

Analysis of Financial Time Series

Second Edition

RUEY S. TSAY

University of Chicago
Graduate School of Business



A JOHN WILEY & SONS, INC., PUBLICATION

Analysis of Financial Time Series

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

Editors: *David J. Balding, Noel A. C. Cressie, Nicholas I. Fisher,
Iain M. Johnstone, J. B. Kadane, Geert Molenberghs, Louise M. Ryan,
David W. Scott, Adrian F. M. Smith, Jozef L. Teugels*

Editors Emeriti: *Vic Barnett, J. Stuart Hunter, David G. Kendall*

A complete list of the titles in this series appears at the end of this volume.

Analysis of Financial Time Series

Second Edition

RUEY S. TSAY

University of Chicago
Graduate School of Business



A JOHN WILEY & SONS, INC., PUBLICATION

Copyright © 2005 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Tsay, Ruey S., 1951–

Analysis of financial time series/Ruey S. Tsay.—2nd ed.

p. cm.

“Wiley-Interscience.”

Includes bibliographical references and index.

ISBN-13 978-0-471-69074-0

ISBN-10 0-471-69074-0 (cloth)

1. Time-series analysis. 2. Econometrics. 3. Risk management. I. Title.

HA30.3T76 2005

332'.01'51955—dc22

2005047030

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To my parents and Teresa

Contents

Preface	xvii
Preface to First Edition	xix
1. Financial Time Series and Their Characteristics	1
1.1 Asset Returns, 2	
1.2 Distributional Properties of Returns, 7	
1.2.1 Review of Statistical Distributions and Their Moments, 7	
1.2.2 Distributions of Returns, 13	
1.2.3 Multivariate Returns, 16	
1.2.4 Likelihood Function of Returns, 17	
1.2.5 Empirical Properties of Returns, 17	
1.3 Processes Considered, 20	
Exercises, 22	
References, 23	
2. Linear Time Series Analysis and Its Applications	24
2.1 Stationarity, 25	
2.2 Correlation and Autocorrelation Function, 25	
2.3 White Noise and Linear Time Series, 31	
2.4 Simple Autoregressive Models, 32	
2.4.1 Properties of AR Models, 33	
2.4.2 Identifying AR Models in Practice, 40	
2.4.3 Goodness of Fit, 46	
2.4.4 Forecasting, 47	

2.5	Simple Moving-Average Models, 50
2.5.1	Properties of MA Models, 51
2.5.2	Identifying MA Order, 52
2.5.3	Estimation, 53
2.5.4	Forecasting Using MA Models, 54
2.6	Simple ARMA Models, 56
2.6.1	Properties of ARMA(1,1) Models, 57
2.6.2	General ARMA Models, 58
2.6.3	Identifying ARMA Models, 59
2.6.4	Forecasting Using an ARMA Model, 61
2.6.5	Three Model Representations for an ARMA Model, 62
2.7	Unit-Root Nonstationarity, 64
2.7.1	Random Walk, 64
2.7.2	Random Walk with Drift, 65
2.7.3	Trend-Stationary Time Series, 67
2.7.4	General Unit-Root Nonstationary Models, 67
2.7.5	Unit-Root Test, 68
2.8	Seasonal Models, 72
2.8.1	Seasonal Differencing, 73
2.8.2	Multiplicative Seasonal Models, 75
2.9	Regression Models with Time Series Errors, 80
2.10	Consistent Covariance Matrix Estimation, 86
2.11	Long-Memory Models, 89
	Appendix: Some SCA Commands, 91
	Exercises, 93
	References, 96

3.	Conditional Heteroscedastic Models	97
3.1	Characteristics of Volatility, 98	
3.2	Structure of a Model, 99	
3.3	Model Building, 101	
3.3.1	Testing for ARCH Effect, 101	
3.4	The ARCH Model, 102	
3.4.1	Properties of ARCH Models, 104	
3.4.2	Weaknesses of ARCH Models, 106	
3.4.3	Building an ARCH Model, 106	
3.4.4	Some Examples, 109	
3.5	The GARCH Model, 113	
3.5.1	An Illustrative Example, 116	

3.5.2	Forecasting Evaluation, 121
3.5.3	A Two-Pass Estimation Method, 121
3.6	The Integrated GARCH Model, 122
3.7	The GARCH-M Model, 123
3.8	The Exponential GARCH Model, 124
3.8.1	An Alternative Model Form, 125
3.8.2	An Illustrative Example, 126
3.8.3	Second Example, 126
3.8.4	Forecasting Using an EGARCH Model, 128
3.9	The Threshold GARCH Model, 130
3.10	The CHARMA Model, 131
3.10.1	Effects of Explanatory Variables, 133
3.11	Random Coefficient Autoregressive Models, 133
3.12	The Stochastic Volatility Model, 134
3.13	The Long-Memory Stochastic Volatility Model, 134
3.14	Application, 136
3.15	Alternative Approaches, 140
3.15.1	Use of High-Frequency Data, 140
3.15.2	Use of Daily Open, High, Low, and Close Prices, 143
3.16	Kurtosis of GARCH Models, 145
Appendix: Some RATS Programs for Estimating Volatility Models, 147	
Exercises, 148	
References, 151	

4. Nonlinear Models and Their Applications

154

4.1	Nonlinear Models, 156
4.1.1	Bilinear Model, 156
4.1.2	Threshold Autoregressive (TAR) Model, 157
4.1.3	Smooth Transition AR (STAR) Model, 163
4.1.4	Markov Switching Model, 164
4.1.5	Nonparametric Methods, 167
4.1.6	Functional Coefficient AR Model, 175
4.1.7	Nonlinear Additive AR Model, 176
4.1.8	Nonlinear State-Space Model, 176
4.1.9	Neural Networks, 177
4.2	Nonlinearity Tests, 183
4.2.1	Nonparametric Tests, 183
4.2.2	Parametric Tests, 186
4.2.3	Applications, 190

