STATISTICAL METHODS

FOR BUSINESS AND ECONOMICS

Roger C. Pfaffenberger
James H. Patterson

Statistical methods

For business and economics

ROGER C. PFAFFENBERGER
University of Maryland

and

JAMES H. PATTERSON
University of Missouri

1977

Richard D. Irwin, Inc. Homewood, Illinois 60430
IRWIN-DORSEY LIMITED Georgetown, Ontario L7G 4B3

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

ISBN 0-256-01797-2
Library of Congress Catalog Card No. 76–28895

Printed in the United States of America

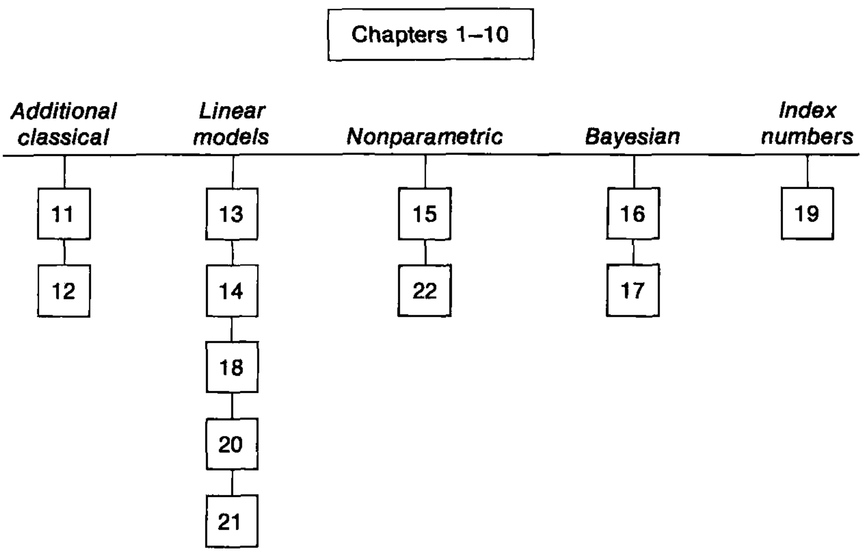
Preface

A note to the instructor

IN MANY DISCIPLINES, including business and economics, there is an increasing awareness of the need to understand basic statistical methods for successful careers in industry, government, and in self-employed positions. With the development and widespread use of computers today, the accumulation and analysis of data for the purposes of decision making is, for many enterprises, among a set of standard operating procedures. While only a few individuals in an organization may be directly involved with the accumulation, dissemination, and statistical analysis of data, many others must be able to interpret the analyses for use in making decisions.

Because of the widespread coverage of statistical methods in nonmathematics curricula, it is no longer possible or, indeed, proper to assume that individuals taking the typical introductory statistics courses are well-trained mathematically. The demand for statistical training has increased so dramatically in the last few years and has affected so many disciplines that most courses *cannot* require more than a knowledge of algebra on the part of the student. We believe the modest mathematics prerequisites to be not only a pragmatic necessity, but also appropriate for the needs of the majority of students. Most students simply do not need the level of knowledge that would enable them to pursue training for a career in statistics. Rather, they need to be exposed to the fundamental ideas of statistical analysis in a competent, but mathematically nonrigorous, fashion. This is the spirit of this text. Among our goals is to motivate students by suggesting the potential importance statistical analysis has in pursuing successful careers. Further, we have emphasized the development of the logic of statistical methods in a thoroughly competent way without resorting to mathematical derivations or proofs in writing this text. We firmly believe that the extensive use of examples, the careful discussion of assumptions, and the proper interpretation of results can competently and effectively replace a mathematically rigorous development of introductory statistical methods. While a knowledge of calculus is not required for the study of this text, we have on occasion noted the ideas of the derivative and the integral to emphasize key points. An example is the delineation of the connection between the summing process for discrete values and the integrating process for a continuum of values. For those students with a calculus background, the learning experience will, we hope, be somewhat enhanced by our treatment. Those students without calculus training are, however, placed at no disadvantage in learning the material involved. We will leave the training of statisticians to mathematics and statistics departments!

