*A private business corporation's questions: Should the corporation buy or lease this building? Should this*

new product be placed on the market? In which of these capital investment proposals should the corporation participate? Should the corporation invest now or wait a year? Or two?

A national government's questions: Should the Federal Reserve Bank place more emphasis on controlling inflation or fighting the possibility of a forthcoming recession? Should a proposed new educational approach to the drug problem be instituted? Should a proposed national long-term nursing home care program be adopted? If so, which one of three competing programs should be selected?

All of these questions require that decisions be made under conditions of uncertainty. In most cases, the costs and benefits associated with these decisions can be estimated only roughly, and the outcomes that will affect these costs and benefits are uncertain. The required choices must be based on incomplete information. Nevertheless, choices must be made. Even a failure to make a decision constitutes a choice—one that may have net benefits far less or far more desirable than those that flow from an explicit decision.

In its modern interpretation, *statistics* is a body of theory and methodology for drawing inferences and making decisions under conditions of uncertainty. From this interpretation, it would seem that the field of statistics has much to contribute toward answering some of the questions posed earlier, and indeed it does. The raw material of statistics is *statistical data* or numbers that represent counts and measurements of events or objects. The theory and methodology of statistics aid in determining what data should be compiled and how they should be collected, analyzed, interpreted, and presented to make the best inferences and decisions.

However, you may protest, "The field of statistics is not going to help me answer some of the earlier questions, such as whom should I marry or what clothes should I wear today?"

For some decisions, a careful scientific approach based on quantitative data does not seem very appropriate. Many decisions require a substantial component of intuitive judgment. Nonetheless, in many areas of human activity, statistical analysis provides a solid foundation for decision making. Statistical concepts and methods bring a logical, objective, and systematic approach to decision making in social, governmental, business, and scientific problems. They are not meant to replace intuition and common-sense judgments. On the contrary, they assist in structuring a problem and in bringing the application of judgment to it.

As an example, let us return to the question that probably seems least likely to be subject to a meaningful solution by statistical reasoning: "Whom should I marry?" In structuring this problem, let us consider the following line of thought. If you marry a (the?) "right" person, you have made no error. Also, if you do not marry someone you should not have married, you have again made a correct decision. On the other hand, if you fail to marry someone with whom you would have been happy, you have made an error. Or, if you marry someone with whom your future life falls far short of a state of connubial bliss, you have also made an error. After studying the subject of hypothesis testing, you will recognize the difference between the two types of errors. Which type of error is more serious? If you are considering the possibility of marrying someone, which is the best course of action, to marry or not to marry? With which action is the higher expected "payoff" associated? Although the problem and questions here are raised tongue in cheek and without the expectation of a serious answer, they illustrate a way in which you

may begin to structure any problem and apply judgment to it. We cannot dispute that someone's romantic sensibilities may be offended by any attempt to use the suggested method of analysis to solve this problem. Our purpose is merely to indicate that even so unlikely a problem for quantitative analysis may be structured in a framework that can assist decision making. It is currently recognized that the field of statistics aids in providing the basis for arriving at rational decisions regarding a tremendous variety of matters related to business, public affairs, science, and other fields. In many of these matters, such quantitative analysis has been found to be not only applicable but also extremely helpful.

The most widely known statistical methods are those that summarize numerical data in terms of **averages** and other descriptive measures.

> If we are interested in the incomes of a group of 1,000 families chosen at random in a particular city, important characteristics of these incomes may be described by calculating an average income and a measure of the spread, or dispersion, of these incomes around the average.

The essence of modern statistics, however, is the theory and the methodology of **drawing inferences** that extend beyond the particular set of data examined and of **making decisions** based on appropriate analyses of such inferential data.

