Contents

Preface **vii**

1 The Nature of Financial Data **1**
1.1 Financial Time Series 5
1.1.1 Autocorrelations 7
1.1.2 Stationarity 7
1.1.3 Time Scales and Data Aggregation 8
1.2 Financial Assets and Markets 12
1.2.1 Markets and Regulatory Agencies 16
1.2.2 Interest 20
1.2.3 Returns on Assets 29
1.2.4 Stock Prices; Fair Market Value 32
1.2.5 Splits, Dividends, and Return of Capital 45
1.2.6 Indexes and "the Market" 48
1.2.7 Derivative Assets 63
1.2.8 Short Positions 65
1.2.9 Portfolios of Assets: Diversification and Hedging . . . 66
1.3 Frequency Distributions of Returns 76
1.3.1 Location and Scale 79
1.3.2 Skewness 81
1.3.3 Kurtosis 81
1.3.4 Multivariate Data 82
1.3.5 The Normal Distribution 87
1.3.6 Q-Q Plots 91
1.3.7 Outliers 93
1.3.8 Other Statistical Measures 94
1.4 Volatility 98
1.4.1 The Time Series of Returns 99
1.4.2 Measuring Volatility: Historical and Implied 102
1.4.3 Volatility Indexes: The VIX 108
1.4.4 The Curve of Implied Volatility 112
1.4.5 Risk Assessment and Management 113
1.5 Market Dynamics 120
1.6 Stylized Facts about Financial Data 129
Notes and Further Reading 130
Exercises and Questions for Review 132

Appendix A1: Accessing and Analyzing Financial Data in R . . . 139
A1.1 R Basics . . . 140
A1.2 Data Repositories and Inputting Data into R . . . 158
A1.3 Time Series and Financial Data in R . . . 172
A1.4 Data Cleansing . . . 183
Notes, Comments, and Further Reading on R . . . 187
Exercises in R . . . 191

2 Exploratory Financial Data Analysis 205
2.1 Data Reduction . . . 207
2.1.1 Simple Summary Statistics . . . 207
2.1.2 Centering and Standardizing Data . . . 208
2.1.3 Simple Summary Statistics for Multivariate Data . . . 208
2.1.4 Transformations . . . 209
2.1.5 Identifying Outlying Observations . . . 210
2.2 The Empirical Cumulative Distribution Function . . . 211
2.3 Nonparametric Probability Density Estimation . . . 217
2.3.1 Binned Data . . . 217
2.3.2 Kernel Density Estimator . . . 219
2.3.3 Multivariate Kernel Density Estimator . . . 220
2.4 Graphical Methods in Exploratory Analysis . . . 221
2.4.1 Time Series Plots . . . 222
2.4.2 Histograms . . . 222
2.4.3 Boxplots . . . 224
2.4.4 Density Plots . . . 226
2.4.5 Bivariate Data . . . 227
2.4.6 Q-Q Plots . . . 229
2.4.7 Graphics in R . . . 234
Notes and Further Reading . . . 238
Exercises . . . 239

3 Probability Distributions in Models of Observable Events 245
3.1 Random Variables and Probability Distributions . . . 247
3.1.1 Discrete Random Variables . . . 248
3.1.2 Continuous Random Variables . . . 252
3.1.3 Linear Combinations of Random Variables; Expectations and Quantiles . . . 256
3.1.4 Survival and Hazard Functions . . . 257
3.1.5 Multivariate Distributions . . . 258
3.1.6 Measures of Association in Multivariate Distributions 261
3.1.7 Copulas . . . 264
3.1.8 Transformations of Multivariate Random Variables . . 267
3.1.9 Distributions of Order Statistics . . . 269
3.1.10 Asymptotic Distributions; The Central Limit Theorem 270
3.1.11 The Tails of Probability Distributions . . . 273

3.1.12 Sequences of Random Variables; Stochastic Processes 278
3.1.13 Diffusion of Stock Prices and Pricing of Options . . . 279
3.2 Some Useful Probability Distributions 282
3.2.1 Discrete Distributions 283
3.2.2 Continuous Distributions 285
3.2.3 Multivariate Distributions 294
3.2.4 General Families of Distributions Useful in Modeling . 295
3.2.5 Constructing Multivariate Distributions 309
3.2.6 Modeling of Data-Generating Processes 311
3.2.7 R Functions for Probability Distributions 311
3.3 Simulating Observations of a Random Variable 314
3.3.1 Uniform Random Numbers 315
3.3.2 Generating Nonuniform Random Numbers 317
3.3.3 Simulating Data in R 321
Notes and Further Reading 323
Exercises 325

