
Contents

List of Figures	xi
List of Tables	xiii
Preface	xv
About the Author	xxi
1 Introduction to R	1
1.1 What is <i>R</i> ?	1
1.2 Steps on Installing <i>R</i> and Updating <i>R</i> Packages	3
1.2.1 First Step: Install <i>R Base System</i>	3
1.2.2 Second Step: Installing and Updating <i>R</i> Packages	4
1.2.3 Steps to Get Help and Documentation	5
1.3 Database Management and Data Manipulations	7
1.3.1 Data Management with <i>RMySQL</i>	7
1.3.2 Data Management with Microsoft Excel and <i>R</i> Package <i>gdata</i>	7
1.3.3 Data Management with Microsoft Excel and <i>R</i> Package <i>xlsx</i>	8
1.3.4 Data Management with Microsoft Excel and <i>R</i> Package <i>readxl</i>	8
1.3.5 Default Methods to Read Data into <i>R</i>	8
1.3.6 <i>R</i> Package <i>foreign</i>	9
1.4 Monte-Carlo Simulation Illustration	10
1.4.1 Central Limit Theorem	10
1.4.2 <i>R</i> Capacity on Random Number Generation	12
1.4.3 <i>R</i> to Prove (Empirically) the Central Limit Theorem	12
1.5 Summary and Recommendations for Further Reading	19
1.6 Exercises	20
2 Linear Regression	21
2.1 Descriptive Data Analysis	21
2.2 Review for <i>Multiple Linear Regression</i> Models	24
2.2.1 Multiple Linear Regression Model	24

2.2.2	The Method of Least Squares Estimation (LSE)	25
2.2.3	The Properties of LSE	26
2.2.4	The Fundamental Assumptions	28
2.3	Simple Linear Regression: One Independent Variable	29
2.3.1	Model Fitting	29
2.3.2	Model Diagnostics	31
2.4	Multiple Linear Regression	32
2.4.1	Model Fitting	33
2.4.2	Regression Model Diagnostics: Checking Assumptions	35
2.5	Model Diagnostics with Monte-Carlo Simulations	40
2.5.1	Violation of <i>Normality of Residuals</i>	40
2.5.2	Violation of <i>Homoscedasticity</i>	49
2.6	Discussions	57
2.7	Exercises	57
3	Transition from Linear to Nonlinear Regression	59
3.1	The Financial Compounding Model	59
3.1.1	The Basic Concept of Compounding	60
3.1.2	Data Available	61
3.1.3	Computation and <i>R</i> Implementation	63
3.2	Depreciation of Non-Current Assets	65
3.2.1	The Reduced Balance Method	65
3.2.2	Data Available	66
3.2.3	Computation and <i>R</i> Implementation	67
3.3	Pros and Cons of Linearization	70
3.3.1	Theoretical Differences in Residual Patterns	71
3.3.2	Numerical Differences in Residual Patterns	72
3.3.3	Additive Errors and Multiplicative Errors	77
3.4	Discussions	80
3.5	Exercises	80
4	Nonlinear Regression Modeling	81
4.1	Logistic Growth Model	81
4.2	Theory of Nonlinear Regression	83
4.2.1	<i>R</i> Packages for Nonlinear Regression	83
4.2.2	<i>R</i> Function <i>nls</i> for Nonlinear Regression	84
4.2.3	Theory of Nonlinear Regression in <i>nls</i>	85
4.3	Nonlinear Regression for Mobile Banking	86
4.3.1	Data on Customer Size on Mobile Banking	86
4.3.2	Fitting the Mobile Banking Data Using <i>R</i> Function <i>nls</i>	88
4.3.3	Possible Consequences Shown in <i>dMobile</i> Data	92
4.4	Monte-Carlo Simulation-Based Investigation	93
4.4.1	Simulation Setup	93
4.4.2	Numerical Investigation with Monte-Carlo Simulations	94

