

STATISTICS FOR BUSINESS AND ECONOMICS

THIRD EDITION



Prentice-Hall International Editions

PAUL NEWBOLD

CONTENTS

PREFACE xiii

1

WHAT IS STATISTICS? 1

- 1.1 Making Sense of Numerical Information 1
- 1.2 Dealing with Uncertainty 2
- 1.3 Sampling 3
- 1.4 Analyzing Relationships 4
- 1.5 Forecasting 5
- 1.6 Decision-Making in an Uncertain Environment 5

2

SUMMARIZING NUMERICAL INFORMATION 7

- 2.1 Populations and Samples 7
- 2.2 Numerical Summary: Measures of Central Tendency 9
- 2.3 Numerical Summary: Measures of Dispersion 16
- 2.4 Grouped Data and Histograms 31
- 2.5 Numerical Summary of Grouped Data 39
- 2.6 Some Other Graphical Methods 56
- 2.7 Lying with Statistics 68

3

PROBABILITY 82

- 3.1 Introduction 82
- 3.2 Random Experiment, Outcomes, Events 83
- 3.3 What Is Probability? 92
- 3.4 Probability and Its Postulates 94

3.5	Permutations and Combinations	100
3.6	Probability Rules	105
3.7	Bivariate Probabilities	117
3.8	Bayes' Theorem	123

53s.

4

DISCRETE RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS 138

4.1	Random Variables	138
4.2	Probability Distributions for Discrete Random Variables	140
4.3	Expectations for Discrete Random Variables	145
4.4	Jointly Distributed Discrete Random Variables	155
4.5	The Binomial Distribution	169
4.6	The Hypergeometric Distribution	176
4.7	The Poisson Distribution	178

48s.

5

CONTINUOUS RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS 191

5.1	Continuous Random Variables	191
5.2	Probability Distributions for Continuous Random Variables	192
5.3	Expectations for Continuous Random Variables	199
5.4	Jointly Distributed Continuous Random Variables	203
5.5	The Normal Distribution	208
5.6	The Central Limit Theorem	222
5.7	The Normal Distribution as an Approximation to the Binomial and Poisson Distributions	226
5.8	The Exponential Distribution	232

6

SAMPLING AND SAMPLING DISTRIBUTIONS 239

6.1	Sampling from a Population	239
6.2	Sampling Distribution of the Sample Mean	244
6.3	Sampling Distribution of a Sample Proportion	251
6.4	Sampling Distribution of the Sample Variance	262
Appendix A6.1		274

7

POINT ESTIMATION 275

- 7.1 Introduction 275
- 7.2 Unbiased Estimators and Their Efficiency 278
- 7.3 Choice of Point Estimator 284

8

INTERVAL ESTIMATION 289

- 8.1 Confidence Intervals 289
- 8.2 Confidence Intervals for the Mean of a Normal Distribution: Population Variance Known 291
- 8.3 The Student's t Distribution 301
- 8.4 Confidence Intervals for the Mean of a Normal Population: Population Variance Unknown 304
- 8.5 Confidence Intervals for the Population Proportion (Large Samples) 311
- 8.6 Confidence Intervals for the Variance of a Normal Population 315
- 8.7 Confidence Intervals for the Difference Between the Means of Two Normal Populations 321
- 8.8 Confidence Intervals for the Difference Between Two Population Proportions (Large Samples) 330
- 8.9 Estimating the Sample Size 336

9

HYPOTHESIS TESTING 346

- 9.1 Concepts of Hypothesis Testing 346
- 9.2 Tests of the Mean of a Normal Population Distribution: Population Variance Known 353
- 9.3 Tests of the Mean of a Normal Distribution: Population Variance Unknown 364
- 9.4 Tests of the Variance of a Normal Distribution 369
- 9.5 Tests of the Population Proportion (Large Samples) 373
- 9.6 Tests for the Difference Between Two Means 377
- 9.7 Tests for the Difference Between Two Population Proportions (Large Samples) 385
- 9.8 Testing the Equality of the Variances of Two Normal Populations 392
- 9.9 Measuring the Power of a Test 396
- 9.10 Some Comments on Hypothesis Testing 402