4 Statistical Models and Methods of Inference 335
4.1 Models 336
4.1.1 Fitting Statistical Models 340
4.1.2 Measuring and Partitioning Observed Variation 340
4.1.3 Linear Models 343
4.1.4 Nonlinear Variance-Stabilizing Transformations 344
4.1.5 Parametric and Nonparametric Models 345
4.1.6 Bayesian Models 346
4.1.7 Models for Time Series 346
4.2 Criteria and Methods for Statistical Modeling 347
4.2.1 Estimators and Their Properties 347
4.2.2 Methods of Statistical Modeling 349
4.3 Optimization in Statistical Modeling; Least Squares and Maximum Likelihood 358
4.3.1 The General Optimization Problem 358
4.3.2 Least Squares 363
4.3.3 Maximum Likelihood 371
4.3.4 R Functions for Optimization 375
4.4 Statistical Inference 376
4.4.1 Confidence Intervals 379
4.4.2 Testing Statistical Hypotheses 381
4.4.3 Prediction 385
4.4.4 Inference in Bayesian Models 386
4.4.5 Resampling Methods; The Bootstrap 393
4.4.6 Robust Statistical Methods 396
4.4.7 Estimation of the Tail Index 399
4.4.8 Estimation of VaR and Expected Shortfall 404
4.5 Models of Relationships among Variables 408

4.5.1 Principal Components 409
4.5.2 Regression Models 413
4.5.3 Linear Regression Models 418
4.5.4 Linear Regression Models: The Regressors 422
4.5.5 Linear Regression Models: Individual Observations and Residuals 428
4.5.6 Linear Regression Models: An Example 435
4.5.7 Nonlinear Models 449
4.5.8 Specifying Models in R 454
4.6 Assessing the Adequacy of Models 455
4.6.1 Goodness-of-Fit Tests; Tests for Normality 456
4.6.2 Cross-Validation 463
4.6.3 Model Selection and Model Complexity 467
Notes and Further Reading 469
Exercises 472

5 Discrete Time Series Models and Analysis 487
5.1 Basic Linear Operations 495
5.1.1 The Backshift Operator 495
5.1.2 The Difference Operator 497
5.1.3 The Integration Operator 500
5.1.4 Summation of an Infinite Geometric Series 500
5.1.5 Linear Difference Equations 501
5.1.6 Trends and Detrending 505
5.1.7 Cycles and Seasonal Adjustment 508
5.2 Analysis of Discrete Time Series Models 510
5.2.1 Stationarity 514
5.2.2 Sample Autocovariance and Autocorrelation Functions; Stationarity and Estimation 518
5.2.3 Statistical Inference in Stationary Time Series 523
5.3 Autoregressive and Moving Average Models 528
5.3.1 Moving Average Models; MA(q) 529
5.3.2 Autoregressive Models; AR(p) 534
5.3.3 The Partial Autocorrelation Function, PACF 547
5.3.4 ARMA and ARIMA Models 549
5.3.5 Simulation of ARMA and ARIMA Models 555
5.3.6 Statistical Inference in ARMA and ARIMA Models . 556
5.3.7 Selection of Orders in ARIMA Models 560
5.3.8 Forecasting in ARIMA Models 561
5.3.9 Analysis of ARMA and ARIMA Models in R 561
5.3.10 Robustness of ARMA Procedures; Innovations with Heavy Tails 566
5.3.11 Financial Data 568
5.3.12 Linear Regression with ARMA Errors 571
5.4 Conditional Heteroscedasticity 575

5.4.1 ARCH Models 576
5.4.2 GARCH Models and Extensions 580
5.5 Unit Roots and Cointegration 584
5.5.1 Spurious Correlations; The Distribution of the Correlation Coefficient 584
5.5.2 Unit Roots 592
5.5.3 Cointegrated Processes 599
Notes and Further Reading 603
Exercises 604

References **615**

Index **623**