

COMPLETE BUSINESS STATISTICS 7e

Complete Business Statistics provides a solid foundation for understanding and using statistics in business. The examples and exercises presented in this book are drawn from a wide variety of current sources; applications are offered in finance, economics, government, technology, operations, and marketing - many with an international context. It also has the added flexibility to use both Excel and Minitab solutions. This new edition retains its global emphasis, maintaining its position of being at the vanguard of international issues in business.

New Features

- ❖ Special emphasis on statistical concepts with many more examples in the Asian and Indian context
- ❖ New and updated problems, examples, and cases
- ❖ Exercise applications and screens updated to Microsoft Excel®2007
- ❖ New Minitab applications and end of chapter tutorials provide 'how to do' information
- ❖ MegaStat® for Excel 2007 has been added to the student CD, offering an outstanding add-in option for Excel

www.mhhe.com/sie-aczel7e

Write to us at
info.india@mheducation.com

**Mc
Graw
Hill**
Education

www.mheducation.co.in

ISBN-13: 978-0-07-107790-3
ISBN-10: 0-07-107790-1



**Mc
Graw
Hill**
Education

COMPLETE BUSINESS STATISTICS

7e

Aczel
Sounderpandian
Saravanan
Joshi

**Mc
Graw
Hill**
Education

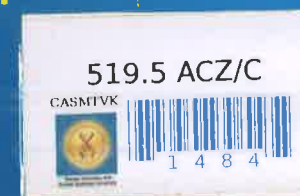


With Free CD



COMPLETE BUSINESS STATISTICS

Amir D Aczel
Jayavel Sounderpandian
Palanisamy Saravanan
Rohit Joshi



7e

**Mc
Graw
Hill**
Education

For sale in India, Pakistan, Nepal, Bangladesh, Sri Lanka and Bhutan only.

With Free CD





Complete Business Statistics

SEVENTH EDITION

Amir D. Aczel

Boston University

University of New Hampshire

Jayavel Sounderpandian

University of Wisconsin-Parkside

Palanisamy Saravanan

*Rajiv Gandhi Indian Institute of Management
(IIM Shillong)*

Rohit Joshi

*Rajiv Gandhi Indian Institute of Management
(IIM Shillong)*



McGraw Hill Education (India) Private Limited

CHENNAI

McGraw Hill Education Offices

Chennai New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto



McGraw Hill Education (India) Private Limited

Complete Business Statistics, 7/e

Indian Adaptation done by arrangement with McGraw-Hill Global Education Holdings, LLC, New York.

Sales Territories: India, Pakistan, Nepal, Bangladesh, Sri Lanka and Bhutan

Copyright © 2012, by McGraw-Hill Companies, Inc.

Tenth reprint 2017
RCQQCRAMDXXZQ

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise or stored in a database or retrieval system without the prior written permission of The McGraw-Hill Companies, Inc., including, but not limited, in any network or other electronic storage or transmission, or broadcast for distance learning.

This edition can be exported from India only by the publishers,
McGraw Hill Education (India) Private Limited.

ISBN (13 digits): 978-0-07-107790-3

ISBN (10 digits): 0-07-107790-1

Managing Director: *Kaushik Bellani*
Head—Higher Education (Publishing and Marketing): *Vibha Mahajan*
Senior Production Executive: *Atul Gupta*
General Manager—Production: *Rajender P Ghansela*
Production Manager: *Reji Kumar*

Information contained in this work has been obtained by McGraw Hill Education, from sources believed to be reliable. However, neither McGraw Hill Education nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw Hill Education nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw Hill Education and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Published by McGraw Hill Education (India) Private Limited,
444/1, Sri Ekambara Naicker Industrial Estate, Alapakkam, Porur, Chennai 600116, Tamil Nadu, India.

Typeset at The Composers, 260, C.A. Apt., Paschim Vihar, New Delhi 110063, and printed at
Sai Printo Pack Pvt. Ltd., A-102/4, Okhla Industrial Area, Phase-II, New Delhi 110020.

Visit us at: www.mheducation.co.in

*To the memories of my father,
Captain E. L. Aczel, and my mother, Miriam Aczel*

Amir D. Aczel

To our families

Palanisamy Saravanan and Rohit Joshi

PREFACE

Regrettably, Professor Jayavel Sounderbandian passed away before the revision of the text commenced. He had been a consistent champion of the book, first as a loyal user and later as a productive co-author. His many contributions and contagious enthusiasm will be sorely missed. In the seventh edition of *Complete Business Statistics*, we focus on many improvements in the text, driven largely by recommendations from dedicated users and others who teach business statistics. In their reviews, these professors suggested ways to improve the book by maintaining the Excel feature while incorporating MINITAB, as well as by adding new content and pedagogy, and by updating the source material. Additionally, there is increased emphasis on good applications of statistics, and a wealth of excellent real-world problems has been incorporated in this edition. The book continues to attempt to instill a deep understanding of statistical methods and concepts with its readers.

The seventh edition, like its predecessors, retains its global emphasis, maintaining its position of being at the vanguard of international issues in business. The economies of countries around the world are becoming increasingly intertwined. Events in Asia and the Middle East have direct impact on Wall Street, and the Russian economy's move toward capitalism has immediate effects on Europe as well as on the United States. The publishing industry, in which large international conglomerates have acquired entire companies; the financial industry, in which stocks are now traded around the clock at markets all over the world; and the retail industry, which now offers consumer products that have been manufactured at a multitude of different locations throughout the world—all testify to the ubiquitous globalization of the world economy. A large proportion of the problems and examples in this new edition are concerned with international issues. We hope that instructors welcome this approach as it increasingly reflects that context of almost all business issues.

CONTENTS

Preface

From the Adaptors' Desk

1. Introduction and Descriptive Statistics

- 1-1 Using Statistics 2
 - Samples and Populations 4
 - Data and Data Collection 4
- 1-2 Percentiles and Quartiles 8
- 1-3 Measures of Central Tendency 10
- 1-4 Measures of Variability 15
- 1-5 Grouped Data and the Histogram 22
- 1-6 Skewness and Kurtosis 24
- 1-7 Relations between the Mean and the Standard Deviation 26
 - Chebyshev's Theorem 26
 - The Empirical Rule 27
- 1-8 Methods of Displaying Data 27
 - Pie Charts 27
 - Bar Charts 28
 - Frequency Polygons and Ogives 28
 - A Caution about Graphs 31
 - Time Plots 31
- 1-9 Exploratory Data Analysis 33
 - Stem-and-Leaf Displays 33
 - Box Plots 35
- 1-10 Using the Computer 40
 - Using Excel for Descriptive Statistics and Plots 40
 - Using MINITAB for Descriptive Statistics and Plots 44
- 1-11 Summary and Review of Terms 46
- Case 1 BSE-Volatility 54

ix

xi

1

2. Probability

- 2-1 Using Statistics 56
- 2-2 Basic Definitions: Events, Sample Space, and Probabilities 58
- 2-3 Basic Rules for Probability 63
 - The Range of Values 63
 - The Rule of Complements 64
 - Mutually Exclusive Events 65
- 2-4 Conditional Probability 67
- 2-5 Independence of Events 73
 - Product Rules for Independent Events 74
- 2-6 Combinatorial Concepts 78
- 2-7 The Law of Total Probability and Bayes' Theorem 81
 - The Law of Total Probability 81
 - Bayes' Theorem 84
- 2-8 The Joint Probability Table 89
- 2-9 Using the Computer 90
 - Excel Templates and Formulas 90
 - Using MINITAB 91
- 2-10 Summary and Review of Terms 93
- Case 2a** Job Applications 98
- Case 2b** Job Applications 99

3. Random Variables

- 3-1 Using Statistics 101
 - Discrete and Continuous Random Variables 106
 - Cumulative Distribution Function 107
 - Simulation: An Application of Random Number 110
- 3-2 Expected Values of Discrete Random Variables 115
 - The Expected Value of a Function of a Random Variable 117
 - Variance and Standard Deviation of a Random Variable 119
 - Variance of a Linear Function of a Random Variable 121
- 3-3 Sum and Linear Composites of Random Variables 121
 - Chebyshev's Theorem 123
 - The Templates for Random Variables 124
- 3-4 Bernoulli Random Variable 127
- 3-5 The Binomial Random Variable 128
 - Conditions for a Binomial Random Variable 129
 - Binomial Distribution Formulas 130
 - The Template 131
 - Problem Solving with the Template 132
- 3-6 Negative Binomial Distribution 135
 - Negative Binomial Distribution Formulas 135
 - Problem Solving with the Template 136