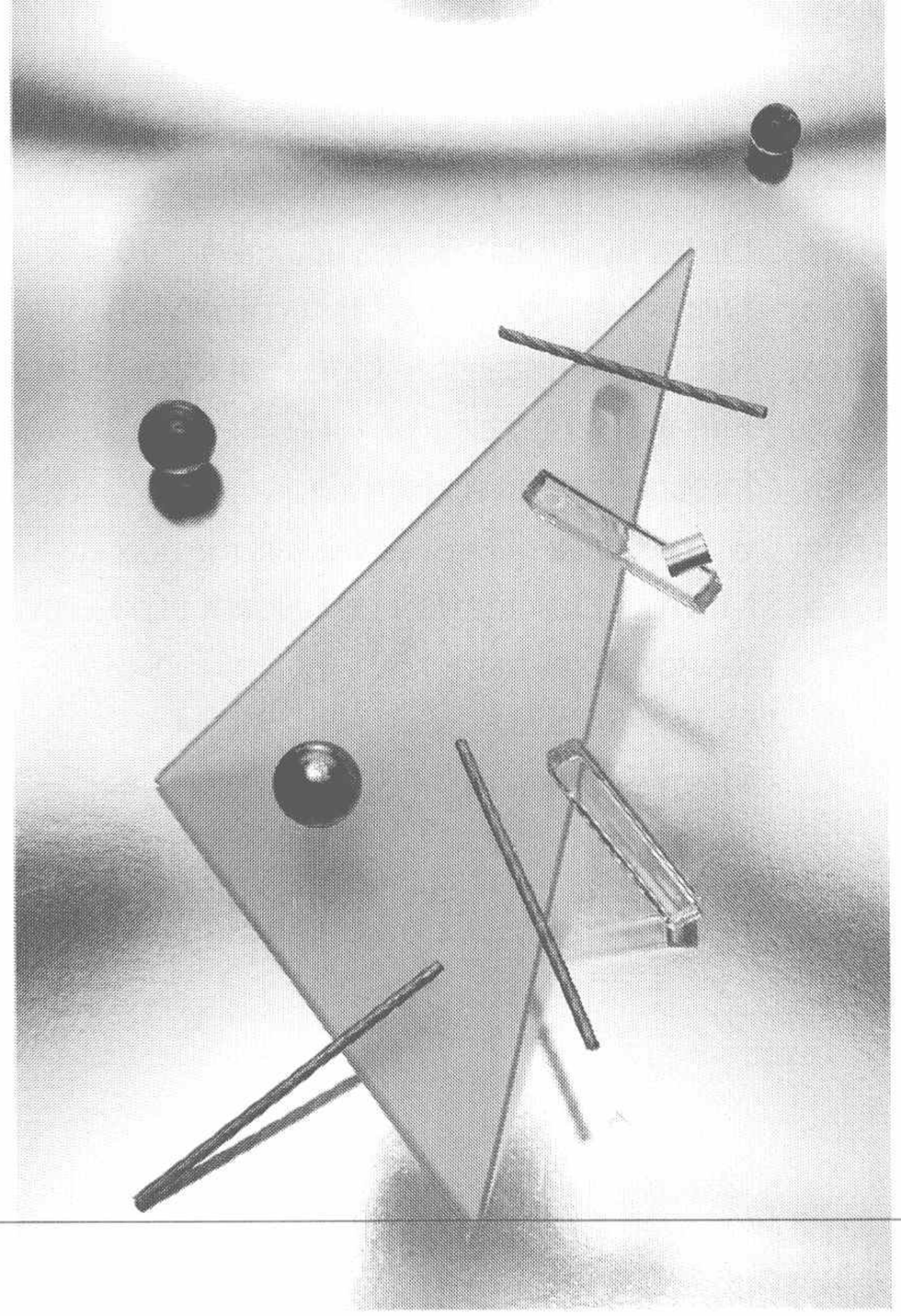
> We are probably not so much interested in the incomes of the particular 1,000 families included in the sample as we are interested in an inference about the incomes of all families in the city from which the sample was drawn.

- Such an inference might be in the form of a **test of a hypothesis** that the average income of all families in the city is $35,000 or less.
- The inference could also be in the form of a single figure, an **estimate** of the average income of all families in the city based on the average income observed in the sample of 1,000 families.
- Or the marketing department of a company may want the information in order to decide among different types of advertising programs based on the identification of the city as a low-, medium-, or high-income area.

The mathematical theory of probability provides the logical framework for the mental leap from the sample of data studied to the inference about all families in the city and for decisions such as the type of advertising program to be used.

The preceding example presents three points:

- We may have wanted an inference about the incomes of *all* families in the city. The totality of families in the city (or, more generally, the totality of the elements about which the inference is desired) is referred to in statistics as the **universe** or **population**.
- Because it would have been too expensive and too time-consuming to obtain the income data for every family in the city, only the sample of 1,000 families was observed. The 1,000 families, which represent a collection of only some elements of the universe, are referred to as a **sample**.
- In statistics, sample data are collected in order to make **inferences** or **decisions** concerning the populations from which samples are drawn.



# CONTENTS