The text is designed to be quite flexible. A typical one-semester course would cover Chapters 1–10, the core of classical statistical inference, and perhaps two or three additional chapters which may be selected to satisfy specific course objectives. The chart below indicates the interdependence of chapters. In a two-semester course, we suggest Chapters 1–10, the linear models chapters (13, 14, 18, 20, 21) and a selection of two or three addi-



tional chapters. We have included what we feel are the most important special methods in this text. Nonparametric statistics (Chapter 22) is increasingly important in business and social sciences applications of statistics where, all too often, the data will not support the appropriate classical statistics method assumptions. Forecasting (Chapter 21) also has become an increasingly used statistical tool in recent years. Most financial and economic institutions depend heavily on the accuracy of forecasts for their success. Bayesian statistics (Chapters 16 and 17) is finding an important place in business applications of statistics, where prior information relevant to a current problem is on hand and must be revised when additional information becomes available.

An Instructor's Manual is available which gives complete solutions to all exercises at the end of the chapters and provides helpful guidance to the instructor on how to present the text material. Computer programs are additionally available from the authors to aid the instructor and the student by removing many of the calculations required to implement a particular statistical procedure and spend more time on the mastery of the concepts involved. These programs are FORTRAN based, and an advantage to using these particular FORTRAN programs is that the output conforms to the format presented in the text.

A note to the student

The competent interpretation of statistical data and analyses is possible only with an understanding of basic statistical concepts. *All research* that involves the collection of data inevitably leads to a statistical analysis of these data. A well-known research effort that has involved considerable statistical analysis in recent years is the study conducted under the authority of the Surgeon General of the United States to assess the effect of smoking on the health of an individual. While this study has received notoriety because of various interpretations of the results reported, it is in no way unique in the sense of using *statistical analysis to interpret data.* Today, the applications and reporting of statistical data abound, and we are frequently called upon to draw *our own* conclusions based upon an analysis of the reported results. It is not surprising, therefore, to find that many disciplines have added introductory statistics courses to their undergraduate curriculums in recent years.

It is the authors' view that statistics is best learned on first exposure by presenting an example of the application of the statistical concept immediately after its introduction. This should solidify an understanding of the statistical concept involved. Accordingly, most discussions of statistical techniques and concepts are followed by at least one example (with solution) of an application of the statistical technique discussed. The inclusion of these example problems and solutions accounts for the text being somewhat larger in size than several introductory statistics texts with which the student may be familiar. We feel that this increased size is warranted by the exposure to actual problem solution. The student is encouraged in reviewing the material in the text to attempt to solve several of the examples given in the various chapters and then compare his or her own results with those presented. This "self-test" feature of the book should provide the student with rapid feedback on an understanding of the material presented, and should lessen the time required for a mastery of the concepts involved. We have also highlighted important definitions, theorems, rules, and so on, and where possible have provided summary computation formulas at the end of selected chapters.

To assist the student in doing the exercises at the end of each chapter, we have provided the answers to selected end-of-chapter problems in the back of the text. Consistent with our belief that statistics can be better learned "by doing," the student is encouraged to work as many problems given at the end of each chapter as possible.

Acknowledgments

The material in this text has been class-tested over a number of years. We are particularly indebted to those students who used early versions of the manuscript—their reactions to the material and suggestions for improvement have been most helpful.

Several individuals have additionally contributed to the writing of this text. Particularly helpful has been Mildred Massey, who has reviewed more drafts of selected chapters than she probably wishes to remember! Her thoughtful analysis and the attention paid to our presentation of ideas in the earlier versions of the manuscript could only have resulted in a vastly improved text. Warren Boe reviewed several chapters in an earlier version of the manuscript and his comments were particularly helpful on the organization of selected chapters. We only regret that Professor Boe's schedule did not permit a final reading of the manuscript. Jack Hayya gave us several good ideas on the presentation of text material, particularly with reference to time series and forecasting. Robert Ressek provided us with an excellent review of the final manuscript, but unfortunately, we were able to incorporate only a portion of his remarks because of impending deadlines on text publication. Particularly noteworthy were his comments on the linear models chapters (Chapters 13, 14, 18, 20, and 21), a few of which did reach us in time to be incorporated in the text.

Special thanks are also extended to Marie Straka, Ildiko Takacs, and Sharalyn Bowersox for an absolutely superb job in typing several iterations of the text material, often on short notice.

We wish to acknowledge the contribution of Tim Kaage and John Thoeming in preparing the art work for the text. They were able to add clarity to many of the statistical concepts discussed in the text through the excellent use of illustrations to convey key ideas.

Finally, Professor Pfaffenberger wishes to express his special appreciation to H. O. Hartley, Director of the Institute of Statistics, Texas A&M University, whose remarkable knowledge of statistics and special gifts as a teacher afforded a rare educational opportunity to learn statistics from a master.