55

4. The Normal Distribution

- 3-7 The Geometric Distribution 137
 - Geometric Distribution Formulas 137
 - Problem Solving with the Template 137
- 3-8 The Hypergeometric Distribution 138
 - Hypergeometric Distribution Formulas 139
 - Problem Solving with the Template 140
- 3-9 The Poisson Distribution 141
 - Problem Solving with the Template 143
- 3-10 Continuous Random Variables 144
- 3-11 The Uniform Distribution 146
 - Problem Solving with the Template 147
- 3-12 The Exponential Distribution 148
 - A Remarkable Property 149
 - The Template 149
 - Value at Risk 151
- 3-13 Using the Computer 152
 - Using Excel Formulas for Some Standard Distributions 152
 - Using MINITAB for Some Standard Distributions 153
- 3-14 Summary and Review of Terms 153
- Case 3** Concepts Testing 165
- 4-1 Using Statistics 167
- 4-2 Properties of the Normal Distribution 168
- 4-3 The Standard Normal Distribution 171
 - Finding Probabilities of the Standard Normal Distribution 172
 - Finding Values of Z Given a Probability 175
- 4-4 The Transformation of Normal Random Variables 177
 - Using the Normal Transformation 179
- 4-5 The Inverse Transformation 185
- 4-6 The Template 190
 - Problem Solving with the Template 191
- 4-7 Normal Approximation of Binomial Distributions 193
- 4-8 Using the Computer 195
 - Using Excel Functions for a Normal Distribution 195
 - Using MINITAB for a Normal Distribution 196
- 4-9 Summary and Review of Terms 197
- Case 4** Acceptable Pins 201
- Case 5** Multicurrency Decision 202
- 5. Sampling and Sampling Distributions**
- 5-1 Using Statistics 205

166

100

204

5-2	Sample Statistics as Estimators of Population Parameters	208
	Random and Nonrandom Sampling	211
	Obtaining a Random Sample	212
	Other Sampling Methods	213
	Nonresponse	214
5-3	Sampling Distributions	216
	The Central Limit Theorem	220
	The History of the Central Limit Theorem	225
	The Standardized Sampling Distribution of the Sample Mean When σ is Not Known	225
	The Sampling Distribution of the Sample Proportion $\hat{p}$	225
5-4	Estimators and Their Properties	229
	Applying the Concepts of Unbiasedness, Efficiency, Consistency, and Sufficiency	231
5-5	Degrees of Freedom	233
5-6	Using the Computer	238
	Using Excel for Generating Sampling Distributions	238
	Using MINITAB for Generating Sampling Distributions	240
5-7	Summary and Review of Terms	241
	Case 6 Acceptance Sampling of Pins	246
6.	Confidence Intervals	247
6-1	Using Statistics	248
6-2	Confidence Interval for the Population Mean When the Population Standard Deviation is Known	249
	The Template	255
6-3	Confidence Intervals for μ When σ is Unknown—The t Distribution	258
	The t Distribution	258
6-4	Large-Sample Confidence Intervals for the Population Proportion p	266
	The Template	268
6-5	Confidence Intervals for the Population Variance	270
	The Template	273
6-6	Sample-Size Determination	275
6-7	The Templates	279
	Optimizing Population Mean Estimates	279
	Determining the Optimal Half-Width	279
	Using the Solver	281
	Optimizing Population Proportion Estimates	281
6-8	Using the Computer	282
	Using Excel Built-In Functions for Confidence Interval Estimation	282
	Using MINITAB for Confidence Interval Estimation	283

6-9	Summary and Review of Terms	283
	Case 7 Presidential Polling	289
	Case 8 Privacy Problem	290
7.	Hypothesis Testing	291
7-1	Using Statistics	292
	The Null Hypothesis	292
7-2	The Concepts of Hypothesis Testing	295
	Evidence Gathering	295
	Type I and Type II Errors	296
	The p -Value	297
	The Significance Level	298
	Optimal α and the Compromise between Type I and Type II Errors	299
	β and Power	300
	Sample Size	300
7-3	Computing the p -Value	302
	The Test Statistic	302
	p -Value Calculations	303
	One-Tailed and Two-Tailed Tests	304
	Computing β	306
7-4	The Hypothesis Test	309
	Testing Population Means	309
	A Note on t Tables and p -Values	310
	The Templates	312
	Testing Population Proportions	314
	Testing Population Variances	316
7-5	Pretest Decisions	328
	Testing Population Means	329
	Manual Calculation of Required Sample Size	330
	Testing Population Proportions	334
	Manual Calculation of Sample Size	335
7-6	Using the Computer	338
	Using Excel for One-Sample Hypothesis Testing	338
	Using MINITAB for One-Sample Hypothesis Testing	340
7-7	Summary and Review of Terms	341
	Case 9 Tiresome Tires I	342
8.	The Comparison of Two Populations	343
8-1	Using Statistics	344
8-2	Paired-Observation Comparisons	345
	The Template	347
	Confidence Intervals	349
	The Template	350

8-3	A Test for the Difference between Two Population Means Using Independent Random Samples	353
	The Templates	357
	Confidence Intervals	359
	The Templates	362
	Confidence Intervals	365
8-4	A Large-Sample Test for the Difference between Two Population Proportions	368
	Confidence Intervals	372
	The Template	373
8-5	The <i>F</i> Distribution and a Test for Equality of Two Population Variances	375
	A Statistical Test for Equality of Two Population Variances	378
	The Templates	382
8-6	Using the Computer	384
	Using Excel for Comparison of Two Populations	384
	Using MINITAB for Comparison of Two Samples	387
8-7	Summary and Review of Terms	388
	Case 10 Tiresome Tires II	393
9.	Analysis of Variance	395
9-1	Using Statistics	396
9-2	The Hypothesis Test of Analysis of Variance	397
	The Test Statistic	398
9-3	The Theory and the Computations of ANOVA	402
	The Sum-of-Squares Principle	405
	The Degrees of Freedom	410
	The Mean Squares	411
	The Expected Values of the Statistics <i>MSTR</i> and <i>MSE</i> under the Null Hypothesis	411
	The <i>F</i> Statistic	412
9-4	The ANOVA Table and Examples	414
9-5	Further Analysis	421
	The Tukey Pairwise-Comparisons Test	424
	Conducting the Tests	426
	The Case of Unequal Sample Sizes, and Alternative Procedures	427
	The Template	427
9-6	Models, Factors, and Designs	429
	One-Factor versus Multifactor Models	430
	Fixed-Effects versus Random-Effects Models	430
	Experimental Design	431
9-7	Two-Way Analysis of Variance	432
	The Two-Way ANOVA Model	434

	The Hypothesis Tests in Two-Way ANOVA	435
	Sums of Squares, Degrees of Freedom, and Mean Squares	435
	The <i>F</i> Ratios and the Two-Way ANOVA Table	437
	The Template	441
	The Overall Significance Level	441
	The Tukey Method for Two-Way Analysis	442
	Extension of ANOVA to Three Factors	443
	Two-Way ANOVA with One Observation per Cell	443
9-8	Blocking Designs	447
	Randomized Complete Block Design	447
	The Template	451
9-9	Using the Computer	453
	Using Excel for Analysis of Variance	453
	Using MINITAB for Analysis of Variance	455
9-10	Summary and Review of Terms	458
	Case 11 Rating Wines	461
	Case 12 Checking Out Checkout	462
	Case 13 Job Satisfaction Study	463
10.	Simple Linear Regression and Correlation	465
10-1	Using Statistics	466
	Model Building	467
10-2	The Simple Linear Regression Model	469
10-3	Estimation: The Method of Least Squares	473
	The Template	480
10-4	Error Variance and the Standard Errors of Regression Estimators	483
	Confidence Intervals for the Regression Parameters	486
10-5	Correlation	490
10-5A	Partial Correlation	495
	Semi-Partial or Part Correlation	497
	Point Bi-Serial Correlation	498
10-6	Hypothesis Tests about the Regression Relationship	500
	Other Tests	503
10-7	How Good is the Regression?	505
10-8	Analysis-of-Variance Table and an <i>F</i> Test of the Regression Model	511
10-9	Residual Analysis and Checking for Model Inadequacies	513
	A Check for the Equality of Variance of the Errors	513
	Testing for Missing Variables	513
	Detecting a Curvilinear Relationship between <i>Y</i> and <i>X</i>	514
	The Normal Probability Plot	516