4.3	Modeling, 191	
4.4	Forecasting, 192	
4.4.1	Parametric Bootstrap, 192	
4.4.2	Forecasting Evaluation, 192	
4.5	Application, 194	
	Appendix A: Some RATS Programs for Nonlinear Volatility Models, 199	
	Appendix B: S-Plus Commands for Neural Network, 200	
	Exercises, 200	
	References, 202	
5.	High-Frequency Data Analysis and Market Microstructure	206
5.1	Nonsynchronous Trading, 207	
5.2	Bid–Ask Spread, 210	
5.3	Empirical Characteristics of Transactions Data, 212	
5.4	Models for Price Changes, 218	
5.4.1	Ordered Probit Model, 218	
5.4.2	A Decomposition Model, 221	
5.5	Duration Models, 225	
5.5.1	The ACD Model, 227	
5.5.2	Simulation, 229	
5.5.3	Estimation, 232	
5.6	Nonlinear Duration Models, 236	
5.7	Bivariate Models for Price Change and Duration, 237	
	Appendix A: Review of Some Probability Distributions, 242	
	Appendix B: Hazard Function, 245	
	Appendix C: Some RATS Programs for Duration Models, 246	
	Exercises, 248	
	References, 250	
6.	Continuous-Time Models and Their Applications	251
6.1	Options, 252	
6.2	Some Continuous-Time Stochastic Processes, 252	
6.2.1	The Wiener Process, 253	
6.2.2	Generalized Wiener Processes, 255	
6.2.3	Ito Processes, 256	
6.3	Ito's Lemma, 256	
6.3.1	Review of Differentiation, 256	
6.3.2	Stochastic Differentiation, 257	

6.3.3	An Application, 258	
6.3.4	Estimation of μ and σ , 259	
6.4	Distributions of Stock Prices and Log Returns, 261	
6.5	Derivation of Black–Scholes Differential Equation, 262	
6.6	Black–Scholes Pricing Formulas, 264	
6.6.1	Risk-Neutral World, 264	
6.6.2	Formulas, 264	
6.6.3	Lower Bounds of European Options, 267	
6.6.4	Discussion, 268	
6.7	An Extension of Ito’s Lemma, 272	
6.8	Stochastic Integral, 273	
6.9	Jump Diffusion Models, 274	
6.9.1	Option Pricing Under Jump Diffusion, 279	
6.10	Estimation of Continuous-Time Models, 282	
	Appendix A: Integration of Black–Scholes Formula, 282	
	Appendix B: Approximation to Standard Normal Probability, 284	
	Exercises, 284	
	References, 285	
7.	Extreme Values, Quantile Estimation, and Value at Risk	287
7.1	Value at Risk, 287	
7.2	RiskMetrics, 290	
7.2.1	Discussion, 293	
7.2.2	Multiple Positions, 293	
7.3	An Econometric Approach to VaR Calculation, 294	
7.3.1	Multiple Periods, 296	
7.4	Quantile Estimation, 298	
7.4.1	Quantile and Order Statistics, 299	
7.4.2	Quantile Regression, 300	
7.5	Extreme Value Theory, 301	
7.5.1	Review of Extreme Value Theory, 301	
7.5.2	Empirical Estimation, 304	
7.5.3	Application to Stock Returns, 307	
7.6	Extreme Value Approach to VaR, 311	
7.6.1	Discussion, 314	
7.6.2	Multiperiod VaR, 316	
7.6.3	VaR for a Short Position, 316	
7.6.4	Return Level, 317	

7.7	A New Approach Based on the Extreme Value Theory, 318
7.7.1	Statistical Theory, 318
7.7.2	Mean Excess Function, 320
7.7.3	A New Approach to Modeling Extreme Values, 322
7.7.4	VaR Calculation Based on the New Approach, 324
7.7.5	An Alternative Parameterization, 325
7.7.6	Use of Explanatory Variables, 328
7.7.7	Model Checking, 329
7.7.8	An Illustration, 330
	Exercises, 335
	References, 337

8. Multivariate Time Series Analysis and Its Applications 339

8.1	Weak Stationarity and Cross-Correlation Matrices, 340
8.1.1	Cross-Correlation Matrices, 340
8.1.2	Linear Dependence, 341
8.1.3	Sample Cross-Correlation Matrices, 342
8.1.4	Multivariate Portmanteau Tests, 346
8.2	Vector Autoregressive Models, 349
8.2.1	Reduced and Structural Forms, 349
8.2.2	Stationarity Condition and Moments of a VAR(1) Model, 351
8.2.3	Vector AR(p) Models, 353
8.2.4	Building a VAR(p) Model, 354
8.2.5	Impulse Response Function, 362
8.3	Vector Moving-Average Models, 365
8.4	Vector ARMA Models, 371
8.4.1	Marginal Models of Components, 375
8.5	Unit-Root Nonstationarity and Cointegration, 376
8.5.1	An Error-Correction Form, 379
8.6	Cointegrated VAR Models, 380
8.6.1	Specification of the Deterministic Function, 382
8.6.2	Maximum Likelihood Estimation, 383
8.6.3	A Cointegration Test, 384
8.6.4	Forecasting of Cointegrated VAR Models, 385
8.6.5	An Example, 385
8.7	Threshold Cointegration and Arbitrage, 390
8.7.1	Multivariate Threshold Model, 391
8.7.2	The Data, 392

