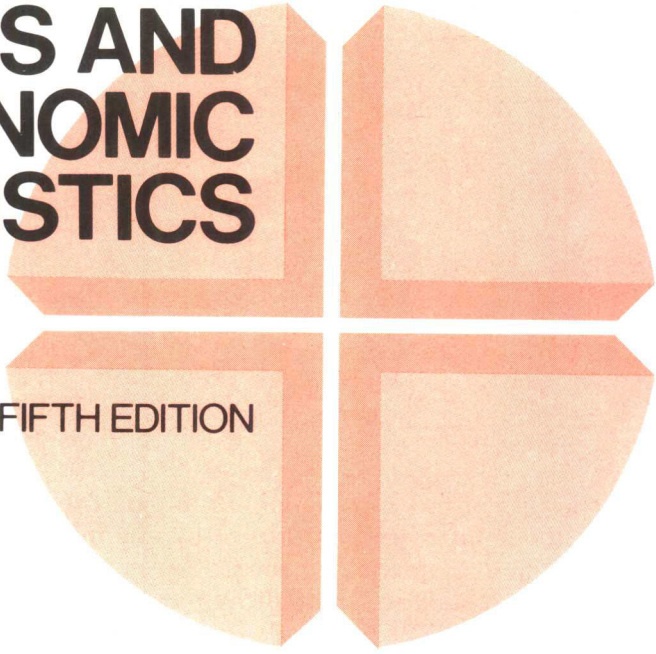
INTRODUCTION TO BUSINESS AND ECONOMIC STATISTICS

STOCKTON & CLARK • FIFTH EDITION

INTRODUCTION TO BUSINESS AND ECONOMIC STATISTICS



FIFTH EDITION

JOHN R. STOCKTON

Professor Emeritus of Business Statistics
Former Director, Bureau of Business Research
The University of Texas at Austin

CHARLES T. CLARK

Associate Professor of Business Statistics
The University of Texas at Austin

M24

Published by

SOUTH-WESTERN PUBLISHING CO.

CINCINNATI WEST CHICAGO, ILL. DALLAS PELHAM MANOR, N.Y.
PALO ALTO, CALIF. BRIGHTON, ENGLAND

ISBN: 0-538-13240-X

Library of Congress Catalog Card Number: 74-24472

34567Ki1098

Printed in the United States of America

Preface

The major objective of a textbook in elementary business statistics should be to inform students how people in business use statistics in making decisions and controlling the operations of a business. Information is needed on the sources of statistical data useful to the businessperson and on the methods of analysis that will make these data of maximum usefulness. This text attempts to describe in detail the important basic principles used by the business statistician and to supplement this description with numerous examples of how the methods have been used.

It has been kept in mind throughout the discussion that many students majoring in business administration are not highly trained mathematicians. Whenever possible a simple explanation of methods and theory has been used in preference to mathematical terminology. The sophisticated methods of mathematics may be preferred by those students with an adequate background, but a fairly large percentage of business students have gone no farther than an introductory course in calculus. Some have not even been introduced to calculus.

The coverage of the methods of analysis of data, published sources of data, and the collection of data are more complete than in the majority of statistics texts. This has been done in the belief that without an understanding of these basic principles the business student will not be equipped to make effective use of these data. Special emphasis is given to the analysis of time series and index numbers. The present turmoil in the world economy has increased to a considerable degree the interest of businesspersons and government officials in statistics of this nature. Unless he has a thorough understanding of the methods of measuring and forecasting changes in the various segments of the economy, the student of business has not received a well-rounded education.

It has been assumed that not all students taking a course in business statistics are familiar with probability. For the benefit of students who have not studied probability, a simple explanation of the most important theorems is given as a basis for the study of sampling and statistical inference. For students who are adequately prepared in probability, this chapter can serve as a brief review.

The major probability distributions are covered in expectation that the majority of business students will have had only a limited introduction to them.

Chapters 5 and 6 will serve as the theoretical basis for understanding the remaining chapters in Part 2, which are devoted to an introduction to the collection of sample data and to some of the most generally used methods of statistical inference. It is impossible to do more than introduce the student to statistical inference in a first course. The methods described were selected on the basis of their frequency of use and the fact that they are typical of the type of analysis that can be carried out in this manner. This subject matter is taught much more satisfactorily in an advanced course for students who have a considerable interest in statistical methods and who are attracted to the field of research.

Chapter 7 has been retained without much change to give the basic theory for estimates from random samples and the computation of confidence intervals. This chapter does not go into the widely used methods of stratified and cluster sampling since it is believed that this subject matter is more properly covered in an advanced text.