10-10	Use of the Regression Model for Prediction	522
	Point Predictions	523
	Prediction Intervals	523
	A Confidence Interval for the Average Y , Given a Particular Value of X	525
10-11	Using the Computer	526
	The Excel Solver Method for Regression	526
	The Excel LINEST Function	530
	Using MINITAB for Simple Linear Regression Analysis	533
10-12	Summary and Review of Terms	533
Case 14	Firm Leverage and Shareholder Rights	536
Case 15	Risk and Return	536
Case 16	Regression	538
11.	Multiple Regression	539
11-1	Using Statistics	540
11-2	The k -Variable Multiple Regression Model	540
	The Estimated Regression Relationship	545
11-3	The F Test of a Multiple Regression Model	546
11-4	How Good is the Regression?	550
11-5	Tests of the Significance of Individual Regression Parameters	556
11-6	Testing the Validity of the Regression Model	569
	Residual Plots	569
	Standardized Residuals	570
	The Normal Probability Plot	571
	Outliers and Influential Observations	571
	Lack of Fit and Other Problems	574
11-7	Using the Multiple Regression Model for Prediction	575
	The Template	578
	Setting Recalculation to "Manual" on the Template	578
11-8	Qualitative Independent Variables	579
	Interactions between Qualitative and Quantitative Variables	587
11-9	Polynomial Regression	590
	Other Variables and Cross-Product Terms	595
11-10	Nonlinear Models and Transformations	599
	Variance-Stabilizing Transformations	606
	Regression with Dependent Indicator Variable	607
11-11	Multicollinearity	610
	Causes of Multicollinearity	611
	Detecting the Existence of Multicollinearity	613
	Solutions to the Multicollinearity Problem	617

11-12	Residual Autocorrelation and the Durbin-Watson Test	619
11-13	Partial F Tests and Variable Selection Methods	623
	Partial F Tests	623
	Variable Selection Methods	626
11-14	Using the Computer	630
	Multiple Regression Using the Solver	630
	LINEST Function for Multiple Regression	632
	Using MINITAB for Multiple Regression	634
11-15	Summary and Review of Terms	636
Case 17a	Return on Capital for Four Different Sectors	639
Case 17b	Predicting the Sales	642
12.	Forecasting, Time Series, and Random Numbers	647
12-1	Forecasting	648
12-2	Types of Forecasting	648
12-3	Trend Analysis	649
12-4	Seasonality and Cyclical Behavior	654
12-5	The Ratio-to-Moving-Average Method	658
	The Template	664
	The Cyclical Component of the Series	664
	Forecasting a Multiplicative Series	666
12-6	Exponential Smoothing Methods	668
	The Template	672
12-7	Index Numbers	674
	The Consumer Price Index	678
	The Template	679
12-8	Using the Computer	681
	Using Microsoft Excel in Forecasting and Time Series	681
	Using MINITAB in Forecasting and Time Series	682
12-9	Summary and Review of Terms	684
Case 18	Auto Parts Sales Forecast	685
13.	Statistical Quality Control and Improvement	688
13-1	Using Statistics	690
13-2	Evaluation of Quality Control and contribution of Quality Gurus	691
13-3	Statistics and Quality	692
	Variations and Causes	693
	Control Charts	693
	Process Capability	697

	Pareto Diagrams	698
	Acceptance Sampling	698
	Analysis of Variance and Experimental Design	699
	Taguchi Methods	700
	The Template	700
13-4	The $\bar{x}$ Chart	702
	The Template	704
13-5	The R Chart and the s Chart	706
	The R Chart	706
	The s Chart	707
13-6	The p Chart	710
	The Template	711
13-7	The c Chart	712
	The Template	713
13-8	The x Chart	714
13-9	Using the Computer	715
	Using MINITAB for Quality Control	715
13-10	Summary and Review of Terms	716
	Case 19 Quality Control and Improvement at Nashua Corporation	718
14.	Nonparametric Methods and Chi-Square Tests	720
14-1	Using Statistics	721
14-2	One Sample Non-parametric Tests	722
14-3	One Sample Chi-Square Test for Goodness of Fit	722
	A Goodness-of-Fit Test for the Multinomial Distribution	724
	The Template	725
	Unequal Probabilities	726
	The Template	730
	Kolmogorov-Smirnov (K-S) One Sample Test	731
	One Sample Sign Test	732
14-4	Two Sample Non-Parametric Test	737
	The Sign Test	737
14-5	The Runs Test—A Test for Randomness	742
	Large-Sample Properties	744
	The Template	746
	The Wald-Wolfowitz Test	747
14-6	The Mann-Whitney U Test	750
	The Computational Procedure	751

14-7	The Wilcoxon Signed-Rank Test	757
	The Paired-Observations Two-Sample Test	757
	Large-Sample Version of the Test	759
	A Test for the Mean or Median of a Single Population	760
	The Template	762
14-8	The Kruskal-Wallis Test—A Nonparametric Alternative to One-Way ANOVA	765
	The Template	767
	Further Analysis	770
14-9	The Friedman Test for a Randomized Block Design	773
	The Template	776
14-10	The Spearman Rank Correlation Coefficient	778
	The Template	781
14-11	Contingency Table Analysis—A Chi-Square Test for Independence	783
	The Template	787
14-12	A Chi-Square Test for Equality of Proportions	790
	The Median Test	793
14-13	Using the Computer	796
	Using MINITAB for Nonparametric Tests	796
14-14	Summary and Review of Terms	798
	Case 20 The Nine Nations of North America	800
15.	Bayesian Statistics and Decision Analysis	803
15-1	Using Statistics	804
15-2	Bayes' Theorem and Discrete Probability Models	805
	The Template	810
15-3	Bayes' Theorem and Continuous Probability Distributions	814
	The Normal Probability Model	814
	Credible Sets	816
	The Template	819
15-4	The Evaluation of Subjective Probabilities	820
	Assessing a Normal Prior Distribution	820
15-5	Decision Analysis: An Overview	822
	Actions	823
	Chance Occurrences	823
	Probabilities	823
	Final Outcomes	823
	Additional Information	823
	Decision	824
15-6	Decision Trees	825
	The Payoff Table	826

- 15-7 Handling Additional Information Using Bayes' Theorem 835
 - Determining the Payoffs 837
 - Determining the Probabilities 837
- 15-8 Analytic Hierarchy Process 847
 - Pair-wise Comparison 847
 - Matrix Formation 848
 - Priority Vectors 849
 - AHP: An Application 851
- 15-9 Utility 856
 - A Method of Assessing Utility 858
- 15-10 The Value of Information 860
- 15-11 Using the Computer 864
 - The Template 864
- 15-12 Summary and Review of Terms 865
- Case 21 Pizzas 'R' Us 868
- Case 22 New Drug Development 869

Appendix A: References	871
Appendix B: Answers to Most Odd-Numbered Problems	875
Appendix C: Statistical Tables	901
Appendix D: Tables	938
Index	943

On the CD

16. Sampling Methods

- 16-1 Using Statistics
- 16-2 Nonprobability Sampling and Bias
- 16-3 Stratified Random Sampling
 - Practical Applications
 - Confidence Intervals
 - The Template
 - Stratified Sampling for the Population Proportion
 - The Template
 - What Do We Do When the Population Strata Weights Are Unknown?
 - How Many Strata Should We Use?
 - Postsampling Stratification
 - Optimum Allocation
 - The Template
- 16-4 Cluster Sampling
 - The Relation with Stratified Sampling
 - Single-Stage Cluster Sampling for the Population Mean

- Single-Stage Cluster Sampling for the Population Proportion
- The Templates
- Two-Stage Cluster Sampling
- 16-5 Systematic Sampling
 - The Advantages of Systematic Sampling
 - Estimation of the Population Mean in Systematic Sampling
 - The Template
- 16-6 Nonresponse
- 16-7 Summary and Review of Terms
- Case 23 The Boston Redevelopment Authority

17. Multivariate Analysis

- 17-1 Using Statistics
- 17-2 The Multivariate Normal Distribution
- 17-3 Discriminant Analysis
 - Developing a Discriminant Function
 - Evaluating the Performance of the Model
 - Discriminant Analysis with More Than Two Groups
- 17-4 Principal Components and Factor Analysis
 - Principal Components
 - The Extraction of the Components
 - Factor Analysis
 - The Extraction of Factors
 - The Rotation of Factors
- 17-5 Using the Computer
 - Using MINITAB for Discriminant and Principal Component Analysis
- 17-6 Summary and Review of Terms
- Case 24 Predicting Company Failure

Working with Templates

- The Idea of Templates
- Working with Templates
- The Autocalculate Command
- The Data Table Command
- The Goal Seek Command
- The Solver Macro
- Some Formatting Tips
- Saving the Templates

Introduction to Excel Basics

- Excel Formulas
- Excel Functions
- The Need for Templates

INDEX

Page numbers followed by n indicate material found in notes.

