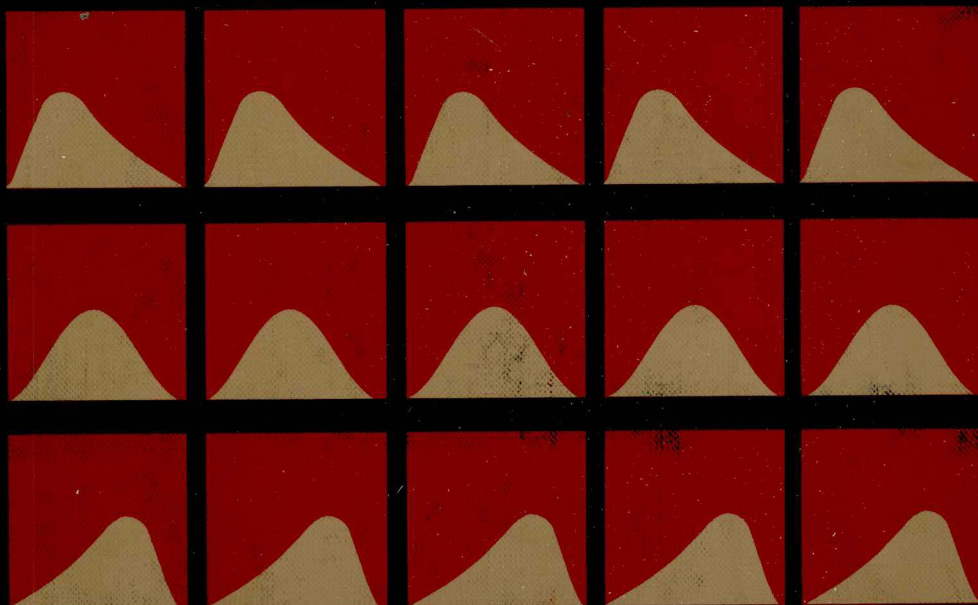


Statistical
Techniques

in Business
and Economics

Mason



Statistical
Techniques

In Business
and Economics

Fifth
Edition

Robert D. Mason
The University of Toledo



1982

RICHARD D. IRWIN, INC. Homewood, Illinois 60430

© RICHARD D. IRWIN, INC., 1967, 1970, 1974, 1978, and 1982

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

ISBN 0-256-02656-4

Library of Congress Catalog Card No. 81-82880

Printed in the United States of America

4 5 6 7 8 9 0 MP 9 8 7 6 5 4 3

Dow Jones-Irwin, a division of Richard D. Irwin, Inc., has developed Personal Learning Aids (PLAIDS) to accompany texts in this subject area. Copies can be purchased through your bookstore or by writing PLAIDS, 1818 Ridge Road, Homewood, IL 60430.

Preface

This text is designed primarily for a first course in business and economic statistics. However, students majoring in any of the social sciences will find the illustrative material quite familiar. The mathematics can be handled by students who do not have an extensive mathematical background.

The book is divided into 24 chapters, and 16 appendixes are included. There is ample material for a one-year course, but considerable latitude is possible in the selection of topics for shorter courses. A one-semester course, for example, could include Chapters 1–5, 9–15, and 20. The primary objective of this fifth edition remains the same as that of the previous editions. Both descriptive statistics and inferential statistics are covered, but the main emphasis is on inferential statistics. In addition, there are a number of special-interest topics such as quality control, time series analysis, and index numbers.

The text includes a number of new special features.

It is now oriented to the example—solution approach.

At various points in each section there are *self-reviews* designed to give students an opportunity to work problems similar to the examples. Answers and methods of solution are given in the margin adjacent to the self-review problems. This work serves to reinforce the student's understanding of the preceding material.

A *chapter self-review examination* at the end of each chapter covers all the material in the chapter. This test allows the student to evaluate his or her overall comprehension of the subject matter.

A number of applications presented use the Statistical Package for the Social Sciences, Minitab, and Basic to illustrate the computer's potential for problem solving.

Both a *chapter summary* and a *chapter outline* are presented at the end of each chapter. Students can use them to pull together the ideas of the chapter.

New material in this fifth edition also includes exponential smoothing, the moving-average method of smoothing time series, and the normal approximation to the binomial. The material on probability and probability distributions has been revised and expanded to three chapters. The presentation of tests of hypotheses has also been revised and is now contained in three chapters.