10

% pensum

SOME NONPARAMETRIC TESTS 413

- 10.1 Introduction 413
- 10.2 The Sign Test 414
- 10.3 The Wilcoxon Test 419
- 10.4 The Mann-Whitney Test 423
- 10.5 Discussion 428

11

GOODNESS-OF-FIT TESTS AND CONTINGENCY TABLES 434

- 11.1 Goodness-of-Fit Tests 434
- 11.2 Goodness-of-Fit Tests: Population Parameters Unknown 440
- 11.3 Contingency Tables 445

12

LINEAR CORRELATION AND REGRESSION 457

- 12.1 Correlation 457
- 12.2 Rank Correlation 466
- 12.3 The Linear Regression Model 472
- 12.4 Least Squares Estimation 477
- 12.5 Standard Assumptions for the Linear Regression Model 481
- 12.6 The Gauss-Markov Theorem 482
- 12.7 The Explanatory Power of a Linear Regression Equation 483
- 12.8 Confidence Intervals and Hypothesis Tests 491
- 12.9 Prediction 496
- Appendix A12.1 510

13

MULTIPLE REGRESSION 512

- 13.1 The Multiple Regression Model 512
- 13.2 Least Squares Estimation 520
- 13.3 Standard Assumptions for the Multiple Regression Model 523
- 13.4 The Gauss-Markov Theorem 524

13.5	The Explanatory Power of a Multiple Regression Equation	525
13.6	Confidence Intervals and Hypothesis Tests for Individual Regression Parameters	533
13.7	Tests on Sets of Regression Parameters	542
13.8	Prediction	548
13.9	Computer Packages for Regression Calculations	554
Appendix A13.1		568

14 *pensum*

ADDITIONAL TOPICS IN REGRESSION ANALYSIS 571

14.1	Model-building Methodology	571
14.2	Dummy Variables	574
14.3	Lagged Dependent Variables	578
14.4	Nonlinear Models	590
14.5	Specification Bias	594
14.6	Multicollinearity	596
14.7	Heteroscedasticity	605
14.8	Autocorrelated Errors	611
14.9	Summary	625

15

ANALYSIS OF VARIANCE 636

15.1	Comparison of Several Population Means	636
15.2	One-Way Analysis of Variance	639
15.3	The Kruskal-Wallis Test	648
15.4	Two-Way Analysis of Variance: One Observation per Cell, Randomized Blocks	655
15.5	Two-Way Analysis of Variance: More than One Observation per Cell	667

% pensum

16 *% pensum*

STATISTICAL QUALITY CONTROL 686

16.1	The Importance of Quality Control	686
16.2	Control Charts for Means and Standard Deviations	688
16.3	Process Capability	700
16.4	Control Charts for Proportions	703
16.5	Summary	707
Appendix A16.1		712

17

% pensum

TIME SERIES ANALYSIS AND FORECASTING 714

- 17.1 Time Series Data: Problems and Opportunities 714
- 17.2 Index Numbers 715
- 17.3 A Nonparametric Test for Randomness 727
- 17.4 Components of a Time Series 731
- 17.5 Moving Averages 736
- 17.6 Extraction of the Seasonal Component through Moving Averages 739
- 17.7 Simple Exponential Smoothing 750
- 17.8 The Holt-Winters Exponential Smoothing Forecasting Model 754
- 17.9 Autoregressive Models 767
- 17.10 Autoregressive Integrated Moving Average Models 773

18

SURVEY SAMPLING METHODS 780

- 18.1 Introduction 780
- 18.2 Sampling and Nonsampling Errors 786
- 18.3 Simple Random Sampling 789
- 18.4 Stratified Sampling 797
- 18.5 Determining the Sample Size 810
- 18.6 Other Sampling Methods 816

19

% pensum

STATISTICAL DECISION THEORY 831

- 19.1 Decision-Making Under Uncertainty 831
- 19.2 Solutions Not Involving Specification of Probabilities 835
- 19.3 Expected Monetary Value 841
- 19.4 Use of Sample Information: Bayesian Analysis 849
- 19.5 The Value of Sample Information 855
- 19.6 Allowing for Risk: Utility Analysis 870

APPENDIX: TABLES 882

- 1 Probability Function of the Binomial Distribution 882
- 2 Values of $e^{-\lambda}$ 887
- 3 Cumulative Distribution Function of the Standard Normal Distribution 888
- 4 Some Uniformly Distributed Random Numbers 890

