

Contents

1	Introduction to Statistical Analysis — 1
1.1	Scales of Measurement — 2
1.2	Data, Data Collection, and Data Presentation — 3
1.3	Data Groupings — 5
1.3.1	Categories of Data Groupings — 5
1.3.2	Calculation of Class Interval for Ungrouped Data — 7
1.3.3	Class Limits and Class Boundaries — 8
1.4	Methods of Visualizing Data — 9
1.5	Exploring Data Analysis (EDA) — 12
1.5.1	Stem-and-Leaf Displays — 13
1.5.2	Box Plots — 13
1.5.2.1	Importance of Box Plots — 14
	Exercises for Chapter 1 — 14
2	Introduction to R Software — 16
2.1	How to Download and Install R — 17
2.2	Using R for Descriptive Statistical Analysis and Plots — 17
2.3	Basics of R — 18
2.3.1	R Is Vectorized — 18
2.3.2	R Data Types — 22
2.3.2.1	Scalars — 22
2.3.2.2	Vectors — 23
2.3.2.3	Matrices — 24
2.3.2.4	Data Frames — 25
2.3.2.5	Lists — 26
2.3.2.6	Factors — 27
2.3.2.7	Coercions — 28
2.3.2.8	Date and Time — 28
2.3.3	Missing Values — 30
2.3.4	Data Creation — 31
2.3.5	Data Type Conversion — 37
2.3.6	Variable Information — 38
2.4	Basic Operations in R — 38
2.4.1	Subsetting — 39
2.4.2	Control Structures — 41
2.4.2.1	Conditional — 41
2.4.2.2	For Loop — 41
2.4.2.3	Repeat Loop — 42

2.4.2.4	While Loop —	43
2.4.3	Built-In Functions in R —	44
2.4.3.1	Numerical Functions —	44
2.4.3.2	Character Functions —	45
2.4.3.3	Statistical Probability Functions —	45
2.4.3.4	Other Statistical Functions —	45
2.4.3.5	Other Useful Functions —	45
2.4.4	User-Written Functions —	46
2.4.5	Importing, Reporting, and Writing Data —	49
2.4.5.1	Packages —	49
2.4.5.2	Working Directory and R Script —	49
2.4.5.3	Reading and Writing Local Flat Files —	49
2.4.5.4	Reading and Writing Excel Files —	51
2.4.5.5	Connection Interfaces —	52
2.4.5.6	Connect to a Database —	52
2.5	Data Exploration —	53
2.5.1	Data Exploration through Visualization —	54
2.5.1.1	Bar Chart —	54
2.5.1.2	Pie Chart —	55
2.5.1.3	Box-Plot Distributions —	56
	Exercises for Chapter 2 —	57
3	Descriptive Statistics —	58
3.1	Central Tendency —	58
3.1.1	The Mean —	58
3.1.2	The Median —	59
3.1.3	The Mode —	60
3.2	Measure of Dispersion —	61
3.3	Shapes of the Distribution—Symmetric and Asymmetric —	63
3.3.1	Skewness Illustrated —	64
3.3.2	Kurtosis —	64
	Exercises for Chapter 3 —	66
4	Basic Probability Concepts —	67
4.1	Experiment, Outcome, and Sample Space —	67
4.1.1	Experiment —	67
4.1.2	Outcome —	67
4.1.3	Sample Space —	67
4.2	Elementary Events —	68
4.3	Complementary Events —	68

4.4	Mutually Exclusive Events — 68
4.5	Mutually Inclusive Events — 69
4.6	Venn Diagrams — 69
4.7	Probability — 70
4.7.1	Simulation of a Random Sample in R — 71
4.8	Axioms of Probability — 73
4.9	Basic Properties of Probability — 74
4.10	Independent Events and Dependent Events — 74
4.11	Multiplication Rule of Probability — 76
4.12	Conditional Probabilities — 76
4.13	Computation of Probability in R — 77
	Exercises for Chapter 4 — 80
5	Discrete Probability Distributions — 82
5.1	Probability Mass Distribution — 82
5.2	Expected Value and Variance of a Discrete Random Variable — 83
5.3	Binomial Probability Distribution — 86
5.4	Expected Value and Variance of a Binomial Distribution — 87
5.4.1	Expected Value (Mean) of a Binomial Distribution — 88
5.4.2	Variance of a Binomial Distribution — 89
5.5	Solve Problems Involving Binomial Distribution Using R — 92
5.5.1	Generating a Random Sample from a Binomial Distribution — 92
5.5.2	Calculate Probability from a Binomial Distribution — 94
	Exercises for Chapter 5 — 96
6	Continuous Probability Distributions — 97
6.1	Normal Distribution and Standardized Normal Distribution — 98
6.1.1	Properties of a Normal Distribution — 99
6.2	Standard Normal Score (Z-Score) — 100
6.3	Approximate Normal Distribution to the Binomial — 103
6.4	Use of the Normal Distribution in Business Problem Solving Using R — 104
	Exercises for Chapter 6 — 105
7	Other Continuous Probability Distributions — 107
7.1	Student's t-Distribution — 107
7.1.1	Properties of the Student's t-Distribution — 110