A

Absolute frequencies, 22
Absolute kurtosis, 22
Absolute zero, 3
Acceptable Pins (case), 202-203
Acceptance Sampling of Pins (case), 246
Actions, 823
Aczel, A. D., 97n, 863n
Additive factors, 383, 568
Adjusted multiple coefficient of determination, 552
Aizenman, Joshua, 654n
All possible regressions, 626
Alternative actions, 824
Alternative hypothesis, 292-294
Analysis of covariance, 509
Analysis of variance (ANOVA), 233, 396-458
 ANOVA diagram, 422
 ANOVA table and examples, 414-419
 ANOVA table for regression, 511
 assumptions of, 398
 blocking designs, 431
 completely randomized design, 431
 computer use and, 453-458
 confidence intervals, 423-424
 defined, 396
 degrees of freedom, 410-411
 error deviation, 405
 Excel for, 453-454
 experimental design, 431
 F statistic, 412-413
 fixed-effects vs. random-effects models, 431
 further analysis, 421-423
 grand mean, 402, 407
 hypothesis test of, 397-402

 main principle of, 403
 mean squares, 411-412
 MINITAB for, 456
 models, factors, and designs, 430-431
 multiple regression, 549, 554
 one-factor model, 430
 one-factor vs. multifactor models, 430-431
 principle of, 405
 random-effects model, 430
 randomized complete block design, 430, 447-450
 repeated-measures design, 450
 sum-of-squares principle, 405-410
 template (single-factor ANOVA), 429
 test statistic for, 398-402, 413
 theory and computations of, 402-405
 three factors extension, 443
 total deviation of data point, 408
 treatment deviation, 405
 Tukey pairwise-comparisons test, 424-427
 two-way ANOVA, 432-435
 two-way ANOVA with one observation per cell, 443-445
 unequal sample size, 427
ANOVA; *see* Analysis of variance (ANOVA)
ANOVA table, 414-419
ANOVA test statistic, 398-402
Arithmetic mean, 10
Asimov, Eric, 284
Auto Parts Sales Forecasts (case), 685-687
Autocorrelation, 619
Average, 10; *see also* Mean
Averaging out and folding back, 827

B

Backward elimination, 627
Bailey, Jeff, 782n
Baker-Said, Stephanie, 741n

Balanced design, 427
 Baland, J. M., 556n
 Banner, Katie, 361n
 Bar charts, 28, 42
 probability bar chart, 103
 Barbaro, Michael, 323n
 Barenghi, C., 863n
 Barr, Susan Learner, 289n
 Barrionuevo, Alexei, 209n
 Base period, 675
 Basic outcome, 60
 Bayes, Thomas, 81
 Bayes' Theorem, 81–82
 additional information and, 835–837
 continuous probability distributions, 814–819
 determining the payoff, 837
 determining the probabilities, 837–841
 discrete probability models, 805–812
 extended Bayes' Theorem, 86–88
 normal probability model, 820–821
 Bayesian analysis, 804–805
 Bayesian statistics, 804–806
 advantages of approach, 809
 classical approaches vs., 805
 computer usage for, 864–865
 subjective probabilities, evaluation of, 820–821
 template for, 810–811
 Bearden, William O., 429n
 Beaver, William H., 511n
 Bell-shaped normal curve, 168
 Berdahl, Robert M., 56
 Berenson, Alex, 833n
 Bernoulli, Jakob, 128
 Bernoulli distribution, 128
 Bernoulli process, 129
 Bernoulli random variable, 128
 Bernoulli trial, 128
 Bertrand, Marianne, 597n
 Best, R., 863n
 Best linear unbiased estimators (BLUE), 473
 B computation of, 306–308
 B and power of test, 300
 Between-treatments deviation, 408
 Bias, 205, 224–231
 nonresponse bias, 5, 205
 Bigda, Caroline, 28n
 Billett, Matthew T., 589n
 Binary variable, 580
 BINOMDIST function, 152
 Binomial distribution, 79, 131

MINITAB for, 153
 negative binomial distribution, 135–137
 normal approximation of, 193–194
 population proportions, 311
 template for, 131–132
 Binomial distribution formulas, 130–131
 Binomial distribution template, 131
 Binomial probability formula, 130
 Binomial random variable, 103, 129–132
 conditions for, 129–130
 Binomial successes, 209
 Biscourp, Pierre, 568n
 Block, 448
 Blocking, 350
 Blocking designs, 431, 447–451
 randomized complete block design, 447–450
 repeated-measures design, 450
 BLUE (best linear unbiased estimators), 473
 Bonferroni method, 427
 Box-and-whisker plot, 35
 Box plots, 35–38, 42
 elements of, 35–36
 uses of, 38
 Brav, James C., 554n
 Briley, Donnel A., 420n
 Brooks, Rick, 323n
 Bruno, Mark, 765n
 Bukey, David, 362n
 Burros, Marian, 763n
 Bush, Jason, 257n
 Business cycle, 721

C

c chart, 712–713
 Caesar, William K., 658n
 Callbacks, 215
 Capital Asset Pricing Model (CAPM), 527
 Carey, John, 359n
 Carlson, Jay P., 429n
 Carter, Erin, 749n
 Cases
 Acceptable Pins, 201–202
 Acceptance Sampling of Pins, 246
 Checking Out Checkout, 462
 Concepts Testing, 165
 Firm Leverage and Shareholder Rights, 536
 Job Applications, 98–99
 Multicurrency Decision, 202–203
 New Drug Development, 869–870
 Nine Nations of North America, 800–802
 Pizzas "R" Us, 868–869
 Presidential Polling, 289
 Privacy Problem, 290
 Quality Control and Improvement at Nashua Corporation, 718–719
 Rating Wines, 461
 Return on Capital for Four Different Sectors, 639–641
 Risk and Return, 536
 Tiresome Tires I, 342
 Tiresome Tires II, 394
 Casey, Susan, 393n
 Cassidy, Michael, 772n
 Categorical variable, 3
 Causality, 494
 Center of mass, 118
 Central limit theorem, 220–225, 249
 effects of, 222
 history of, 225
 population standard deviation and, 225
 sample size and, 222
 Central tendency; *see* Measures of central tendency
 Centrality of observations, 10, 116
 Chance node, 824
 Chance occurrences, 823
 Chance outcome, 836
 Chart Wizard, 41
 Charts; *see* Methods of displaying data
 Chatzky, Jean, 195, 289, 338n
 Chebyshev's theorem, 26, 123–124
 Checking Out Checkout (case), 462
 Chi-square distribution, 270, 283
 mean of, 270
 values and probabilities of, 272
 Chi-square random variable, 378
 Chi-square test for equality of proportions, 790–793
 CHIINV function, 283
 Christen, Markus, 429n, 555n, 637n
 Classes, 22
 Classical approach, 804–805
 Classical probability, 57
 Cluster, 214
 Cluster sampling, 214
 Coefficient of determination (r^2), 505–511
 Collinearity, 610–611; *see also* Multicollinearity
 Combinations, 79, 91
 Combinatorial concepts, 78–80
 Comparison of two populations, 344–388
 computer templates for, 385–387

difference (population-means/independent random samples), 353–366
 equality of two population variances, 379–384
 F distribution, 376–378
 large-sample test (two population proportions), 368–374
 paired-observation comparisons, 345–350
 Complement, 58–59
 rule of complements, 64
 Completely randomized design, 406, 431
 Computational formula for the variance of a random variable, 119
 Computers; *see also* Excel; Templates
 bar charts, 42
 Bayesian statistics/decision analysis, 864–865
 box plots, 42
 confidence interval estimation, 282–284
 decision analysis, 864–865
 for descriptive statistics and plots, 40–46
 in forecasting and time series, 681–684
 histograms, 40–41
 hypothesis testing, 339–341
 multiple regression using Solver, 630–634
 normal distribution, 195–196
 one-way ANOVA, 453
 paired-difference test, 385–387
 percentile/percentile rank computation, 40
 pie charts, 41
 probability, 90–91
 for quality control, 715–716
 sampling distributions, 238–241
 scatter plots, 44
 for standard distributions, 153
 time plots, 44
 two-way ANOVA, 453–454
 Concepts Testing (case), 165
 Concomitant variables, 586
 Conditional probability, 67–70, 82
 Confidence, 248
 Confidence coefficient, 252
 CONFIDENCE function, 282
 Confidence intervals, 190, 248–284
 Bayesian approach, 249n
 classical/frequentist interpretation, 249n
 defined, 248
 80% confidence interval, 254
 Excel functions for, 282–284
 expected value of Y for given X, 522
 half-width, determining optimal, 280
 important property of, 253