5	Cutoff Points of the Chi-Square Distribution Function	891
6	Cutoff Points for the Student's t Distribution	893
7	Cutoff Points for the F Distribution	895
8	Cutoff Points for the Distribution of the Wilcoxon Test Statistic	898
9	Cutoff Points for the Distribution of Spearman's Rank Correlation Coefficient	899
10	Cutoff Points for the Distribution of the Durbin-Watson Test Statistic	900
11	Cumulative Distribution Function of the Runs Test Statistic	902

ANSWERS TO SELECTED EVEN-NUMBERED EXERCISES 903

INDEX 921

881 s.

pensum

746 s.

PREFACE

42-93 21.30

An understanding of statistical ideas, and the potential for their application, is so important to the manager that statistics courses are offered in virtually all college business programs and are required in the great majority. In the interests of honesty, it cannot be claimed that required statistics classes are the most popular in the curriculum. However, they are certainly among the most useful. Statistical methods are widely used in many areas of business, where numerical information needs to be collected and sensibly interpreted as an aid to decision-making in an uncertain environment.

In preparing the third edition of this text, I have kept in mind that not all statistical concepts are easily grasped. The book introduces a wide range of procedures appropriate for the analysis of various types of business data. A very poor strategy for attempting to master the subject would be to try to memorize the details of these procedures "cookbook" style. The student's goal should be an understanding of the rationale underlying the various statistical techniques. The author's goal has been to try to facilitate that understanding. The vast majority of readers of this book will become consumers rather than producers of statistical information. It is intended that this text will help them become well-informed consumers, with an understanding of the basis for, and the background to critically judge, the statistical claims that they will later meet. As an aid to understanding the methodology, I have included a very large number of numerical examples. These are intended not merely as specific applications of the general computations, but also as realistic illustrations of the situations in which particular methods are useful. Because understanding of the material is, however, best achieved through hands-on experience, the text contains a great many exercises. The student's understanding of the material will be sharpened by working through several exercises on each topic.

In writing this text, I have been conscious of the importance of persuading the student of the practical relevance of the subject, hoping thereby to induce the motivation to put in the time and effort needed to adequately understand statistical methods. My strategy has been to include many real and realistic examples and exercises taken from a range of business fields, including accounting, economics, finance, marketing, industrial organization, and organizational behavior. It is hoped that the accumulation of this illustrative material will convince the reader of the importance of statistical methods in the modern business environment, and so heighten his or her interest in the subject.

This text is suitable for a one- or two-semester course aimed at business or economics majors. I have included more than enough material for a two-semester course, and many instructors will not want to cover every chapter in detail, particularly if a large amount of course time is devoted to project work. All, or parts, of Chapters 10, 11, and 14–19 can be omitted without loss of continuity. One of many possibilities for a one-semester course is to cover Chapters 1–9, 12, and perhaps 13. In such a course, instructors may prefer to omit some of the following sections: 2.6, 4.6, 4.7, 5.8, 7.3, 9.9, and 12.2.

This book developed from a two-semester course at the University of Illinois, one covering most of the material in Chapters 1–9, 11–14, and 17. I am very grateful to my colleagues, teaching assistants, and, most of all, students for helpful discussions, suggestions, and criticisms. Many of these prompted changes from the previous edition, which I hope have improved the clarity of the exposition. The major addition in this edition is a new chapter, Chapter 16 on statistical quality control. The importance of this topic in production management has become increasingly widely recognized in the last few years, and this new Chapter 16 provides an introduction to the principles involved, while requiring only elementary statistical concepts. I have tried to maintain and enhance the two crucial features of previous editions. These were the care and clarity in explaining, using no mathematics beyond basic algebra, how and why the various procedures work, and the provision of large numbers of examples and exercises involving real business and economic data.

I continue to acknowledge the many people without whose help previous editions of this book would have been less successful. In addition, in preparing the third edition, thanks also go to my Prentice-Hall editor Valerie Ashton and Kate Scheinman of The Total Book. Finally, I am grateful for the expert typing of Naoko Miki.

To the small minority of students who actually read prefaces, I wish you success in your studies of business statistics. In particular, I hope that I will be able to convince you of the value and relevance of statistics in many branches of business.

Paul Newbold