8.7.3	Estimation, 393	
	Appendix A: Review of Vectors and Matrices, 395	
	Appendix B: Multivariate Normal Distributions, 399	
	Appendix C: Some SCA Commands, 400	
	Exercises, 401	
	References, 402	
9.	Principal Component Analysis and Factor Models	405
9.1	A Factor Model, 406	
9.2	Macroeconometric Factor Models, 407	
9.2.1	A Single-Factor Model, 408	
9.2.2	Multifactor Models, 412	
9.3	Fundamental Factor Models, 414	
9.3.1	BARRA Factor Model, 414	
9.3.2	Fama–French Approach, 420	
9.4	Principal Component Analysis, 421	
9.4.1	Theory of PCA, 421	
9.4.2	Empirical PCA, 422	
9.5	Statistical Factor Analysis, 426	
9.5.1	Estimation, 428	
9.5.2	Factor Rotation, 429	
9.5.3	Applications, 430	
9.6	Asymptotic Principal Component Analysis, 436	
9.6.1	Selecting the Number of Factors, 437	
9.6.2	An Example, 437	
	Exercises, 440	
	References, 441	
10.	Multivariate Volatility Models and Their Applications	443
10.1	Exponentially Weighted Estimate, 444	
10.2	Some Multivariate GARCH Models, 447	
10.2.1	Diagonal VEC Model, 447	
10.2.2	BEKK Model, 451	
10.3	Reparameterization, 454	
10.3.1	Use of Correlations, 454	
10.3.2	Cholesky Decomposition, 455	
10.4	GARCH Models for Bivariate Returns, 459	
10.4.1	Constant-Correlation Models, 459	
10.4.2	Time-Varying Correlation Models, 464	

- 10.4.3 Some Recent Developments, 470
- 10.5 Higher Dimensional Volatility Models, 471
- 10.6 Factor–Volatility Models, 477
- 10.7 Application, 480
- 10.8 Multivariate t Distribution, 482
- Appendix: Some Remarks on Estimation, 483
- Exercises, 488
- References, 489

11. State-Space Models and Kalman Filter 490

- 11.1 Local Trend Model, 490
 - 11.1.1 Statistical Inference, 493
 - 11.1.2 Kalman Filter, 495
 - 11.1.3 Properties of Forecast Error, 496
 - 11.1.4 State Smoothing, 498
 - 11.1.5 Missing Values, 501
 - 11.1.6 Effect of Initialization, 503
 - 11.1.7 Estimation, 504
 - 11.1.8 S-Plus Commands Used, 505
- 11.2 Linear State-Space Models, 508
- 11.3 Model Transformation, 509
 - 11.3.1 CAPM with Time-Varying Coefficients, 510
 - 11.3.2 ARMA Models, 512
 - 11.3.3 Linear Regression Model, 518
 - 11.3.4 Linear Regression Models with ARMA Errors, 519
 - 11.3.5 Scalar Unobserved Component Model, 521
- 11.4 Kalman Filter and Smoothing, 523
 - 11.4.1 Kalman Filter, 523
 - 11.4.2 State Estimation Error and Forecast Error, 525
 - 11.4.3 State Smoothing, 526
 - 11.4.4 Disturbance Smoothing, 528
- 11.5 Missing Values, 531
- 11.6 Forecasting, 532
- 11.7 Application, 533
- Exercises, 540
- References, 541

12. Markov Chain Monte Carlo Methods with Applications 543

- 12.1 Markov Chain Simulation, 544
- 12.2 Gibbs Sampling, 545
- 12.3 Bayesian Inference, 547
 - 12.3.1 Posterior Distributions, 547
 - 12.3.2 Conjugate Prior Distributions, 548
- 12.4 Alternative Algorithms, 551
 - 12.4.1 Metropolis Algorithm, 551
 - 12.4.2 Metropolis–Hasting Algorithm, 552
 - 12.4.3 Griddy Gibbs, 552
- 12.5 Linear Regression with Time Series Errors, 553
- 12.6 Missing Values and Outliers, 558
 - 12.6.1 Missing Values, 559
 - 12.6.2 Outlier Detection, 561
- 12.7 Stochastic Volatility Models, 565
 - 12.7.1 Estimation of Univariate Models, 566
 - 12.7.2 Multivariate Stochastic Volatility Models, 571
- 12.8 A New Approach to SV Estimation, 578
- 12.9 Markov Switching Models, 588
- 12.10 Forecasting, 594
- 12.11 Other Applications, 597
- Exercises, 597
- References, 598

Index 601

Preface

The subject of *financial time series analysis* has attracted substantial attention in recent years, especially with the 2003 Nobel awards to Professors Robert Engle and Clive Granger. At the same time, the field of financial econometrics has undergone various new developments, especially in high-frequency finance, stochastic volatility, and software availability. There is a need to make the material more complete and accessible for advanced undergraduate and graduate students, practitioners, and researchers. The main goals in preparing this second edition have been to bring the book up to date both in new developments and empirical analysis, and to enlarge the core material of the book by including consistent covariance estimation under heteroscedasticity and serial correlation, alternative approaches to volatility modeling, financial factor models, state-space models, Kalman filtering, and estimation of stochastic diffusion models.

The book therefore has been extended to 10 chapters and substantially revised to include S-Plus commands and illustrations. Many empirical demonstrations and exercises are updated so that they include the most recent data.

The two new chapters are Chapter 9, Principal Component Analysis and Factor Models, and Chapter 11, State-Space Models and Kalman Filter. The factor models discussed include macroeconomic, fundamental, and statistical factor models. They are simple and powerful tools for analyzing high-dimensional financial data such as portfolio returns. Empirical examples are used to demonstrate the applications. The state-space model and Kalman filter are added to demonstrate their applicability in finance and ease in computation. They are used in Chapter 12 to estimate stochastic volatility models under the general Markov chain Monte Carlo (MCMC) framework. The estimation also uses the technique of forward filtering and backward sampling to gain computational efficiency.

A brief summary of the added material in the second edition is:

1. To update the data used throughout the book.
2. To provide S-Plus commands and demonstrations.
3. To consider unit-root tests and methods for consistent estimation of the covariance matrix in the presence of conditional heteroscedasticity and serial correlation in Chapter 2.