A new feature of this edition is the increased emphasis on nonparametric methods. These methods of testing for statistical significance are becoming increasingly important in certain types of business research and a few of the most useful methods have been described in Chapter 9 to supplement the coverage of the more traditional parametric methods in Chapter 8.

One of the most widely used applications of statistical inference is in the use of sampling to control the quality of manufactured product. Some of the techniques of statistical quality control are described as an introduction to a subject that has become a highly specialized use of the methods of statistical inference in making industrial decisions.

Regression and correlation analysis is given fairly complete coverage in the belief that even the most elementary course in statistics should explain the theory and methods of this very useful tool with its many applications to business problems.

Various types of study material are given at the end of each chapter to aid the student in applying the principles discussed in the text. This material has been grouped into two categories: study questions and problems. The study questions require the student to give some thought to the meaning of the various topics discussed in the chapter. The problems are to be used in developing skill in performing the calculations needed for the various methods of analysis. Various aids to calculation are provided in the tables in the appendixes.

We are indebted to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., Cambridge; to Dr. Frank Yates, F.R.S., Rothamsted; and to Messrs. Oliver & Boyd, Ltd., Edinburgh, for permission to reprint in Appendixes K and L the tables from their book *Statistical Tables for Biological, Agricultural, and Medical Research.* We are also indebted to Professor Stephen P. Shao, Old Dominion University, and the publisher for permission to reprint in Appendix E the table from his book *Mathematics for Management and Finance;* and to Professor Charles T. Clark, The University of Texas at Austin, Professor

Lawrence L. Schkade, The University of Texas at Arlington, and the publisher for permission to reprint in Appendixes F and J the tables from their book *Statistical Analysis for Administrative Decisions.*

J.R.S.
C.T.C.

Contents

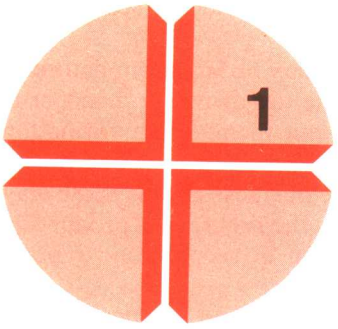
INTRODUCTION

Part 1 DATA ANALYSIS

Part 2 STATISTICAL INFERENCE

Part 4 ANALYSIS OF BUSINESS CHANGE

APPENDIXES



Statistics in Business Decision Making

Statistical techniques are used today in almost every area of human enterprise. Whether the concern is with predicting the weather, fighting disease with new drugs, preventing crime, evaluating a new food product, or forecasting population growth, there is an underlying core of uncertainty which must be dealt with in an objective and scientific manner to obtain the best results. The use of statistics provides some capability for dealing with this uncertainty.

The application of statistical techniques to business decisions are many and result from the fact that virtually all important business decisions are made under conditions of uncertainty. Business has become increasingly complex, and the responsibilities of the business executive have become correspondingly greater. Often the business executive must choose one definite course of action from all those open to him, although the consequences of each course of action cannot be fully known at the time the decision must be made.

To make such difficult decisions in the face of uncertainty demands two things. First, there must be an elaborate information system to gather and to supply all the facts needed by the executive. Second, there must be available a tested, and often sophisticated, set of tools for evaluating those facts. The nature and application of this set of tools is the subject of this book.

When most business units were small, the facts needed by the manager were not only fewer in number but were also more easily obtained. Since the market was close at hand and the customers were, for the most part, personal friends of the owner of the business, there was little need for an elaborate analysis of the market. The owner of the business could find out what his customers thought about his product and service simply by listening to their comments or asking them questions. However, in today's large business enterprises, goods are manufactured long before they are offered for sale. Producers and consumers are strangers to each other, and manufacturers must try to anticipate not only what consumers will buy but also the quantity they will buy. In order to get the information needed for intelligent planning of the output

of a factory, management must rely on a systematic method of securing information from consumers.

The personnel problems of the small business may be similar in nature to those of the large concern, but in a small enterprise the owner or manager is close to his employees and knows a great deal about them. The information that must now be collected in elaborate personnel records was at one time a matter of personal knowledge of the manager of the business. In the same way the owner of a small factory could keep informed on the progress of production in an informal manner without the elaborate system of records and reports that is necessary in a large organization.

The increasing reliance on information collected and summarized in reports has increased the time required in getting the information to the businessmen who are to use it. Not only has the preparation time for completing reports increased the cost of the information, but the collection time has reduced its value in making decisions and controlling the business. The pressure for timely information has revolutionized the techniques and the machinery for collecting and summarizing business data.