4.5	Discussions	98
4.6	Exercises	99
5	The Logistic Regression	101
5.1	The Logistic Regression	101
5.1.1	Common Misconceptions	101
5.1.2	Logistic Regression with MLE	103
5.2	S&P Stock Market Data Analysis	105
5.2.1	Descriptive Data Analysis	105
5.2.2	Model Fitting with Logistic Regression	108
5.2.3	Prediction using Logistic Regression	111
5.3	Monte-Carlo Simulation-Based Validation Study	112
5.3.1	Simulation Design	113
5.3.2	Simulation Implementation in <i>R</i>	114
5.3.3	A Full-Scale Monte-Carlo Simulation Validation.	118
5.4	Discussions	121
5.5	Exercises	122
6	The Poisson Regression: Models for Count Data	124
6.1	Why Poisson Regression to Model Count Data	125
6.1.1	Data Description	125
6.1.2	Why Not Linear Regression	127
6.1.3	Why Not the Log-Linear Regression	129
6.2	Poisson Regression	134
6.2.1	Conventional Poisson Regression	134
6.2.2	Models for Overdispersed Count Data	136
6.2.3	Models for Underdispersed Count Data	141
6.3	Monte-Carlo Simulation Study	143
6.3.1	Scenarion 1: No Dispersion	144
6.3.2	Scenario 2: Overdispersion	147
6.3.3	Scenario 3: Underdispersion	151
6.4	Discussions	155
6.5	Exercises	156
7	Autoregressive Integrated Moving-Average Models	157
7.1	Autocorrelation Function and Partial Autocorrelation Function	158
7.1.1	Autocorrelation Function	158
7.1.2	Partial Autocorrelation Function	158
7.1.3	Ljung-Box Test for Significance	159
7.1.4	Illustration with <i>Wages</i> Data	160
7.2	Auto-Regressive Integrated Moving-Average Time Series	161
7.2.1	Auto-Regressive Process with Order p	161
7.2.2	Moving-Average Process with Order q	162
7.2.3	ARMA Process with Order of p and q	163

7.2.4	Auto-Regressive Integrated Moving-Average Process	163
7.2.5	ARIMA with Regression	164
7.3	Fitting <i>ARIMA</i> using <i>R</i>	165
7.3.1	Fitting AR(p) Model	165
7.3.2	Fitting MA(q) Models	168
7.3.3	Fitting ARIMA model	169
7.3.4	Fitting Time Series Regression	171
7.4	Monte-Carlo Simulation Validation	173
7.4.1	Simulation Design	174
7.4.2	Simulation Implementation	175
7.4.3	Summary	180
7.5	Discussions	180
7.5.1	Stationary Process	181
7.5.2	White Noise Process is a Stationary Process	182
7.5.3	Random Walk Process is a Non-Stationary Process	182
7.5.4	Stationary to Heteroskedasticity	182
7.6	Exercises	183
8	Generalized AutoRegressive Conditional Heteroskedasticity Model	185
8.1	The Volatility Index Data	186
8.1.1	The Volatility Index Briefly Explained	186
8.1.2	The VIX Data	187
8.1.3	Descriptive Analysis	188
8.1.4	ARIMA Modeling	190
8.2	The GARCH Modeling	193
8.2.1	From General Regression to <i>ARCH</i> Models	193
8.2.2	From <i>ARCH</i> to <i>GARCH</i>	195
8.2.3	R Packages	196
8.3	Data Analysis with GARCH Modeling	197
8.3.1	Model Fitting	198
8.3.2	Summary of Model Fitting	214
8.4	Discussions and Recommendations	218
8.4.1	Discussions	218
8.4.2	Recommendations	219
8.5	Exercises	221
9	Cointegration	222
9.1	What R Packages for <i>Cointegration</i>	222
9.2	Statistical Methods	223
9.2.1	The Augmented Dickey-Fuller Test for Stationarity	223
9.2.2	Johansen Test Procedure	224
9.3	Standard & Poor Data	226
9.3.1	Descriptive Analysis	226
9.3.2	Test for Non-Stationarity	227

<i>Contents</i>	ix
9.3.3 Cointegration Regression	230
9.3.4 Vector Error Correction Model	233
9.4 The Volatility Index Data	238
9.4.1 Descriptive Analysis	238
9.4.2 Test for Non-Stationarity	240
9.4.3 Vector Error Correction Model	244
9.5 Discussions	249
9.5.1 How to Do Cointegration in R	250
9.5.2 Cointegration in Financial Analysis	251
9.5.3 Cointegration in Investment among Stocks to Mutual Funds	252
10 Financial Statistical Modeling in Risk and Wealth Management	254
10.1 Statistical Modeling in Risk Management	254
10.1.1 Understanding Risk	254
10.1.2 Key Financial Statistics	255
10.1.3 Statistical Tools and Models	258
10.2 Statistical Modeling in Wealth Management	259
10.2.1 The Role of Financial Statistical Modeling	259
10.2.2 Key Statistical Models in Wealth Management	260
10.2.3 Portfolio Construction and Optimization	261
10.3 Discussions and Recommendations	262
Bibliography	265
Index	269