4. To describe alternative approaches to volatility modeling, including use of high-frequency transactions data and daily high and low prices of an asset in Chapter 3.
5. To give more applications of nonlinear models and methods in Chapter 4.
6. To introduce additional concepts and applications of value at risk in Chapter 7.
7. To discuss cointegrated vector AR models in Chapter 8.
8. To cover various multivariate volatility models in Chapter 10.
9. To add an effective MCMC method for estimating stochastic volatility models in Chapter 12.

The revision benefits greatly from constructive comments of colleagues, friends, and many readers on the first edition. I am indebted to them all. In particular, I thank J. C. Artigas, Spencer Graves, Chung-Ming Kuan, Henry Lin, Daniel Peña, Jeff Russell, Michael Steele, George Tiao, Mark Wohar, Eric Zivot, and students of my MBA classes on financial time series for their comments and discussions, and Rosalyn Farkas, production editor, at John Wiley. I also thank my wife and children for their unconditional support and encouragement. Part of my research in financial econometrics is supported by the National Science Foundation, the High-Frequency Finance Project of the Institute of Economics, Academia Sinica, and the Graduate School of Business, University of Chicago.

Finally, the website for the book is:

gsbwww.uchicago.edu/fac/ruey.tsay/teaching/fts2.

RUEY S. TSAY

*University of Chicago
Chicago, Illinois*

Preface for the First Edition

This book grew out of an MBA course in analysis of financial time series that I have been teaching at the University of Chicago since 1999. It also covers materials of Ph.D. courses in time series analysis that I taught over the years. It is an introductory book intended to provide a comprehensive and systematic account of financial econometric models and their application to modeling and prediction of financial time series data. The goals are to learn basic characteristics of financial data, understand the application of financial econometric models, and gain experience in analyzing financial time series.

The book will be useful as a text of time series analysis for MBA students with finance concentration or senior undergraduate and graduate students in business, economics, mathematics, and statistics who are interested in financial econometrics. The book is also a useful reference for researchers and practitioners in business, finance, and insurance facing value at risk calculation, volatility modeling, and analysis of serially correlated data.

The distinctive features of this book include the combination of recent developments in financial econometrics in the econometric and statistical literature. The developments discussed include the timely topics of value at risk (VaR), high-frequency data analysis, and Markov chain Monte Carlo (MCMC) methods. In particular, the book covers some recent results that are yet to appear in academic journals; see Chapter 6 on derivative pricing using jump diffusion with closed-form formulas, Chapter 7 on value at risk calculation using extreme value theory based on a nonhomogeneous two-dimensional Poisson process, and Chapter 9 on multivariate volatility models with time-varying correlations. MCMC methods are introduced because they are powerful and widely applicable in financial econometrics. These methods will be used extensively in the future.

Another distinctive feature of this book is the emphasis on real examples and data analysis. Real financial data are used throughout the book to demonstrate applications of the models and methods discussed. The analysis is carried out by using several computer packages; the SCA (the Scientific Computing Associates)

for building linear time series models, the RATS (regression analysis for time series) for estimating volatility models, and the S-Plus for implementing neural networks and obtaining postscript plots. Some commands required to run these packages are given in appendixes of appropriate chapters. In particular, complicated RATS programs used to estimate multivariate volatility models are shown in Appendix A of Chapter 9. Some Fortran programs written by myself and others are used to price simple options, estimate extreme value models, calculate VaR, and carry out Bayesian analysis. Some data sets and programs are accessible from the World Wide Web at <http://www.gsb.uchicago.edu/fac/ruey.tsay/teaching/fts>.

The book begins with some basic characteristics of financial time series data in Chapter 1. The other chapters are divided into three parts. The first part, consisting of Chapters 2 to 7, focuses on analysis and application of univariate financial time series. The second part of the book covers Chapters 8 and 9 and is concerned with the return series of multiple assets. The final part of the book is Chapter 10, which introduces Bayesian inference in finance via MCMC methods.

A knowledge of basic statistical concepts is needed to fully understand the book. Throughout the chapters, I have provided a brief review of the necessary statistical concepts when they first appear. Even so, a prerequisite in statistics or business statistics that includes probability distributions and linear regression analysis is highly recommended. A knowledge of finance will be helpful in understanding the applications discussed throughout the book. However, readers with advanced background in econometrics and statistics can find interesting and challenging topics in many areas of the book.

An MBA course may consist of Chapters 2 and 3 as a core component, followed by some nonlinear methods (e.g., the neural network of Chapter 4 and the applications discussed in Chapters 5–7 and 10). Readers who are interested in Bayesian inference may start with the first five sections of Chapter 10.

Research in financial time series evolves rapidly and new results continue to appear regularly. Although I have attempted to provide broad coverage, there are many subjects that I do not cover or can only mention in passing.

I sincerely thank my teacher and dear friend, George C. Tiao, for his guidance, encouragement, and deep conviction regarding statistical applications over the years. I am grateful to Steve Quigley, Heather Haselkorn, Leslie Galen, Danielle LaCouriere, and Amy Hendrickson for making the publication of this book possible, to Richard Smith for sending me the estimation program of extreme value theory, to Bonnie K. Ray for helpful comments on several chapters, to Steve Kou for sending me his preprint on jump diffusion models, to Robert E. McCulloch for many years of collaboration on MCMC methods, to many students in my courses on analysis of financial time series for their feedback and inputs, and to Jeffrey Russell and Michael Zhang for insightful discussions concerning analysis of high-frequency financial data. To all these wonderful people I owe a deep sense of gratitude. I am also grateful for the support of the Graduate School of Business, University of Chicago and the National Science Foundation. Finally, my heartfelt thanks to my wife, Teresa, for her continuous support, encouragement, and

understanding; to Julie, Richard, and Vicki for bringing me joy and inspirations; and to my parents for their love and care.