NATURE OF STATISTICAL DATA

Statistical data are facts expressed in quantitative form. They are the building blocks of statistical analysis. As business has grown more complex and the demand for information needed to make decisions has increased, a change has come about in the nature of information required. Greater emphasis is being placed on information expressed precisely as quantities rather than in more general qualitative terms.

In many cases the characteristics of the item for which information is desired can be expressed in measurable or countable quantities such as dollar sales, tons of coal, or bushels of wheat. Statistical data of this quantitative type can be analyzed in many ways to make the facts more meaningful and more useful to the business executive.

Some kinds of facts are not directly quantifiable; that is, they cannot be measured on any kind of a scale. This type of statistical data is produced by grouping individual observations into qualitative categories or classes. Good examples are hospital patients classified according to their ailments or machined parts classified as either acceptable or not acceptable. The kinds of statistical analysis possible with qualitative data are not the same as with quantitative data, but useful information can still be obtained.

Confusion frequently arises over the use of the word *statistics* because it has three meanings: (1) it is used to refer to statistical data; (2) it is used to refer to a body of methods and techniques for dealing with statistical data; and (3) the term "statistic," used in the singular, refers to a measure derived from a sample of data. The third meaning of the term will be discussed in Chapter 7, which deals with statistical sampling.

Anything that exhibits differences of magnitude or number is called a *variable* and serves as the raw material for statistical analysis. A company's retirement and pension program probably requires the age of each employee

on the payroll. The value of this variable is obtained by measuring the age of each employee. The sales department may be concerned with the cost of processing small orders and may ask for information about the size of the orders received in a selected period of time. This characteristic may be measured by computing the amount of each order in either physical quantities or in total cost of the goods bought. Either of these units is a value of a variable.

It is also possible to derive the values of a variable by computing the ratio between two amounts such as the number of deaths per thousand population. This is a ratio between the number of deaths in a given period and the thousands of persons in the population. Derived variables may be in many forms. A few examples are per capita income for each state, average yield of cotton per acre for various cotton-growing regions of the south, and fatalities on airlines per million passenger-miles flown.

ANALYSIS OF STATISTICAL DATA

After a mass of measurements has been made to secure the statistical data related to a given problem, the data must be processed in order to be in the most useful form for management. Generally the individual using the data does not want all the details that have been collected; he usually wants summaries and as much analysis of the figures as possible. The methods used to measure the characteristics of the data, to put the individual measurements into some kind of summary form, and to analyze the data to bring out the greatest amount of information are called *statistical methods*.

The theory underlying the methods of analysis used on statistical data is referred to as *statistical theory*. Any discussion of the methods of analyzing data will also include a discussion of the logic underlying the methods, but the rigorous development of this logic is mathematical in nature. Courses entitled mathematical statistics stress the mathematical theory underlying the methods of analysis.

The technical statistical methods used to make the needed measurements of the characteristics that are important to the problem and to bring them together in a summary form may be called *descriptive statistics*. This term is appropriate because the purpose of these methods is to make effective use of statistical data in communicating with the executive making a decision. The better the data describe the situation facing the executive the more useful they are. Methods that have been developed to collect, summarize, and in other ways manipulate the various types of data needed to describe situations are discussed in detail in Part 1.

It is not always necessary to compile all the data concerning a particular problem in order to secure valuable information for decision making. Under certain conditions the analysis of a sample taken from a large group will give the information needed at a small fraction of the cost required to compile complete data. The methods used in selecting the sample and in drawing inferences from sample data are powerful tools of research and are used with great effectiveness in making business decisions. Reaching conclusions from a sample is known as *statistical induction* and will be discussed in Part 2.

Parts 1 and 2 are concerned with methods of analyzing data involving only one variable at a time. That is to say, an observation consists of only one measurement or count on the variable under study.

Part 3 of this text discusses problems dealing with the relationship between two or more variables considered simultaneously. An observation consists of a measurement on several variables at one time. When it can be shown that variations in one or more independent variables are related to the variations in another variable called the dependent variable, this fact may be useful in predicting the action of the dependent variable based on prior knowledge of the independent variables to which it is related.

For example, if it is known that gasoline consumption in any state is related to the number of cars registered and the number of miles of highway in that state, this fact can be used to estimate gasoline consumption based on a knowledge of the other two factors long before the gasoline is purchased for use by the consumer. How this estimate is made and how accurate it may be is a matter of statistical analysis.

Part 4 deals with the analysis of business change. Because much of the data with which the businessman works are observations measured at various points in time, a special kind of statistical analysis is required to understand them.