individual population means, 473
 MININTAB for, 283-285
 95% confidence interval, 250-252
 paired-observation comparisons, 349-350
 population mean (known standard deviation), 249-256
 population means, difference between, 354
 population proportion (large sample), 266-268
 population proportions, difference between, 372
 population variance, 270-273
 regression parameters, 486-488
 sample-size determination, 275-278
t distribution, 258-263
 templates, 255-256
 Confidence level, 253, 299
 Conlin, Michelle, 328n
 Consistency, 231
 Contingency table, 69, 783-784
 Contingency table analysis, 783-786
 chi-square test for independence, 783-785
 chi-square test statistic for independence, 784
 degrees of freedom, chi-square statistic, 784
 expected count in cell, 785
 hypothesis test for independence, 784
 template, 787-788
 Yates correction, 787
 Continuity correction, 193-194
 Continuous probability distributions, Bayes' theorem
 and, 814-819
 Continuous random variable, 106-107, 144-146
 Control chart, 695-697
 centerline, 706-707, 712
 lower control limit (LCL), 704, 706-707, 712
 for process mean, 703
 for process proportion, 711
 upper control limit (UCL), 704, 706-707, 712
 Control treatment (placebo), 397
 Cook, R. Dennis, 592n
 Cordoba, Jose de, 374
 Correlation, 490-494
 Correlation analysis, 490
 Correlation coefficient, 490
 Correlation matrix, 613-614
 Counts of data points, 22
 Covariance, 491
 CPI (Consumer price index), 674, 678-679
 Creamer, Matthew, 413n
 Credible sets, 806, 816-819
 Creswell, Julie, 114n
 Crockett, Roger O., 94n
 Cross-product terms, 595-596

Cumulative distribution function, 107-109
 Cumulative frequency plots (ogives), 28, 30
 Cumulative probability function, 108
 Curved trends, 650
 Curvilinear relationship, 471, 514-515
 Cveykus, Renee, 749n
 Cycle, 654
 Cyclical behavior, 654-658
 Cyclical variation, 655

D

Darwin, Charles, 466
 Dash, Eric, 286n
 Data, 3, 5
 grouped data, 22-24
 Data collection, 4
 Data set, 4, 116
 de Fermat, Pierre, 57
 de Finetti, Bruno, 57n
 de Mère, Chevalier, 57
 de Moivre, Abraham, 57, 168, 225
 Decision, 824, 826
 Decision analysis, 805, 822-825
 actions, 823
 additional information, 823-824
 chance occurrences, 823
 decision, 824
 decision tree, 825-833
 elements of, 822
 final outcomes, 823
 overview of, 822-825
 payoff table, 826-829
 probabilities, 823
 utility, 856-859
 value of information, 860-863
 Decision node, 824
 Decision tree, 825-833
 Deflators, 676
 DeGraw, Irv, 554n
 Degree of linear association, 490-492
 Degrees of freedom (df), 225, 233-237
 ANOVA and, 410-411, 436-437
 sum-of-squares for error (SSE), 411
 sum-of-squares total (SST) and, 411
 sum-of-squares for treatment (SSTR), 411
 Degrees of freedom of the denominator, 376
 Degrees of freedom of the numerator, 376
 Demers, Elizabeth, 534n
 Deming's 14 Points, 694-695

Dependent indicator variable, regression with,
 607-608
 Dependent variable, 466
 Descriptive graphs, 28
 Descriptive statistics, 3-46, 205n
 computer use for, 40-44
 exploratory data analysis, 33-38
 grouped data and histogram, 22-24
 mean-standard deviation relations, 26-27
 measures of central tendency, 10-14
 measures of variability, 10-14
 methods of displaying data, 27-33
 MINITAB for, 44-46
 percentiles and quartiles, 8-9, 40
 random variable, 101-105
 skewness and kurtosis, 24-26, 37
 templates for random variables, 124-125
 Deseasonalizing a time series, 661-663
 df, *see* Degrees of freedom (df)
 Diffuse priors, 816
 Discrete probability models, 805-806
 Discrete random variable, 106-107
 Bayes' theorem for, 806
 cumulative distribution function of, 108
 expected values of, 117-121
 probability distribution of, 107
 variance of, 119
 Disjoint sets, 59
 Dispersion, 15-16, 120; *see also* Measures of variability
 Displaying data; *see* Methods of displaying data
 Distribution of the data, 9
 Distribution-free methods, 798; *see also* Nonparametric
 tests
 Distributions; *see also* Normal distribution; Probability
 distribution
 Bernoulli distribution, 128
 cumulative distribution function, 107-111
 exponential distribution, 148-152
 geometric distribution, 137-138
 hypergeometric distribution, 138-141
 kurtosis of, 24-26
 Poisson distribution, 141-143
 sampling distributions, 216-227
 skewness of, 24-25
 uniform distribution, 146-148
 Dow Jones Industrial Average, 674-675
 Dummy variable, 580, 583, 657
 Dummy variable regression technique, 657
 Durbin-Watson test, 513, 619-622
 Durbin-Watson test statistic, 620

E

Eccles, Robert G., 668n
 EDA (Exploratory data analysis), 33-38
 Efficiency, 229, 230, 865
 80% confidence interval, 254
 Elementary event, 60
 Elements of a set, 58
 Elliot, Stuart, 368n
 Empirical rule, 27, 186n
 Empty set, 58
 Epstein, Edward, 75n
 Error deviation, 405, 407
 Error probability, 252
 Estimated regression relationship, 472
 Estimators, 208-209, 229
 consistency of, 224, 231
 efficiency of, 209, 231
 of population parameter, 209-210
 properties of, 229-232, 472
 sufficiency of, 229, 231
 as unbiased, 229-231
 Event, 60, 895
 EVPI (expected value of perfect information), 860
 Excel; *see also* Solver Macro
 ANOVA and, 453-454
 Bayesian revision of probabilities, 91
 descriptive statistics and plots, 27-46
 F-test, 387
 in forecasting and time series, 681-682
 graphs, 31
 histograms, 40-41
 LINEST function, 530-532
 normal distribution, 195-197
 one-sample hypothesis testing, 338-340
 paired-difference test, 385-387
 percentile/percentile rank computation, 40
 probabilities, 90-91
 Random Number Generation analysis, 240
 regression, 526-529, 533
 Sampling analysis tool, 239
 sampling distributions and, 238-241
 standard distributions and, 152-153
 t-test, 387
 Excel Analysis Toolpack, 40
 Expected net gain from sampling, 862
 Expected payoff, 827, 831-832
 Expected value of a discrete random variable, 115-116
 Expected value of a function of a random variable,
 115-116
 Expected value of a linear composite, 122

- Expected value of a linear function of a random variable, 118
 - Expected value of perfect information (EVPI), 860
 - Expected value of sample mean, 218
 - Expected value of the sum of random variables, 122
 - Experiment, 60
 - Experimental design, 431, 699
 - Experimental units, 350
 - Explained deviation, 506–507
 - Explained variation, 410, 507
 - Exploratory data analysis (EDA), 33–38
 - box plots, 35–38
 - stem-and-leaf displays, 34–35
 - EXPONDIST function, 153
 - Exponential distribution, 148–152
 - common examples of, 148–149
 - remarkable property of, 149
 - template for, 149–151, 152
 - Exponential model, 603
 - Extended Bayes' Theorem, 86
 - Extra sum of squares, 624
 - Extrapolation, 574
- F**
- F* distribution, 375–378, 398, 512
 - degrees of freedom of the denominator, 375, 398
 - degrees of freedom of the numerator, 375, 398
 - equality of two population variances, 379–380
 - templates for, 384
 - F* ratio, two-way ANOVA, 437
 - F* statistic, 412–414
 - F* test, 357, 387
 - multiple regression model, 546–549
 - partial *F* tests, 623–626
 - of regression model, 511–512, 517
 - Factor, 429
 - Factorial, 78, 91
 - Fair games, 117
 - Fairley, W., 292n
 - Farley, Amy, 112n
 - Farzad, R., 49n
 - Fass, Allison, 264n, 398n
 - Feller, W., 78, 225, 741
 - Ferry, John, 33n, 77n
 - Fialka, John J., 286n
 - 50th percentile, 9
 - Final outcomes, 823
 - Firm Leverage and Shareholder Rights (case), 536–537

- First quartile, 9
- Fisher, Anne, 287n
- Fisher, Sir Ronald A., 375, 396
- Fixed-effects vs. random-effects models, 430
- Flight simulators, 34
- Fong, Eric A., 434n
- Forbes, Malcolm, 3
- Forbes, Steve, 820n
- Forecasting
 - Excel/MINITAB in, 681–684
 - exponential smoothing methods, 661–674
 - index numbers, 674–680
 - multiplicative series, 666–668
- Forward selection, 626–627
- Frame, 8, 211
- Frequency, 22
- Frequency distribution, 208
- Frequency polygon, 28–30
- Frequentist approach, 804
- Friedman test, 451
 - data layout for, 774
 - null and alternative hypotheses of, 714
 - template, 776–777
 - test statistic, 775
- Fulcrum, 11
- Full model (*F* test), 623–624