RUEY S. TSAY

University of Chicago
Chicago, Illinois

Financial Time Series and Their Characteristics

Financial time series analysis is concerned with the theory and practice of asset valuation over time. It is a highly empirical discipline, but like other scientific fields theory forms the foundation for making inference. There is, however, a key feature that distinguishes financial time series analysis from other time series analysis. Both financial theory and its empirical time series contain an element of uncertainty. For example, there are various definitions of asset volatility, and for a stock return series, the volatility is not directly observable. As a result of the added uncertainty, statistical theory and methods play an important role in financial time series analysis.

The objective of this book is to provide some knowledge of financial time series, introduce some statistical tools useful for analyzing these series, and gain experience in financial applications of various econometric methods. We begin with the basic concepts of asset returns and a brief introduction to the processes to be discussed throughout the book. Chapter 2 reviews basic concepts of linear time series analysis such as stationarity and autocorrelation function, introduces simple linear models for handling serial dependence of the series, and discusses regression models with time series errors, seasonality, unit-root nonstationarity, and long-memory processes. The chapter also provides methods for consistent estimation of the covariance matrix in the presence of conditional heteroscedasticity and serial correlations. Chapter 3 focuses on modeling conditional heteroscedasticity (i.e., the conditional variance of an asset return). It discusses various econometric models developed recently to describe the evolution of volatility of an asset return over time. The chapter also discusses alternative methods to volatility modeling, including use of high-frequency transactions data and daily high and low prices of an asset. In Chapter 4, we address nonlinearity in financial time series, introduce test statistics that can discriminate nonlinear series from linear ones, and discuss several nonlinear models. The chapter also introduces nonparametric

estimation methods and neural networks and shows various applications of non-linear models in finance. Chapter 5 is concerned with analysis of high-frequency financial data and its application to market microstructure. It shows that nonsynchronous trading and bid–ask bounce can introduce serial correlations in a stock return. It also studies the dynamic of time duration between trades and some econometric models for analyzing transactions data. In Chapter 6, we introduce continuous-time diffusion models and Ito’s lemma. Black–Scholes option pricing formulas are derived and a simple jump diffusion model is used to capture some characteristics commonly observed in options markets. Chapter 7 discusses extreme value theory, heavy-tailed distributions, and their application to financial risk management. In particular, it discusses various methods for calculating value at risk of a financial position. Chapter 8 focuses on multivariate time series analysis and simple multivariate models with emphasis on the lead–lag relationship between time series. The chapter also introduces cointegration, some cointegration tests, and threshold cointegration and applies the concept of cointegration to investigate arbitrage opportunity in financial markets. Chapter 9 discusses ways to simplify the dynamic structure of a multivariate series and methods to reduce the dimension. It introduces and demonstrates three types of factor model to analyze returns of multiple assets. In Chapter 10, we introduce multivariate volatility models, including those with time-varying correlations, and discuss methods that can be used to reparameterize a conditional covariance matrix to satisfy the positiveness constraint and reduce the complexity in volatility modeling. Chapter 11 introduces state-space models and the Kalman filter and discusses the relationship between state-space models and other econometric models discussed in the book. It also gives several examples of financial applications. Finally, in Chapter 12, we introduce some newly developed Markov chain Monte Carlo (MCMC) methods in the statistical literature and apply the methods to various financial research problems, such as the estimation of stochastic volatility and Markov switching models.

The book places great emphasis on application and empirical data analysis. Every chapter contains real examples and, on many occasions, empirical characteristics of financial time series are used to motivate the development of econometric models. Computer programs and commands used in data analysis are provided when needed. In some cases, the programs are given in an appendix. Many real data sets are also used in the exercises of each chapter.

1.1 ASSET RETURNS

Most financial studies involve returns, instead of prices, of assets. Campbell, Lo, and MacKinlay (1997) give two main reasons for using returns. First, for average investors, return of an asset is a complete and scale-free summary of the investment opportunity. Second, return series are easier to handle than price series because the former have more attractive statistical properties. There are, however, several definitions of an asset return.

Let P_t be the price of an asset at time index t . We discuss some definitions of returns that are used throughout the book. Assume for the moment that the asset pays no dividends.

One-Period Simple Return

Holding the asset for one period from date $t - 1$ to date t would result in a *simple gross return*

$$1 + R_t = \frac{P_t}{P_{t-1}} \quad \text{or} \quad P_t = P_{t-1}(1 + R_t). \quad (1.1)$$

The corresponding one-period *simple net return* or *simple return* is

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (1.2)$$

Multiperiod Simple Return

Holding the asset for k periods between dates $t - k$ and t gives a k -period simple gross return

$$\begin{aligned} 1 + R_t[k] &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}} \\ &= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}) \\ &= \prod_{j=0}^{k-1} (1 + R_{t-j}). \end{aligned}$$

Thus, the k -period simple gross return is just the product of the k one-period simple gross returns involved. This is called a compound return. The k -period simple net return is $R_t[k] = (P_t - P_{t-k})/P_{t-k}$.

In practice, the actual time interval is important in discussing and comparing returns (e.g., monthly return or annual return). If the time interval is not given, then it is implicitly assumed to be one year. If the asset was held for k years, then the annualized (average) return is defined as

$$\text{Annualized}\{R_t[k]\} = \left[\prod_{j=0}^{k-1} (1 + R_{t-j}) \right]^{1/k} - 1.$$

This is a geometric mean of the k one-period simple gross returns involved and can be computed by

$$\text{Annualized}\{R_t[k]\} = \exp \left[\frac{1}{k} \sum_{j=0}^{k-1} \ln(1 + R_{t-j}) \right] - 1,$$

where $\exp(x)$ denotes the exponential function and $\ln(x)$ is the natural logarithm of the positive number x . Because it is easier to compute arithmetic average than

geometric mean and the one-period returns tend to be small, one can use a first-order Taylor expansion to approximate the annualized return and obtain

$$\text{Annualized}\{R_t[k]\} \approx \frac{1}{k} \sum_{j=0}^{k-1} R_{t-j}. \quad (1.3)$$

Accuracy of the approximation in Eq. (1.3) may not be sufficient in some applications, however.