7.1.2	R Code to Obtain t-values Using Example 7.1 —	111
7.2	Chi-square Distribution —	112
7.2.1	Properties of the Chi-square Distribution —	115
7.2.2	<i>Uses of the Chi-square Distribution</i> —	116
7.3	F-Distribution —	118
7.3.1	Properties of the F-Distribution —	118
7.3.2	Uses of the F-Distribution —	119
	Exercises for Chapter 7 —	121
8	Sampling and Sampling Distribution —	122
8.1	Probability and Non-Probability Sampling —	122
8.2	Probability Sampling Techniques—Simple Random, Systematic, Cluster, and Stratified —	123
8.2.1	Simple Random Sampling (SRS) —	123
8.2.2	Systematic Sampling —	127
8.2.2.1	Advantages of Systematic Sampling —	127
8.2.2.2	Disadvantages of Systematic Sampling —	127
8.2.3	Cluster Sampling —	127
8.2.3.1	Advantages of Cluster Sampling —	128
8.2.3.2	Disadvantages of Cluster Sampling —	128
8.2.4	Stratified Random Sampling —	128
8.2.4.1	Advantages of Stratified Random Sampling —	129
8.2.4.2	Disadvantages of Stratified Random Sampling —	129
8.3	Non-Probability Sampling Techniques—Judgment, Quota, and Convenience —	129
8.3.1	Judgment Sampling —	129
8.3.1.1	Advantages of Judgment Sampling —	130
8.3.1.2	Disadvantages of Judgment Sampling —	130
8.3.2	Quota Sampling —	130
8.3.2.1	Advantages of Quota Sampling —	130
8.3.2.2	Disadvantages of Quota Sampling —	130
8.3.3	Convenience Sampling —	131
8.3.3.1	Advantages of Convenience Sampling —	131
8.3.3.2	Disadvantages of Convenience Sampling —	131
8.4	Sampling Distribution of the Mean —	131
8.5	Central Limit Theorem and Its Significance —	134
	Exercises for Chapter 8 —	137

9	Confidence Intervals for Single Population Mean and Proportion — 138
9.1	Point Estimates and Interval Estimates — 138
9.2	Confidence Intervals for Mean — 139
9.3	Confidence Intervals for Proportion — 142
9.4	Calculating the Sample Size — 143
9.5	Factors That Determine Margin of Error — 145
	Exercises for Chapter 9 — 146
10	Hypothesis Testing for Single Population Mean and Proportion — 147
10.1	Null and Alternative Hypotheses — 147
10.2	Type I and Type II Error — 147
10.3	Acceptance and Rejection Regions — 148
10.4	Hypothesis Testing Procedures — 148
	Exercises for Chapter 10 — 154
11	Regression Analysis and Correlation Analysis — 157
11.1	Construction of Line Fit Plots — 157
11.2	Types of Regression Analysis — 157
11.2.1	Uses of Regression Analysis — 158
11.3	Simple Linear Regression — 158
11.3.1	Assumptions of Simple Linear Regression — 160
11.4	Multiple Linear Regression — 164
11.4.1	Significance Testing of Each Variable — 168
11.4.2	Interpretation of Regression Coefficients and Other Output — 168
11.5	Pearson Correlation Coefficient — 171
11.5.1	Assumptions of Correlation Test — 172
11.5.2	Types of Correlation — 172
11.5.3	Coefficient of Determination — 172
11.5.4	Test for the Significance of Correlation Coefficient (r) — 173
	Exercises for Chapter 11 — 175
12	Poisson Distribution — 179
12.1	Poisson Distribution and Its Properties — 179
12.2	Mean and Variance of a Poisson Distribution — 180
12.2.1	Mean — 181
12.2.2	Variance — 182
12.3	Application of the Poisson Distribution — 187

12.4	Poisson to Approximate the Binomial — 188
	Exercises for Chapter 12 — 191
13	Uniform Distributions — 193
13.1	Uniform Distribution and Its Properties — 193
13.2	Mean and Variance of a Uniform Distribution — 193
13.2.1	Mean — 194
13.2.2	Variance — 194
	Exercises for Chapter 13 — 200
14	Time Series Analysis — 202
14.1	Concept of Time Series Data — 202
14.1.1	Uses and Application of Time Series Analysis — 202
14.1.1.1	Time Domain Approach — 203
14.1.1.2	Frequency Domain Approach — 204
14.2	Univariate Time Series Model — 204
14.2.1	Generating a Time-Series Object in R — 207
14.2.2	Smoothing and Seasonal Decomposition — 208
14.2.2.1	Smoothing with Simple Moving Averages — 208
14.2.2.2	Seasonal Decomposition — 211
14.2.2.3	Exponential Forecasting Models — 217
14.2.2.4	Simple Exponential Smoothing — 218
14.2.2.5	Holt and Holt-Winters Exponential Smoothing — 222
14.2.2.6	The ets() Function and Automated Forecasting — 224
14.2.2.7	Autoregressive Integrated Moving Average (ARIMA) Forecasting Models — 226
14.3	Multivariate Time Series Models — 230
14.3.1	ARMA and ARIMA Models — 230
14.3.2	Automated ARIMA Predictions — 237
	Exercises for Chapter 14 — 238
Appendix — 240	
Index — 261	