For example, a personnel director keeps records of the number of employees on the payroll each month. The statistical data with which he deals are part of a time series, and the observations are not independent of each other. While the numbers may change from month to month, each month is related to what was observed the previous month. By careful statistical analysis it may be possible to measure the trend or average growth in the work force, the seasonal pattern if one exists, and the effect of swings in the business cycle, and perhaps forecast future changes in the size of the working group.

MODERN DECISION MAKING

Developments in the use of statistical data in business decision making have created some new and powerful tools that can be used by the businessman. The pressure of more complex business problems and the scale of operation have stimulated research in methods of analyzing data to meet these growing problems. Many fields of knowledge have contributed to the improvement in business decision making, but the advances in science on many fronts during the twentieth century have had an unusually significant influence on the thinking about the decision-making process in business.

At first glance it might appear that the business executive faces problems that bear no resemblance to those with which the scientist deals. However, the use of objective, unbiased methods that attempt to make use of all the facts in a given situation may justify the executive in believing that he is using the methods of science in making his decisions. The use of more objective methods and more accurate information may be expected to improve the quality of business decisions.

World War II gave tremendous impetus to efforts to improve decision-making techniques, and this trend has continued at an accelerated pace since the end of the war. The increasing use of statistical methods of analysis has provided a larger volume of precise information to use in the analysis of business problems, and in recent years a great deal of experimentation has been carried on in the hope of finding better methods of analysis in making decisions. During World War II the methods used by scientists were applied with outstanding success to certain military problems, such as determining the best program for strategic bombing, for mining enemy waters, and for searching out enemy ships. Business organizations now apply these same principles to such problems as determining the best policy to follow with respect to product mix, inventory fluctuations, and shipments of goods to various markets from different warehouses. The names commonly applied to this scientific or mathematical approach to the solution of business problems are *management science* or *operations research.*

Operations Research

Operations research is the application of scientific methods of analysis to executive-type problems involving the operations of man-machine systems in industrial organizations in order to provide those in control of the operations with optimum solutions to problems. Because business is so complex, operations research has usually been able to make effective use of a team whose members have been drawn from a number of special fields, including mathematics, statistics, physics, economics, and engineering. The methods used are those employed by scientists, and in general the approach is quantitative. The methods involve the prediction of various courses of action and thus provide a basis that management can use in choosing between alternative courses of action. The complex problems for which solutions are sought involve many important aspects of a firm's operations.

One of the most effective techniques used in operations research is *linear programming*, a body of techniques for solving problems dealing with many variables in which an objective that can be expressed quantitatively is attained, subject to certain restraining conditions. This mathematical method has succeeded in giving the best solution to problems that previously had no solution except through the judgment of the executive. It is reasonable to conclude that progress will continue to be made in the solution of more complex problems by techniques of this type, although this does not mean that it will be possible to optimize all decisions in the near future. Fairly simple methods of analysis of statistical data in making decisions will continue to be important, and as the use of quantitative methods by business executives becomes more widespread, better decisions will be made.

The Role of the Computer

The computer is involved in analyses such as operations research, for without the computer most of the calculations needed could not be made.

During the last two decades the businessman has found the computer to be a valuable aid to decision making, has revamped his organization and his paperwork routines to adapt to the computer, and is now digesting the gains and looking ahead to the changes a new generation of computers will bring. The advanced systems currently being developed are likely to have an impact on management planning and control equal to the revolution of paperwork processing by early computers.

The computer has been readily and enthusiastically accepted to perform mathematical computations with great speed and accuracy. Some business problems which are routinely solved via computer in just a few seconds would require years of manual computation time. A few of the jobs performed by computers are controlling inventories, preparing payrolls, handling checkouts in a supermarket, making airplane reservations, and handling credit card billing. In recent years decreasing costs and the use of remote terminals have enabled even small businesses to take advantage of the savings of automated data processing that were once available only to large firms.

OUTLINE OF STEPS IN STATISTICAL METHODS

The methods of analysis of quantitative data may be broken into the following six steps:

1. Definition of the problem
2. Assembly of information
3. Collection of original data
4. Classification
5. Presentation
6. Analysis.

A brief description of these steps is given in the following paragraphs.

Definition of the Problem

Before starting to assemble the information needed to solve a problem, the problem should be clearly defined so that quantitative methods can be used whenever possible in reaching a solution. If a problem is defined using quantitative analysis, there is a strong possibility that a better decision can be reached than would otherwise have been possible. Use of newer techniques, such as operations research, to make the decision requires a precise statement of the problem in quantitative terms.

Assembly of Available Information

The first step to take after defining the problem is to make a search for all readily available information that relates to the problem. Although sometimes the facts will be predominantly quantitative, on other occasions the most