Continuous Compounding

Before introducing continuously compounded return, we discuss the effect of compounding. Assume that the interest rate of a bank deposit is 10% per annum and the initial deposit is \$1.00. If the bank pays interest once a year, then the net value of the deposit becomes $\$1(1 + 0.1) = \1.1 one year later. If the bank pays interest semiannually, the 6-month interest rate is $10\%/2 = 5\%$ and the net value is $\$1(1 + 0.1/2)^2 = \1.1025 after the first year. In general, if the bank pays interest m times a year, then the interest rate for each payment is $10\%/m$ and the net value of the deposit becomes $\$1(1 + 0.1/m)^m$ one year later. Table 1.1 gives the results for some commonly used time intervals on a deposit of \$1.00 with interest rate of 10% per annum. In particular, the net value approaches \$1.1052, which is obtained by $\exp(0.1)$ and referred to as the result of continuous compounding. The effect of compounding is clearly seen.

In general, the net asset value A of continuous compounding is

$$A = C \exp(r \times n), \quad (1.4)$$

where r is the interest rate per annum, C is the initial capital, and n is the number of years. From Eq. (1.4), we have

$$C = A \exp(-r \times n), \quad (1.5)$$

which is referred to as the *present value* of an asset that is worth A dollars n years from now, assuming that the continuously compounded interest rate is r per annum.

Table 1.1. Illustration of the Effects of Compounding^a

Type	Number of Payments	Interest Rate per Period	Net Value
Annual	1	0.1	\$1.10000
Semiannual	2	0.05	\$1.10250
Quarterly	4	0.025	\$1.10381
Monthly	12	0.0083	\$1.10471
Weekly	52	0.1/52	\$1.10506
Daily	365	0.1/365	\$1.10516
Continuously	∞		\$1.10517

^aThe time interval is 1 year and the interest rate is 10% per annum.

Continuously Compounded Return

The natural logarithm of the simple gross return of an asset is called the continuously compounded return or *log return*:

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1}, \quad (1.6)$$

where $p_t = \ln(P_t)$. Continuously compounded returns r_t enjoy some advantages over the simple net returns R_t . First, consider multiperiod returns. We have

$$\begin{aligned} r_t[k] &= \ln(1 + R_t[k]) = \ln[(1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1})] \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \cdots + \ln(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \cdots + r_{t-k+1}. \end{aligned}$$

Thus, the continuously compounded multiperiod return is simply the sum of continuously compounded one-period returns involved. Second, statistical properties of log returns are more tractable.

Portfolio Return

The simple net return of a portfolio consisting of N assets is a weighted average of the simple net returns of the assets involved, where the weight on each asset is the percentage of the portfolio's value invested in that asset. Let p be a portfolio that places weight w_i on asset i . Then the simple return of p at time t is $R_{p,t} = \sum_{i=1}^N w_i R_{it}$, where R_{it} is the simple return of asset i .

The continuously compounded returns of a portfolio, however, do not have the above convenient property. If the simple returns R_{it} are all small in magnitude, then we have $r_{p,t} \approx \sum_{i=1}^N w_i r_{it}$, where $r_{p,t}$ is the continuously compounded return of the portfolio at time t . This approximation is often used to study portfolio returns.

Dividend Payment

If an asset pays dividends periodically, we must modify the definitions of asset returns. Let D_t be the dividend payment of an asset between dates $t - 1$ and t and P_t be the price of the asset at the end of period t . Thus, dividend is not included in P_t . Then the simple net return and continuously compounded return at time t become

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1, \quad r_t = \ln(P_t + D_t) - \ln(P_{t-1}).$$

Excess Return

Excess return of an asset at time t is the difference between the asset's return and the return on some reference asset. The reference asset is often taken to be riskless such as a short-term U.S. Treasury bill return. The simple excess return and log excess return of an asset are then defined as

$$Z_t = R_t - R_{0t}, \quad z_t = r_t - r_{0t}, \quad (1.7)$$

where R_{0t} and r_{0t} are the simple and log returns of the reference asset, respectively. In the finance literature, the excess return is thought of as the payoff on an arbitrage portfolio that goes long in an asset and short in the reference asset with no net initial investment.

Remark. A long financial position means owning the asset. A short position involves selling an asset one does not own. This is accomplished by borrowing the asset from an investor who has purchased it. At some subsequent date, the short seller is obligated to buy exactly the same number of shares borrowed to pay back the lender. Because the repayment requires equal shares rather than equal dollars, the short seller benefits from a decline in the price of the asset. If cash dividends are paid on the asset while a short position is maintained, these are paid to the buyer of the short sale. The short seller must also compensate the lender by matching the cash dividends from his own resources. In other words, the short seller is also obligated to pay cash dividends on the borrowed asset to the lender. $\square$

Summary of Relationship

The relationships between simple return R_t and continuously compounded (or log) return r_t are

$$r_t = \ln(1 + R_t), \quad R_t = e^{r_t} - 1.$$

If the returns R_t and r_t are in percentages, then

$$r_t = 100 \ln \left(1 + \frac{R_t}{100} \right), \quad R_t = 100(e^{r_t/100} - 1).$$

Temporal aggregation of the returns produces

$$1 + R_t[k] = (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}),$$

$$r_t[k] = r_t + r_{t-1} + \cdots + r_{t-k+1}.$$

If the continuously compounded interest rate is r per annum, then the relationship between present and future values of an asset is

$$A = C \exp(r \times n), \quad C = A \exp(-r \times n).$$

Example 1.1. If the monthly log return of an asset is 4.46%, then the corresponding monthly simple return is $100[\exp(4.46/100) - 1] = 4.56\%$. Also, if the monthly log returns of the asset within a quarter are 4.46%, -7.34%, and 10.77%, respectively, then the quarterly log return of the asset is $(4.46 - 7.34 + 10.77)\% = 7.89\%$.