- G**
- Galilei, Galileo, 57
 - Galton, Sir Francis, 466
 - Gambling models, 57
 - Ganguly, Ananda, 447n
 - Gauss, Carl Friedrich, 168
 - Gauss-Markov theorem, 473
 - Gaussian distribution, 168
 - Generalized least squares (GLS), 622
 - Geometric distribution, 137–138
 - formulas for, 137
 - template for, 138
 - Geometric progression, 137
 - Gleason, Kimberly C., 536n, 566n
 - GLS (Generalized least squares), 622
 - Goal seek command, 132, 140, 189
 - Goldstein, Matthew, 757n
 - Gomez, Paulo, 535n
 - Good, I. J., 56
 - Goodness-of-fit test, 785
 - for multinomial distribution, 786–787
 - Goodstein, Laurie, 5n

- Gossett, W. D., 258
- Grand mean, 403, 407, 430
- Graphs; *see* Methods of displaying data
- Gray, Patricia B., 287n
- Green, Heather, 270n
- Grouped data, 22–24
- Grover, Ronald, 254n
- Gruley, Bryan, 164n

H

- Hall, Kenji, 269n
- Hammand, S., 28
- Hansell, Saul, 327n
- Hardesty, David M., 429n
- Harris, Elizabeth, 228n, 318n
- Harris, Marlys, 318n, 324n
- HDP (highest-posterior-density), 816
- Hellmich, Nancy, 76n
- Helm, Burt, 96n, 215n
- Herbold, Joshua, 447n
- Heteroscedasticity, 513, 570, 578
- Highest-posterior-density (HPD), 816
- Hinges (of box plot), 35–36
- Histogram, 22–24, 28, 40–41
- Holson, Laura M., 351n
- Homogeneity, tests of, 791
- Hovanesian, Marader, 50n
- HSD (honestly significant differences) test, 425
- Huddleston, Patricia, 534n
- Hui, Jerome Kuch Swee, 458n
- Hypergeometric distribution, 138–141
 - formulas for, 139–140
 - problem solving with template, 140–141, 153
 - schematic for, 139
- HYPGEOMDIST function, 153
- Hypothesis, 292
- Hypothesis testing, 292–341, 344
 - alternative hypothesis, 294–295, 401, 402
 - ANOVA, 397–402
 - association between two variables, 779
 - B* and power of test, 308
 - common types of, 309
 - computing *B*, 306
 - concepts of, 295–301
 - confidence level, 299–300
 - evidence gathering, 295, 296
 - Excel/MINITAB for, 338–340
 - for independence, 784
 - individual regression slope parameters, 558

- Kruskal-Wallis test, 766
- left-tailed test, 304–306
- linear relationship between *X* and *Y*, 500
- median test, 793
- null hypothesis, 292–293, 400–402
- one-tailed and two-tailed tests, 304–306
- operating characteristic (OC) curve, 332–333
- optimal significance level, 299–300
- p*-value, 297–298, 300
- p*-value computation, 302–303
- paired-observations two-sample test, 757
- population means, 309–310
- population proportions, 314–315, 334–335
- population variance, 316–317
- power curve, 331–332, 335
- power of the test, 300
- pretest decisions, 328–329
- regression relationship, 495–505
- required sample size (manual calculation), 330–331, 335
- right-tailed test, 304, 308
- sample size, 300–301, 335
- significance level, 298–299
- t* tables, 316
- templates, 312–313
- test statistic, 310
- two-tailed test, 304, 305, 308
- two-way ANOVA, 435, 439
- type I/II errors, 296–297, 299–300

I

- Ihlwan, Moon, 269n
- Independence of events, 73–76;
 - see also* Contingency table analysis
 - conditions for, 73
 - product rules for, 73–75
- Independent events, 76
- Independent variable, 466
- Indicator variable, 580–589
- Indifference, 859
- Inferential statistics, 57, 205
- Influential observation, 574
- Information, 3
 - expected net gain from sampling, 862
 - expected value of perfect information (EVPI), 860
 - qualitative vs. quantitative, 3
 - value of, 860–863
- Initial run, 702
- Inner fence, 36

Interaction effects, 433-434, 787
 Interrival time, 149
 Intercept, 478, 480
 Interquartile range (IQR), 9, 15-16, 36
 Intersection, 59
 Intersection rule, 75
 Interval estimate, 209
 Interval scale, 3
 Intrinsically linear models, 599
Introduction to Probability Theory and Its Applications (Feller), 742
 Inverse transformation, 179, 185-188
 Irregular components models, 684

J

Jiraporn, Pornsit, 536n, 566n
 Jo, Hoje, 567n
 Job Applications (case), 98-99
 Johar, Gita Venkataramani, 445n
 Johnson, George, 804, 804n
 Johnston, J., 615n
 Joint confidence intervals, 487
 Joint density, 815
 Joint hypothesis, 397
 Joint probability, 65
 Joint probability table, 89
 Joint test, 397
 Joos, Philip, 534n
 Josephy, N. H., 97n
 Juran, J. M., 698

K

k -variable multiple regression model, 540-546
 Kacperczyk, Marcin, 421n
 Kang, Jun-Koo, 789n
 Kendall's tau, 781
 Keynes, John Maynard, 3
 Kim, Yongtae, 567n
 Kimball's inequality, 441
 King, Tao-Hsien Dolly, 589n
 Kirkland, R., 33n
 Knapp, Volker, 266n, 322n
 Knox, Noelle, 391n
 Kondratieff definition, 721
 Kramarz, Francis, 568n
 Kranhold, Kathryn, 286n
 Krishnamurthy, Arvind, 555n
 Kroll, Lovisa, 265n, 328n
 Kruskal-Wallis test, 398, 430, 765-771

further analysis, 770-771
 template for, 767-770
 test statistic, 766
 Kurtosis, 24-25
 Kwon, Young Sun, 421n

L

Lack of fit, 574-575
 Lamey, Lien, 472n
 Large sample confidence intervals for population proportion, 368
 Large-sample properties, 744
 Lav, Kong Cheen, 637n
 Law of total probability, 81-84
 LCL (lower control limit), 704, 706-707, 713
 Least-squares estimates, 543-544
 Lee, Alan J., 592n
 Lee, Hyun-Joo, 534n
 Lee, Louise, 795n
 Left-skewed distribution, 24
 Left-tailed test, 304-306
 Lehman, Paula, 795n
 Leptokurtic distribution, 26
 Lerner, Josh, 696n
 Lettav, Martin, 535n
 Level of significance, 298-300
 Li, Peter Ping, 833n
 Likelihood function, 805-806
 Linear composite, 121-125
 expected value of, 122
 LINEST function, 530-532, 632
Literary Digest presidential poll, 205-208
 Lo, May Chiun, 459n
 Location of observations, 10, 117
 Logarithmic model, 603
 Logarithmic transformation, 599-602, 607
 Logistic function, 608
 Logistic regression model, 607
 Loss, 823
 Loss function, 700
 Lower control limit (LCL), 704, 706-707, 711
 Lower quartile, 9

M

Malkiel, Burton G., 228n
 Mann-Whitney U test, 357, 750-756
 computational procedure, 751
 MINITAB for, 755-756
 null and alternative hypothesis for, 750
 U statistic, 751