A new, 200-page Study Guide now accompanies the text. The chapter

organization of this comprehensive study guide is the same as the text's. For each chapter there are:

Chapter objectives.

An extensive summary.

A glossary of terms.

Problems with detailed solutions.

Student practice exercises, with answers.

A two- or four-page tear-out chapter assignment.

A Teacher's Manual contains the full solutions to all the exercises in the textbook and all the chapter assignments in the Study Guide.

I wish to express special thanks to colleagues William G. Marchal and Douglas A. Lind for their many helpful comments in organizing this fifth edition. Special acknowledgment is given to Jeff Hooper, Mesa Community College; Charles N. Keppler, Miami Dade Community College; Charles E. Leitle, Missouri Southern State College; William G. Marchal, University of Toledo; Douglas W. Morrill, Centenary College of Louisiana; John C. Shannon, Suffolk University, Boston, Massachusetts; and Lee C. Wilson, Mason Community College; who made many valuable comments and criticisms in their reviews of the manuscript. I am 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 and Boyd, Ltd., for permission to reprint Tables III and IV of their book *Statistical Tables for Biological, Agriculture and Medical Research*.

Robert D. Mason

A note to the student

This textbook is constructed to aid you in your study by presenting the subject matter in small, easy-to-take steps with frequent checks to determine whether you fully understood the preceding material.

As you progress through each chapter you will be asked to solve a problem and interpret your findings. The answer to each *self-review* is always given on the left side adjacent to the problem. Cover the answer before starting work. Following is an example.

SELF-REVIEW

- a.* The range is 54.3, found by
b. Highest value – Lowest value
= 297.9 – 243.6
= 54.3.

(Cover the answers in the left hand column.)

The consumer price indexes for May in selected cities are: Cleveland, 285.1; Buffalo, 297.9; Denver, 256.7; Atlanta, 284.7; Chicago, 243.6; Dallas, 277.2.

- a.* What is the range?
b. Now verify your answer against the one given in the margin.
-

There are *exercises* after each section, if you need additional practice. The answers to selected exercises and the methods of solution are given in Appendix A.

Another distinctive feature of this text is the *Chapter Self-Review Examination* with answers and method of solution given in Appendix A. By completing this test you can better evaluate your comprehension of all the material in the chapter.

Contents

1. What is statistics?	1
Introduction, 1	
Descriptive statistics, 2	
Inductive statistics, 3	
2. Collection and presentation of data	5
Introduction, 5	
Sources of published data, 6	
Standard business and economic sources, 6	
<i>Publications of the federal government, 6</i>	
<i>Publications of private organizations, 7</i>	
<i>Publications of international organizations, 8</i>	
Consulting reference aids, 8	
The survey as a source of data, 9	
Criteria for questionnaire construction, 10	
Levels of measurement, 11	
Nominal level of measurement, 11	
Ordinal level of measurement, 12	
Interval and ratio levels of measurement, 12	
Graphical presentation of data, 13	
The arithmetic scale, 13	
<i>Simple line charts and simple bar charts, 15</i>	
<i>Multiple-line charts and multiple-bar charts, 16</i>	
<i>Component bar charts, 17</i>	
<i>Two-directional bar charts, 18</i>	
<i>Pie charts, 20</i>	
Semilogarithmic (ratio) scaling, 21	
<i>Design of semilogarithmic (ratio) paper, 21</i>	
<i>Application of semilogarithmic paper, 22</i>	
Chapter summary, 25	
Chapter outline, 25	
Chapter exercises, 27	
3. Describing data—frequency distributions	33
Introduction, 33	
The frequency distribution, 34	
True class limits, 36	
Midpoints, 37	

- Class intervals, 37
- Hints on constructing a frequency distribution, 39
- Graphical representation of a frequency distribution, 42
 - Histogram, 42
 - Frequency polygon, 44
 - Percent frequency polygon, 46
 - Cumulative frequency polygon, 49
- Chapter summary, 53
- Chapter outline, 53
- Chapter exercises, 56

4. Describing data—measures of central tendency 59

- Introduction, 59
- Mean, median, and mode of ungrouped data, 60
 - The arithmetic mean, 60
 - The weighted arithmetic mean, 61
 - The median, 62
 - The mode, 64
 - Ungrouped data, 64*
- Mean, median, and mode of grouped data, 67
 - The arithmetic mean, 67
 - The median, 72
 - The mode, 76
- Relationship of the mean, median, and mode, 79
- Geometric mean, 81
- Chapter summary, 85
- Chapter outline, 85
- Chapter exercises, 88
- Chapter self-review examination, 91