1.2 DISTRIBUTIONAL PROPERTIES OF RETURNS

To study asset returns, it is best to begin with their distributional properties. The objective here is to understand the behavior of the returns across assets and over time. Consider a collection of N assets held for T time periods, say, $t = 1, \dots, T$. For each asset i , let r_{it} be its log return at time t . The log returns under study are $\{r_{it}; i = 1, \dots, N; t = 1, \dots, T\}$. One can also consider the simple returns $\{R_{it}; i = 1, \dots, N; t = 1, \dots, T\}$ and the log excess returns $\{z_{it}; i = 1, \dots, N; t = 1, \dots, T\}$.

1.2.1 Review of Statistical Distributions and Their Moments

We briefly review some basic properties of statistical distributions and the moment equations of a random variable. Let R^k be the k -dimensional Euclidean space. A point in R^k is denoted by $\mathbf{x} \in R^k$. Consider two random vectors $\mathbf{X} = (X_1, \dots, X_k)'$ and $\mathbf{Y} = (Y_1, \dots, Y_q)'$. Let $P(\mathbf{X} \in A, \mathbf{Y} \in B)$ be the probability that $\mathbf{X}$ is in the subspace $A \subset R^k$ and $\mathbf{Y}$ is in the subspace $B \subset R^q$. For most of the cases considered in this book, both random vectors are assumed to be continuous.

Joint Distribution

The function

$$F_{X,Y}(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta}) = P(\mathbf{X} \leq \mathbf{x}, \mathbf{Y} \leq \mathbf{y}; \boldsymbol{\theta}),$$

where $\mathbf{x} \in R^p$, $\mathbf{y} \in R^q$, and the inequality “ $\leq$ ” is a component-by-component operation, is a joint distribution function of $\mathbf{X}$ and $\mathbf{Y}$ with parameter $\boldsymbol{\theta}$. Behavior of $\mathbf{X}$ and $\mathbf{Y}$ is characterized by $F_{X,Y}(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta})$. If the joint probability density function $f_{x,y}(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta})$ of $\mathbf{X}$ and $\mathbf{Y}$ exists, then

$$F_{X,Y}(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta}) = \int_{-\infty}^{\mathbf{x}} \int_{-\infty}^{\mathbf{y}} f_{x,y}(\mathbf{w}, \mathbf{z}; \boldsymbol{\theta}) d\mathbf{z} d\mathbf{w}.$$

In this case, $\mathbf{X}$ and $\mathbf{Y}$ are continuous random vectors.

Marginal Distribution

The marginal distribution of $\mathbf{X}$ is given by

$$F_X(\mathbf{x}; \boldsymbol{\theta}) = F_{X,Y}(\mathbf{x}, \infty, \dots, \infty; \boldsymbol{\theta}).$$

Thus, the marginal distribution of $\mathbf{X}$ is obtained by integrating out $\mathbf{Y}$. A similar definition applies to the marginal distribution of $\mathbf{Y}$.

If $k = 1$, X is a scalar random variable and the distribution function becomes

$$F_X(x) = P(X \leq x; \boldsymbol{\theta}),$$

which is known as the cumulative distribution function (CDF) of X . The CDF of a random variable is nondecreasing (i.e., $F_X(x_1) \leq F_X(x_2)$ if $x_1 \leq x_2$) and satisfies

$F_X(-\infty) = 0$ and $F_X(\infty) = 1$. For a given probability p , the smallest real number x_p such that $p \leq F_X(x_p)$ is called the p th quantile of the random variable X . More specifically,

$$x_p = \inf_x \{x | p \leq F_X(x)\}.$$

We use the CDF to compute the p value of a test statistic in the book.

Conditional Distribution

The conditional distribution of X given $Y \leq y$ is given by

$$F_{X|Y \leq y}(x; \theta) = \frac{P(X \leq x, Y \leq y; \theta)}{P(Y \leq y; \theta)}.$$

If the probability density functions involved exist, then the conditional density of X given $Y = y$ is

$$f_{x|y}(x; \theta) = \frac{f_{x,y}(x, y; \theta)}{f_y(y; \theta)}, \quad (1.8)$$

where the marginal density function $f_y(y; \theta)$ is obtained by

$$f_y(y; \theta) = \int_{-\infty}^{\infty} f_{x,y}(x, y; \theta) dx.$$

From Eq. (1.8), the relation among joint, marginal, and conditional distributions is

$$f_{x,y}(x, y; \theta) = f_{x|y}(x; \theta) \times f_y(y; \theta). \quad (1.9)$$

This identity is used extensively in time series analysis (e.g., in maximum likelihood estimation). Finally, X and Y are independent random vectors if and only if $f_{x|y}(x; \theta) = f_x(x; \theta)$. In this case, $f_{x,y}(x, y; \theta) = f_x(x; \theta)f_y(y; \theta)$.

Moments of a Random Variable

The ℓ th moment of a continuous random variable X is defined as

$$m'_\ell = E(X^\ell) = \int_{-\infty}^{\infty} x^\ell f(x) dx,$$

where E stands for expectation and $f(x)$ is the probability density function of X . The first moment is called the *mean* or *expectation* of X . It measures the central location of the distribution. We denote the mean of X by μ_x . The ℓ th central moment of X is defined as

$$m_\ell = E[(X - \mu_x)^\ell] = \int_{-\infty}^{\infty} (x - \mu_x)^\ell f(x) dx$$

provided that the integral exists. The second central moment, denoted by σ_x^2 , measures the variability of X and is called the *variance* of X . The positive square root, σ_x , of variance is the *standard deviation* of X . The first two moments of a random variable uniquely determine a normal distribution. For other distributions, higher order moments are also of interest.