Manual recalculation, 578-579
 Marcial, Gene G., 245n
 Margin of error, 250
 Marginal probabilities, 89
 Marketing research, 5
 Martin, Mitchell, 184n
 Mauer, David, 589n
 Mean, 10-13, 116
 defined, 10
 extreme observations and, 12
 grand mean, 403
 population mean, 11, 16, 208
 sample mean, 10, 218, 220, 403
 standard deviation and, 26-27
 Mean square error (MSE), 411-412
 multiple regression, 550-551
 simple linear regression, 484
 Mean square treatment (MSTR), 411-412
 Mean time between failures (MTBF), 149
 Measurements, 3
 scales of, 3
 Measures of central tendency, 10-14
 mean, 10-13
 median, 9-13
 mode, 10-13
 Measures of variability, 15-21, 115
 interquartile range, 9, 15-16
 range, 16
 standard deviation, 16
 variance, 16
 Median, 9-12
 Median test, 793
 Mehring, James, 367n
 Method of least squares, 473
 Methods of displaying data, 27-32
 bar charts, 28
 cautionary note to, 31
 exploratory data analysis (EDA), 33-38
 frequency polygons, 28-30
 histogram, 22-24, 28, 40-41
 ogives, 28-30
 pie charts, 27-28, 41
 time plots, 31
 Middle quartile, 9
 MINITAB
 ANOVA and, 455-458
 comparison of two samples, 387-388
 confidence interval estimation, 283-284
 for descriptive statistics/plots, 44-46
 for factorial, combination, and permutation, 91

in forecasting and time series, 682-684
 Mann-Whitney test, 755
 multicollinearity, 613
 multiple regression, 634-636
 nonparametric tests, 796-798
 normal distribution, 196-197
 one-sample hypothesis testing, 340-341
 for quality control, 715-716
 regression analysis, 574
 sampling distributions, 240-241
 simple linear regression analysis, 533
 standard distributions, 153
 stepwise regression, 628
 Missing variables test, 513
 Mode, 11-13, 24
 Montgomery, D., 578n
 Moskin, Julia, 48n
 Mosteller, F., 292n
 Mound-shaped distribution, 26
 MSE; *see* Mean square error (MSE)
 MSTR (mean square treatment), 411-412
 MTBF (mean time between failures), 149
 Mukhopadhyay, Anirban, 445n
 Multicollinearity, 557-558, 610-617
 causes of, 592, 611-612
 detecting existence of, 613-616
 effects of, 616
 solutions to problem, 617
 Multicollinearity set, 611
 Multicurrency Decision (case), 202-203
 Multifactor ANOVA models, 430
 Multiple coefficient of determination (R^2), 551
 Multiple correlation coefficient, 551
 Multiple regression, 466, 544-636
 adjusted multiple coefficient of determination, 552
 ANOVA table for, 549, 553
 assumptions for model, 541
 cross-product terms, 595-596
 decomposition of total deviation, 547
 dependent indicator variable and, 607-608
 Durbin-Watson test, 619-622
 estimated regression relationship, 545-546
 F test, 546-549
 how good is the regression, 550-554
 influential observation, 574
 k -variable model, 540-546
 lack of fit and other problems, 574
 least-squares regression surface, 545
 LINEST function for, 632-634
 mean square error (MSE), 550

measures of performance of, 553
 MINITAB and, 634
 multicollinearity, 557–558, 610–617
 multiple coefficient of determination R^2 , 550–551
 multiple correlation coefficient, 551
 nonlinear models and transformations, 599–608
 normal equations, two independent variables, 542
 normal probability plot, 571
 other variables, 595–596
 outliers and influential observations, 571–574
 partial F test, 623–626
 polynomial regression, 590–596
 prediction and, 575–579
 qualitative independent variables, 579–589
 qualitative/quantitative variables interactions, 587–589
 residual autocorrelation, 619–621
 residual plots, 569–570
 significance of individual regression parameters, 556–566
 Solver, 630–634
 standard error of estimate, 551
 standardized residuals, 570–573
 template for, 544, 561, 565, 578, 594
 validity of model, 569–574
 variable selection methods, 626–628
 Multiplicative model, 599
 Multistage cluster sampling, 214
 Murphy, Dean E., 56n
 Mutually exclusive events, 65, 76–77
 Mutually independent, 122

N

N factorial ($n!$), 78
 Negative binomial distribution, 135–137
 problem solving with template, 136–137, 153
 Negative correlation, 491
 Negative skewness, 25
 NEGBINOMDIST function, 152
 Nelson, Lloyd S., 703n, 718n
 Nelson, Melissa, 772n
 Net regression coefficients, 544
 New Drug Development (case), 869–870
 Newquist, Scott C., 668n
 Newton, Sir Isaac, 690
 Nine Nations of North America (case), 800–802
 95% confidence interval, 251–253
 Nominal scale, 3
 Noninformative, 804
 Nonlinear models, 590, 599–608

Nonparametric tests, 357, 721–799
 defined, 721
 Friedman test, 773–777
 Kruskal-Wallis test, 765–771
 Mann-Whitney U test, 750–756
 median test, 793
 MINITAB for, 796–798
 paired-observations two-sample test, 757–759
 runs test, 742–746
 sign test, 721–728
 Spearman rank correlation coefficient, 778–782
 summary of, 798
 Wald-Wolfowitz test, 747–748
 Wilcoxon signed-rank test, 757–762
 Nonresponse, 214–215
 Nonresponse bias, 5, 205
 Normal approximation of binomial distributions, 193–194
 template, 196
 Normal distribution, 167–197; *see also* Standard normal distribution
 absolute kurtosis of, 25
 Excel functions for, 195–197
 inverse transformation, 185–188
 MINITAB for, 196
 normal approximation of binomial distributions, 193–194
 probability density function, 167
 properties of, 168–171
 sampling and, 218, 225
 standard normal distribution, 172–176
 template for, 190–192
 testing population proportions, 314
 transformation of normal random variables, 177–183
 Normal equations, 476
 Normal prior distribution, 820–821
 Normal probability model, 814–821
 Normal probability plot, 516–518
 Normal random variables
 inverse transformation, 185–188
 inverse transformation of Z to X , 179
 obtaining values, given a probability, 188
 transformation of X to Z , 177, 183
 transformation of, 177–183
 using the normal transformation, 179
 Normal sampling distribution, 219–220
 NORMDIST function, 195
 NORMSINV function, 283
 Null hypothesis, 292, 400–402
 Chi-square test for equality of proportions, 790

of Friedman test, 773
 Mann-Whitney U test, 750

O

Objective probability, 57
 OC (Operating characteristic curve), 332
 Odds, 64
 Ogives, 28–30
 1 standard deviation, 26
 One-factor ANOVA model, 429
 Excel/MINITAB for, 453, 456
 multifactor models vs., 430
 One-tailed test, 304–305
 One-variable polynomial regression model, 590
 Operating characteristic curve (OC curve), 332–333
 Optimal decision, 866
 Optimal sample size, 342
 Ordinal scale, 3
 Ordinary least squares (OLS) estimation method, 570
 Outcomes, 60
 Outer fence, 36
 Outliers, 13, 37, 571–574

P

p chart, 710–711
 template for, 711–712
 p -value, 297–298, 310
 computation of, 302–303
 definition of, 298
 test statistic, 302
 Paired-observation comparisons, 345–350
 advantage of, 345
 confidence intervals, 349–350
 Excel for, 385–387
 template for, 347–350
 test statistic for, 347
 Paired-observation t test, 345–347
 Paired-observations two-sample test, 757
 Palmeri, Christopher, 392n
 Parameters, 208, 798
 Pareto diagrams, 698
 template for, 700
 Park, Myung Seok, 567n
 Parsimonious model, 468
 Partial F statistic, 624
 Partial F tests, 623–626
 Partition, 82–83
 Pascal, Blaise, 57
 Passy, Charles, 287n
 Payoff, 823, 837
 Payoff table/matrix, 826–829
 Pearson product-moment correlation coefficient, 491, 779
 Peck, F., 578n
 Peecher, Mark E., 447n
People v. Collins, 292
 Percentile, 8–9, 40
 Percentile rank computation, 40
 Pereira, Pedro, 673n
 Permutations, 79, 91
 Personal probability, 58
 Peters, Ruth, 163n
 Phav, Ian, 637n
 Pie chart, 27, 41
 Pissaeides, Christopher A., 549n
 Pizzas “R” Us (case), 868–869
 Platykurtic distribution, 26
 Point estimate, 208
 Point predictions, 523
 Poisson distribution, 141–143
 formulas for, 142
 problem solving with template, 143–153
 Poisson formula, 142
 POISSON function, 153
 Polynomial regression, 590–596
 Pooling, 791
 Population, 4, 205, 218, 396
 defined, 4, 208
 sampling from the, 4, 74, 206
 Population correlation coefficient, 490–491
 Population intercept, 469
 Population mean, 11, 16, 208
 cases not covered by Z or t , 357
 confidence interval, 423
 confidence interval (known standard deviation), 249–253
 difference using independent random samples, 359
 hypothesis tests of, 309, 330
 population mean differences, 359
 templates, 280, 313
 test statistic is t , 309, 355–357
 test statistic is Z , 309, 353–354
 Population parameter, 208–209
 comparison of; *see* Comparison of two populations
 point estimate of, 208
 sample statistics as estimators of, 207–211
 Population proportion, 208
 binomial distribution/normal distribution, 315