February 1977

ROGER C. PFAFFENBERGER
JAMES H. PATTERSON

Contents

Statistical methods
For business and economics

1

Introduction

1.1 EXAMPLES OF STATISTICAL PROBLEMS

The word "statistics" conveys a variety of meanings to people, many of which are inaccurate or, at the very least, misleading. To some, the word suggests only a plethora of mind-boggling tables, charts, and figures. Other people consider statistics to be an imposing form of mathematics. The use of the word certainly had an inauspicious beginning, as might be suspected from a cursory study of the word, for it was originally a term used to denote a collection of figures, graphs and the like which contained useful information for the state (primarily budget information such as taxation figures).

Used in the context of its original meaning, statistics generally refers to information about an activity or a process that is expressed in numbers listed in tables or illustrated in figures. But, since its early connotation, statistics has grown to encompass a larger role than presenting us with charts, graphs, and tables or figures. In a modern setting, statistics refers to the science of collecting, presenting, and analyzing numerical data. A statistician is a person who engages in one or more of the following tasks: (1) the clerical activities of tabulating, summarizing, and displaying statistical data; (2) analyzing data by using statistical methods, usually for the purposes of decision making; or (3) advancing the science of statistics by developing new and better analysis methods. The levels of expertise required by statisticians ranges from mastering simple clerical operations with data to advanced training in applied mathematics, and statisticians are needed at all levels.

The use of statistics has permeated almost every facet of our lives. The daily newspapers and the televised news reports supply us with numerous summaries of data such as stock market reports, financial summaries, and crime statistics—and with the results of statistical analyses—weather forecasts, political election outcome predictions, and so on.

Governments, businesses, and individuals collect statistical data required to carry out their activities efficiently and effectively. The rate at which statistical data are being collected is staggering and is primarily due to the realization that better decisions are possible with more information and, perhaps more importantly, to technological advances that have enabled the efficient collection and analysis of large bodies of data. The most important technological advance in this area has, of course, been the development of the electronic digital computer. Statistical concepts and methods, and the use of computers in statistical analyses, have affected virtually all disciplines—physics, engineering, economics, sociology, psychology, business, and others. In business and economics, the development and application of statistical methods have led to greater production efficiency, to better forecasting techniques, and to better management practices. It is becoming increasingly apparent that some knowledge of statistics and computers is essential for careers in economics, business, administration, and many other fields as well. To gain an appreciation for the breadth of applications of statistics to business and economic problems in particular, let us consider five examples.

Example 1.1

Most major city newspapers in the Sunday edition provide a summary of the New York Stock Exchange transactions during the previous week. For example, on a particular Sunday, we may find that shares (rounded to the nearest 100) of IBM (International Business Machines) were traded during the previous week with a weekly high of 263⅛ and a low of 252½ ($ per share). The value of a share at the end of the week is listed as 255⅝ with a change of 6⅜ from the closing value at the end of the prior week. This information is portrayed in the newspaper as in the excerpt from a typical Sunday copy of the *Washington Post* in Figure 1.1. Since it would be infeasible to list the values of *all* 382,100 shares traded, the weekly high, low, closing price, and net change *statistics* are used to summarize meaningfully the entire set of transactions. The statistics presented in these tables in the Sunday newspaper provide important information for investment decisions.

FIGURE 1.1 Excerpt from a summary of weekly N.Y. Stock Exchange transactions

1976				*Weekly*			
High	*Low*	*Stock*	*Volume in 100s*	*High*	*Low*	*Weekly closing*	*Change from last week*
8¼	4⅛	Interctl Div	70	8¼	7¼	8¼	+⅞
42½	24	IntrikeInc	737	42½	38¼	40⅜	+1⅞
264½	157¼	IBM	3821	263⅛	252½	255⅝	−6⅜
35⅛	22½	IntFlavF	1344	25⅜	23¼	23¼	−1⅞
30½	19¾	IntHarv	2249	28⅛	27⅛	27⅛	−¾

Source: *Washington Post*, February 29, 1976.

Example 1.2

Annually, the Bureau of the Census publishes the *Statistical Abstract of the United States* which contains economic and business data on a variety of subjects. The data are compiled from the U.S. census of the population (conducted every 10 years), annual samples drawn from the U.S. population by the Bureau, surveys conducted by other governmental agencies, and government supported research projects conducted by private organizations. In Table 1.1, extracted from the 1975 *Statistical Abstract,* statistics relating to the receipts and outlays of the federal government are presented.