5. Measures of dispersion and skewness 95

- Introduction, 95
- Why study dispersion? 96
- Measures of dispersion—ungrouped data, 96
 - The range, 96
 - The average deviation, 97
 - The variance, 99
 - The population standard deviation, 102
 - The sample variance as an estimator of the population variance, 103
 - The sample standard deviation as an estimator of the population standard deviation, 105
- Measures of dispersion for data grouped into a frequency distribution, 106
 - The range, 106
 - The average deviation, 107
 - The variance and the standard deviation, 107
- Interpretation and uses of the standard deviation, 112

- Some other measures of dispersion, 117
 - The interquartile range, 117
 - The quartile deviation, 119
 - The percentile range, 119
- Relative dispersion, 121
- Measures of skewness, 123
- A computer application, 124
- Chapter summary, 125
- Chapter outline, 126
- Chapter exercises, 129
- Chapter self-review examination, 131

6. Index numbers 133

- Introduction, 133
- Some basic concepts, 134
 - Price indexes, 137*
 - Quantity indexes, 137*
 - Value indexes, 137*
 - Special-purpose indexes, 137*
- Construction of index numbers, 137
 - Unweighted index numbers, 138
 - Weighted indexes, 140
 - Price indexes, 140*
 - Quantity indexes, 142*
 - Value indexes, 145*
 - Special-purpose indexes, 146*
- Consumer price index, 148
- Special uses of the consumer price index, 149
 - Real income, 149
 - Deflated sales, 150
 - Purchasing power of the dollar, 151
 - Cost-of-living increases, 152
- Shifting the base, 152
- Chapter summary, 154
- Chapter outline, 155
- Chapter exercises, 156
- Chapter self-review examination, 160

7. Time series analysis 163

- Introduction, 163
- Components of a time series, 165
 - Secular trend, 165
 - Seasonal variation, 165
 - Cyclical variation, 165
 - Irregular variation, 166
- Linear trend, 167

- Freehand method, 168
- Least squares method, 170
- Nonlinear trends, 177
 - Logarithmic straight-line trend, 177
 - Other trend equations, 182
- The moving average method, 183
- Exponential smoothing, 187
- Chapter summary, 192
- Chapter outline, 193
- Chapter exercises, 195
- Chapter self-review examination, 198
- BASIC program for exponential smoothing, 199

8. Seasonal analysis 201

- Introduction, 201
- Deseasonalizing data, 204
- Methods of determining the seasonal indexes, 205
 - A method using averages, 205
 - Ratio-to-moving-average method, 208
 - A technical note, 208*
- Chapter summary, 215
- Chapter outline, 216
- Chapter exercises, 216
- Chapter self-review examination, 219

9. A survey of probability concepts 221

- Introduction, 221
- What is probability? 222
- Approaches to probability, 223
- Uses of probability, 227
- Some basic rules of probability, 227
 - Rules of addition, 227
 - Special rule of addition, 227*
 - Venn diagrams, 229
 - Complement rule, 229*
 - General rule of addition, 231*
- Rules of multiplication, 234
 - Special rule of multiplication, 234*
 - General rule of multiplication, 235*
- Tree diagrams, 238
- Principles of counting, 241
 - The multiplication principle, 242
 - Permutations, 243
 - Combinations, 247
- Chapter summary, 249
- Chapter outline, 251
- Chapter exercises, 252
- Chapter self-review examination, 253

10. Discrete probability distributions 255

Introduction, 255
What is a probability distribution? 256
Discrete probability distributions, 257
Continuous probability distributions, 258
Binomial probability distributions, 258
 How is a binomial distribution constructed? 260
 Binomial probability tables, 262
The importance of probability distributions in decision making, 265
Poisson probability distribution, 268
Chapter summary, 272
Chapter outline, 273
Chapter exercises, 274
Chapter self-review examination, 276

11. The normal probability distribution 279

Introduction, 279
Some features of the normal distribution, 280
The standard normal distribution, 281
Applications of the standard normal distribution, 283
Other uses of the standard normal distribution, 291
The normal approximation to the binomial, 293
Chapter summary, 296
Chapter outline, 297
Chapter exercises, 298
Chapter self-review examination, 300