Ryan, Patricia A., 554n
 Ryan, T. P., 712n

S

s chart, 707–708
 Sample, 4
 small vs. large samples, 222, 263
 Sample correlation coefficient, 493
 Sample mean, 11, 216, 220
 expected value of, 218
 standard deviation of, 218
 standardized sampling distribution of, 225
 Sample proportion, 209, 225
 Sample-size determination, 275–278, 282
 hypothesis test, 300–301, 334
 manual calculation of, 330–331, 335
 template for, 329
 Sample space, 60, 102
 Sample standard deviation, 17
 Sample statistic, 208
 as estimator of population parameters, 208–211
 Sample variance, 16, 18, 233
 Sampling analysis tool (Excel), 240
 Sampling distribution, 208, 216–227
 defined, 216
 MINITAB for generating, 240–241
 normal sampling distribution, 219–220
 sample proportion and, 225
 template for, 238–239
 Sampling error, 250
 Sampling from the population, 4, 205
 Sampling methods, 211–215
 cluster sampling, 214
 multistage cluster sampling, 214
 other methods, 213–214
 single-stage cluster sampling, 214
 stratified sampling, 213
 systematic sampling, 214
 two-stage cluster sampling, 214
 Sampling with replacement, 130
 Sampling and sampling distributions, 205–241
 central limit theorem, 220–225
 degrees of freedom, 233–236
 estimators and their properties, 229–232
 expected net gain from, 862
 Literary Digest sampling error, 205–208
 nonresponse, 214–215
 obtaining a random sample, 210–211
 as population parameters estimators, 208–211
 small vs. large samples, 222
 standardized sampling distribution of sample mean, 225
 template, 238–241
 uses of, 206
 with/without replacement, 130
 Sarvary, Miklos, 429n, 555n, 637n
 Scales of measurement, 3
 interval scale, 3
 nominal scale, 3
 ordinal scale, 3
 ratio scale, 4
 Scatter plots, 44, 466
 Schank, Thorsten, 505n
 Schatz, Ronald, 668n
 Scheffé method, 427
 Schnabel, Claus, 505n
 Schoar, Antoinette, 597n, 637n
 Schoenfeld, Bruce, 265n
 Schwartz, Nelson D., 321n
 Sciolino, Elaine, 78n
 Seasonal variation, 655
 Seasonality, 654–658
 multiplicative model, 658
 regression model with dummy variables for, 657
 Seber, George A. F., 592n
 Seitz, Thomas, 658n
 Semi-infinite intervals, 172
 Set, 58
 75th percentile, 9
 Shah, Jagar, 269n
 Shakespeare, Catherine, 511n
 Shrum, J. L., 420n
 Sialm, Clemens, 420n
 Sigma squared, 16
 Sign test, 721–742
 possible hypotheses for, 722
 template for, 723
 test statistic, 724
 Significance level, 298–299
 Sikora, Martin, 771n
 Silverman, Rachel Emma, 269n
 Simple exponential smoothing, 668
 Simple index number, 674
 Simple linear regression, 466, 471–472
 analysis-of-variance table, 511–512
 coefficient of determination, 505
 conditional mean of Y , 470
 confidence intervals for regression parameters, 486–488

correlation, 490–494
 curvilinear relationship between Y and X , 471, 514–515
 distributional assumptions of errors, 471
 error variance, 483–488
 estimation: method of least squares, 473–481
 Excel Solver for, 526–529, 532
 F test of, 511–512
 goodness of fit, 505–507
 heteroscedasticity, 513
 how good is the regression, 505–510
 hypothesis tests about, 495–504
 linear relationship between X and Y , 500
 mean square error (MSE), 483–484
 MINITAB for, 533
 missing variables test, 514
 model assumptions, 470
 model building, 467–468
 model inadequacies, 513–517
 model parameters, 469
 normal equations, 475
 normal probability plot, 516–518
 population regression line, 470
 population simple linear regression model, 469
 residual analysis, 513–518
 slope and intercept, 477, 487
 Solver method for, 536–539
 standard error of estimate, 485–488
 steps in, 468
 sum of squares for error (SSE), 473–475, 485
 t test, 500, 512
 template for, 480–481
 use for prediction, 522–526
 Single mode, 821
 Single random sample, 4
 Single-stage cluster sampling, 214
 Single variable, 674
 Skedevold, Gretchen, 653n
 Skewness, 24–25, 37
 Slope, 477, 479, 500
 Smith, Craig S., 367n
 Soliman, Mark T., 511n
 Solver Macro, 281
 multiple regression and, 630–636
 regression, 526–529
 Sorkin, Andrew Ross, 161n
 Spearman rank correlation coefficient, 778–782
 hypothesis test for association, 781
 large-sample test statistic for association, 780
 template for, 781–782
 Spread, 116
 Standard deviation, 16, 116
 defined, 17
 mean and, 26–27
 population standard deviation, 17, 225
 of random variable, 121
 of sample mean, 218
 sample standard deviation, 17
 Standard error, 219, 225
 Standard error of estimate, 485–486
 Standard normal distribution, 172–176
 finding probabilities of, 172–174
 finding values of Z given a probability, 175–176
 importance of, 177
 table area, 172
 Standard normal probabilities (table), 173
 Standard normal random variable Z , 171
 Standard normal test statistic, 744
 State of nature, 826
 Statista, 3
 Statistic, 208
 Statistical analysis, information from, 3
 Statistical control, 696
 Statistical inference, 4–5, 31, 207–208
 business applications of, 5–6
 Statistical model, 424
 checking for inadequacies in, 513
 for control of a variable, 468
 as parsimonious, 468
 for prediction of a variable, 468
 steps in building, 468
 to explain variable relationships, 468
 Statistical process control (SPC), 696
 Statistical test for randomness, 743
 Statistics
 derivation of word, 3
 quality and, 693–701
 as science of inference, 4–5, 31
 use of, 5–6, 167, 205, 248, 292, 344
 Stellin, Susan, 257n
 Stem-and-leaf displays, 34–35
 Stepwise regression, 627–628
 Stigler, S., 690n
 Stokes, Martha, 721n
 Stone, Brad, 66n, 326n
 Story, Louise, 390n
 Straight-line relationship, 466
 Strata, 213
 Stratified sampling, 213
 Studentized range distribution, 424
 Student's distribution/Student's t distribution, 258, 283

- Uncorrelated variables, 500
 - Unexplained deviation (error), 506–508
 - Unexplained variation, 410, 507
 - Uniform distribution, 146–147
 - formulas, 147
 - problem solving with template, 147–148
 - Union, 58–59
 - rule of unions, 75
 - Union rule, 75
 - Universal set, 58
 - Universe, 4
 - Updegrave, Walter, 367n, 756n
 - Upper control limit (UCL), 695, 704, 706–707, 711
 - Upper quartile, 8
 - Useen, Jerry, 51n
 - Utility, 856–859
 - method of assessing, 858
 - Utility function, 856–858, 863
 - Utility scale, 856
- V**
- Value at risk, 151–152
 - Value of information, 860–863
 - Variability; *see* Measures of variability
 - Variable selection methods, 626–628
 - all possible regressions, 626
 - backward elimination, 627–628
 - forward selection, 621–622
 - stepwise regression, 627
 - Variance, 16, 116; *see also* Analysis of variance (ANOVA)
 - defined, 16
 - of discrete random variable, 119–120
 - of linear composite, 123
 - population variance, 16
 - quality control, 696
 - sample variance, 16, 18, 233
 - Variance inflation factor (VIF), 615
 - Variance-stabilizing transformations, 606–607
 - Vella, Matt, 867n
 - Venn diagram, 58–59
 - Vigneron, Olivier, 555n
 - Vining, G. G., 578n
 - Virtual reality, 34
 - Volatility, 19, 780n

W

- Wachter, Jessica A., 535n
- Wagner, Joachim, 505n
- Wain, Daniel, 701n
- Wald-Wolfowitz test, 747–748
- Wallendorf, Melanie, 504n, 510n
- Wang, Jeff, 504n, 510n
- Weak test, 748
- Weighted average, 116
- Weighted least squares (WLS), 570
- Weighting factor, 668
- Weintraub, Arlene, 320n
- Weisberg, Sanford, 590n
- Whiskers, 36
- Wilcoxon rank sum test, 750
- Wilcoxon signed-rank test, 757–762
 - decision rule, 759
 - large-sample version, 759
 - paired-observations two-sample test, 757–758
 - template for, 762
 - test for mean/median of single population, 760–761
- Within-treatment deviation, 408
- Wolff, Edward N., 482n
- Wongsunwai, Wan, 637n
- Wyer, Robrt S., Jr., 420n

X

- $\bar{x}$ -bar chart, 702–705
 - template for, 704–705
- $\bar{x}$ chart, 714
- Xia, Yihong, 510n

Y

- Yates correction, 787

Z

- z distribution, 262
- z standard deviations, 185
- Z test statistic, 309, 353–354
- z value, 185
- Zero skewness, 24
- Zheng, Lu, 421n
- ZTEST function, 338–339