Although the sources and functions are listed only by major sources, the table contains valuable information for business decision making. For example, the increase in health percentage outlay and decrease in defense percentage outlay may dictate a company's future strategies for new product and services development. The negative percentages recorded as outlays reflect the failure of the federal budget to balance in recent years.

TABLE 1.1 Federal budget receipts, by major source, and outlays, by function—percent distribution: 1960 to 1975 (for years ending June 30)

MAJOR SOURCE OR FUNCTION	1960	1965	1968	1969	1970	1971	1972	1973	1974	1975 est.
Total receipts, by source	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Individual income taxes	44.0	41.8	44.7	46.5	46.7	45.8	45.4	44.4	44.9	42.2
Corporation income taxes	23.2	21.8	18.7	19.5	16.9	14.2	15.4	15.6	14.6	13.8
Social ins. taxes and contributions	15.9	19.1	22.5	21.3	23.4	25.8	25.8	27.8	28.9	30.9
Excise taxes	12.6	12.5	9.2	8.1	8.1	8.8	7.4	7.0	6.4	7.2
Customs, estate, and gift taxes	2.9	3.6	3.3	3.1	3.1	3.4	4.2	3.5	3.2	3.1
Miscellaneous receipts	1.3	1.4	1.6	1.5	1.8	2.0	1.7	1.7	2.0	2.8
Total outlays, by function	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
National defense	49.0	41.0	44.4	43.5	40.3	36.3	33.4	30.5	29.3	27.2
International affairs	3.2	3.5	2.6	2.1	1.8	1.5	1.6	1.2	1.3	1.6
General science, space, and technology	.7	5.0	3.1	2.8	2.4	2.0	1.9	1.7	1.6	1.3
Agriculture	2.8	3.3	2.5	3.1	2.6	2.1	2.3	2.0	.8	.6
Natural resources, environment, and energy	2.0	2.5	2.0	1.9	1.8	2.1	2.2	2.2	2.4	3.0
Commerce and transportation	6.2	5.8	6.0	3.8	4.6	4.9	4.6	4.0	4.9	3.8
Community and regional develop.	.5	1.1	1.2	1.4	1.8	1.9	2.0	2.4	1.8	1.6
Education, manpower and social services	1.0	1.8	3.9	3.7	4.0	4.3	5.0	4.8	4.3	4.7
Health	.8	1.5	5.4	6.4	6.7	7.0	7.5	7.6	8.2	8.5
Income security	19.8	21.7	18.8	20.2	21.9	26.2	27.5	29.6	31.5	34.0
Veterans benefits and services	5.9	4.8	3.9	4.1	4.4	4.6	4.6	4.9	5.0	4.9
Law enforcement and justice	.4	.5	.4	.4	.5	.6	.7	.9	.9	1.0
General government	1.2	1.2	1.0	.9	1.0	1.0	1.1	1.1	1.2	.8
Revenue sharing and general purpose fiscal assistance	.2	.2	.2	.2	.2	.2	.2	2.9	2.5	2.2
Interest	9.0	8.7	7.7	8.5	9.3	9.3	8.9	9.2	10.5	10.0
Undistributed offsetting receipts	−2.7	−2.6	−3.1	−3.0	−3.3	−4.0	−3.5	−5.0	−6.2	−5.2

Source: The Statistical Abstract of the United States, 1975.

Caution should be exercised in utilizing compiled statistics such as those in Table 1.1. Although the *percentage* outlay for national defense decreases during the period from 1960 to 1975, the *total* expenditure on defense did not necessarily decrease. Indeed, the total federal outlays for defense increased from approximately $45 billion to $85 billion during this period. Thus, while the percentage expended for defense decreased, the total defense outlay increased from 1960 to 1975.

Example 1.3

In operations management, a primary concern is controlling the quality of the items being produced. If the product is a transistor radio battery, for example, we may be concerned with the longevity of the batteries. Suppose it is desired that at least 95 percent of the batteries last through at least 20 hours of continuous use. The actual percentage of batteries lasting more than 20 hours could be determined by inserting each and every battery produced into a transistor radio and recording its time to failure, but then there would be no batteries to sell. Rather, a manager may wisely decide in a day's production to pull every 100th battery off the production line, insert the sampled batteries in electrical test circuits and record their times to failure. The percentage of these batteries lasting through more than 20 hours of