12. An introduction to sampling 301

Introduction, 301
Reasons for sampling, 303
Methods of selecting a sample, 304
 Probability sampling, 304
 Simple random sampling, 304
 Systematic sampling, 305
 Stratified random sampling, 306
 Nonprobability sampling, 308
Sampling error, 310
Sampling distribution of the mean, 311
The central limit theorem, 315
 Some applications of the central limit theorem, 316
Size of the sample, 319
 Degree of confidence, 320
 Maximum error allowed, 320
 Estimating the variation in the population, 322
Confidence interval for a population proportion, 323
Size of the sample for proportions, 325
Chapter summary, 327

Chapter outline, 328
Chapter exercises, 329
Chapter self-review examination, 330

13. Tests of hypotheses: Large samples 333

Introduction, 333
The procedure in hypothesis testing, 334
 The null and alternate hypotheses, 334
 The level of significance, 335
 The test statistic, 335
 The decision rule, 336
 Making a decision, 338
Testing a hypothesis for the population mean, 338
Hypothesis testing: Two means, 344
Type II errors, operating characteristic curves, and power curves, 349
 Type II errors, 349
 Operating characteristic curves, 352
 Power curves, 353
Chapter summary, 353
Chapter outline, 354
Chapter exercises, 356
Chapter self-review examination, 358

14. Tests of hypotheses about proportions 359

Introduction, 359
A test about a population proportion, 360
A test for the difference between two population proportions, 364
Chapter summary, 370
Chapter outline, 371
Chapter exercises, 372
Chapter self-review examination, 373

15. Student's t test: Small samples 375

Introduction, 375
The t distribution, 376
A test for the population mean, 377
A computer output, 382
A test for the difference between two population means, 383
Chapter summary, 388
Chapter outline, 388
Chapter exercises, 390
Chapter self-review examination, 391

16. Analysis of variance: One-way and two-way classifications	393
Introduction, 393	
ANOVA: The general idea, 394	
One-way analysis of variance, 395	
Two-way analysis of variance, 404	
A production problem, 407	
Chapter summary, 409	
Chapter outline, 409	
Chapter exercises, 410	
Chapter self-review examination, 415	
 17. Nonparametric tests of hypotheses: One-sample and two-sample tests	 417
Introduction, 417	
Chi-square: One-sample test of significance, 418	
Imitations of chi-square, 422	
Mann-Whitney U test, 427	
Small samples, 427	
Large samples, 431	
Kolmogorov-Smirnov two-sample test, 433	
The Wilcoxon matched-pair signed rank test, 439	
Chapter summary, 443	
Chapter outline, 444	
Chapter exercises, 447	
Chapter self-review examination, 450	
 18. Nonparametric tests of hypotheses: More than two samples	 451
Introduction, 451	
The Cochran Q test, 452	
The Friedman two-way analysis of variance, 457	
The chi-square test for k samples, 460	
Chapter summary, 464	
Chapter outline, 466	
Chapter exercises, 468	
Chapter self-review examination, 469	
 19. Statistical quality control	 471
Introduction, 471	
The control chart viewpoint, 473	
Causes of variation, 474	
Purpose and types of quality control charts, 476	

- Charts for variables, 477
 - Bar \bar{X} chart (mean chart), 478
 - Range chart, 481
- Charts for attributes, 485
 - Percent defective chart, 485
 - \bar{c} chart, 488
- Chapter summary, 491
- Chapter outline, 492
- Chapter exercises, 493
- Chapter self-review examination, 495

20. Simple regression and correlation analysis 497

- Introduction, 497
- Regression analysis, 498
 - Scatter diagram, 498
 - The regression equation, 499
 - Least squares method of determining the regression equation, 499
 - Determining the points on the straight line, 503
 - Errors in predicting, 504
 - Confidence limits for predictions, 508
 - Review of regression analysis, 510
- Correlation analysis, 512
 - Coefficients of determination, nondetermination, and correlation, 513
 - Testing the significance of the correlation coefficient, 516
 - Small samples, 516*
 - Large samples, 517*
- An application using a computer package, 518
- Chapter summary, 519
- Chapter outline, 519
- Chapter exercises, 521
- Chapter self-review examination, 525

21. Multiple regression and correlation analysis 527

- Introduction, 527
- Multiple regression analysis, 528
 - Solving for the multiple regression equation, 529
 - A shortcut method, 529*
 - An alternative shortcut procedure, 531*
 - Predicting Y , 532
 - Standard error of estimate, 533
- Multiple correlation analysis, 535
 - Coefficient of multiple determination, 535
 - Coefficient of multiple nondetermination, 536
 - Coefficient of multiple correlation, 536
- An application using a computer package, 536

A salary-prediction problem, 538
Chapter summary, 543
Chapter outline, 543
Chapter exercises, 544
Chapter self-review examination, 548

22. Nonparametric measures of association 551

Introduction, 551
The Spearman rank-order correlation coefficient, 552
 Testing the significance of R_{ρ} , 556
The Kendall coefficient of concordance: W , 558
 Testing the significance of W , 560
 Small samples, 560
 Large samples, 561
Chapter summary, 563
Chapter outline, 563
Chapter exercises, 565
Chapter self-review examination, 567

23. An introduction to decision making under uncertainty 569

Introduction, 569
An illustrative case of decision making under uncertainty, 572
 Payoff table, 572
 Expected payoff, 572
 Opportunity loss, 573
 Expected opportunity loss, 575
Maximin, maximax, and minimax regret strategies, 576
Value of perfect information, 576
Sensitivity analysis, 577
Decision trees, 579
Chapter summary, 581
Chapter outline, 581
Chapter exercises, 582
Chapter self-review examination, 585

24. Bayes' theorem and decision making 587

Introduction, 587
The rule of Bayes, 589
A case involving the best level of production for Giggle dolls, 591
Chapter summary, 597
Chapter outline, 597
Chapter exercises, 598
Chapter self-review examination, 599

Appendixes 601

- Appendix A: Answers to selected exercises and the chapter self-review examinations, 602
- Appendix B: Logarithms, 642
- Appendix C: Squares and square roots, 644
- Appendix D: Table of random numbers, 654
- Appendix E: Critical values of t , 655
- Appendix F: Wilcoxon T values, 656
- Appendix G: Table of areas and ordinates of the normal curve, 657
- Appendix H: Critical values of chi-square, 662
- Appendix I: Binomial distribution, 663
- Appendix J: Poisson distribution for selected values of μ , 669
- Appendix K: Factors for control charts, 670
- Appendix L: Critical values of U in the Mann-Whitney test, 671
- Appendix M: Critical values of ρ , the Spearman rank correlation coefficient, 672
- Appendix N: Critical values of s in the Kendall coefficient of concordance: W , 673
- Appendix O: Critical values of the F distribution at a 5 percent level and a 1 percent level of significance, 674
- Appendix P: Critical values of D in the Kolmogorov-Smirnov two-sample test (large samples: two-tailed test), 676

Index 677

INTRODUCTION

Many television viewers visualize a **statistician** as one who records each yard gained by rushing in the Dallas Cowboys-Denver Broncos Monday night football game, or who records each foul shot attempted and points made in the Boston Celtics-Detroit Piston game. We are constantly bombarded during and after the game with **statistics**: Larry Bird of the Celtics scored a career high of 28 points in the first half, Tony Dorsett gained 103 yards before retiring in the third quarter, Chris Evert Lloyd beat Virginia Ruzici of Romania 7-5 and 6-1, and Gorman Thomas of the Brewers scored two runs, and had one hit in three times at bat.

Likewise, reference to *The Wall Street Journal*, *Business Week*, annual reports of industry, and other sources of business data revealed that Arvida Realty Sales, Inc.,¹ is handling the sale of a four-bedroom home in the harbor section, priced at \$547,500; the average hourly earnings in contract construction are \$10.35,² Connecticut produced 4,950,000 pounds of tobacco in 1981;³ and IBM has 341,279 employees, 737,230 stockholders, and an annual gross income from sales, rentals, and services of \$26,213,000,000.⁴ Fifty-one percent of the millionaires in the United States are women, and 95 percent of the money in the United States is managed by women. A final statistic from Norton Simon: the total compensation of the chairman of the board, David J. Mahoney, was \$2,037,000 in 1981.

Statistics has a more precise meaning, however. It is defined as the *science of collecting, organizing, presenting, analyzing, and interpreting numerical data for the purpose of making better decisions in the face of uncertainty*.

Just as lawyers have "rules of evidence" and accountants have "commonly accepted practices," statisticians follow some standard guidelines when dealing with data. The basic techniques employed by statisticians are discussed in the following chapters.

¹ *The Sarasota (Florida) Herald-Tribune*, April 12, 1981, p. 11D.

² *Survey of Current Business*, February 1981, p. S-13.

³ Department of Agriculture.

⁴ 1980 Annual Report, International Business Machines Corporation, p. 1.