The third central moment measures the symmetry of X with respect to its mean, whereas the fourth central moment measures the tail behavior of X . In statistics, *skewness* and *kurtosis*, which are normalized third and fourth central moments of X , are often used to summarize the extent of asymmetry and tail thickness. Specifically, the skewness and kurtosis of X are defined as

$$S(x) = E \left[\frac{(X - \mu_x)^3}{\sigma_x^3} \right], \quad K(x) = E \left[\frac{(X - \mu_x)^4}{\sigma_x^4} \right].$$

The quantity $K(x) - 3$ is called the *excess kurtosis* because $K(x) = 3$ for a normal distribution. Thus, the excess kurtosis of a normal random variable is zero. A distribution with positive excess kurtosis is said to have heavy tails, implying that the distribution puts more mass on the tails of its support than a normal distribution does. In practice, this means that a random sample from such a distribution tends to contain more extreme values. Such a distribution is said to be *leptokurtic*. On the other hand, a distribution with negative excess kurtosis has short tails (e.g., a uniform distribution over a finite interval). Such a distribution is said to be *platykurtic*.

In application, skewness and kurtosis can be estimated by their sample counterparts. Let $\{x_1, \dots, x_T\}$ be a random sample of X with T observations. The sample mean is

$$\hat{\mu}_x = \frac{1}{T} \sum_{t=1}^T x_t, \quad (1.10)$$

the sample variance is

$$\hat{\sigma}_x^2 = \frac{1}{T-1} \sum_{t=1}^T (x_t - \hat{\mu}_x)^2, \quad (1.11)$$

the sample skewness is

$$\hat{S}(x) = \frac{1}{(T-1)\hat{\sigma}_x^3} \sum_{t=1}^T (x_t - \hat{\mu}_x)^3, \quad (1.12)$$

and the sample kurtosis is

$$\hat{K}(x) = \frac{1}{(T-1)\hat{\sigma}_x^4} \sum_{t=1}^T (x_t - \hat{\mu}_x)^4. \quad (1.13)$$

Under the normality assumption, $\hat{S}(x)$ and $\hat{K}(x) - 3$ are distributed asymptotically as normal with zero mean and variances $6/T$ and $24/T$, respectively; see Snedecor

and Cochran (1980, p. 78). These asymptotic properties can be used to test the normality of asset returns. Given an asset return series $\{r_1, \dots, r_T\}$, to test the skewness of the returns, we consider the null hypothesis $H_o : S(r) = 0$ versus the alternative hypothesis $H_a : S(r) \neq 0$. The t -ratio statistic of the sample skewness in Eq. (1.12) is

$$t = \frac{\hat{S}(r)}{\sqrt{6/T}}.$$

The decision rule is as follows. Reject the null hypothesis at the α significance level, if $|t| > Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the upper $100(\alpha/2)$ th quantile of the standard normal distribution. Alternatively, one can compute the p -value of the test statistic t and reject H_o if and only if the p -value is less than α .

Similarly, one can test the excess kurtosis of the return series using the hypotheses $H_o : K(r) - 3 = 0$ versus $H_a : K(r) - 3 \neq 0$. The test statistic is

$$t = \frac{\hat{K}(r) - 3}{\sqrt{24/T}},$$

which is asymptotically a standard normal random variable. The decision rule is to reject H_o if and only if the p -value of the test statistic is less than the significance level α . Jarque and Bera (1987) combine the two prior tests and use the test statistic

$$JB = \frac{\hat{S}^2(r)}{6/T} + \frac{(\hat{K}(r) - 3)^2}{24/T},$$

which is asymptotically distributed as a chi-squared random variable with 2 degrees of freedom, to test for the normality of r_t . One rejects H_o of normality if the p -value of the JB statistic is less than the significance level.

Example 1.2. Consider the daily simple returns of the IBM stock used in Table 1.2. The sample skewness and kurtosis of the returns are parts of the descriptive (or summary) statistics that can be obtained easily using various statistical software packages. Both SCA and S-Plus are used in the demonstration, where ‘d-ibmvwewsp6203.txt’ is the data file name. Note that in SCA the *kurtosis* denotes excess kurtosis. From the output, the excess kurtosis is high, indicating that the daily simple returns of IBM stock have heavy tails. To test the symmetry of return distribution, we use the test statistic

$$t = \frac{0.0775}{0.024} = 3.23,$$

which gives a p -value of about 0.001, indicating that the daily simple returns of IBM stock are significantly skewed to the right at the 5% level.

Table 1.2. Descriptive Statistics for Daily and Monthly Simple and Log Returns of Selected Indexes and Stocks^a

Security	Start	Size	Mean	Standard Deviation	Skewness	Excess Kurtosis	Minimum	Maximum
<i>Daily Simple Returns (%)</i>								
SP	62/7/3	10446	0.033	0.945	−0.95	25.76	−20.47	9.10
VW	62/7/3	10446	0.045	0.794	−0.76	18.32	−17.14	8.66
EW	62/7/3	10446	0.085	0.726	−0.89	13.42	−10.39	6.95
IBM	62/7/3	10446	0.052	1.648	−0.08	10.21	−22.96	13.16
Intel	72/12/15	7828	0.131	2.998	−0.16	5.85	−29.57	26.38
3M	62/7/3	10446	0.054	1.465	−0.28	12.87	−25.98	11.54
Microsoft	86/3/14	4493	0.157	2.505	−0.25	8.75	−30.12	19.57
Citi-Group	86/10/30	4333	0.110	2.289	−0.10	6.79	−21.74	20.76
<i>Daily Log Returns (%)</i>								
SP	62/7/3	10446	0.029	0.951	−1.41	36.91	−22.90	8.71
VW	62/7/3	10446	0.041	0.895	−1.06	23.91	−18.80	8.31
EW	62/7/3	10446	0.082	0.728	−1.29	14.70	−10.97	6.72
IBM	62/7/3	10446	0.039	1.649	−0.25	12.60	−26.09	12.37
Intel	72/12/15	7828	0.086	3.013	−0.54	7.54	−35.06	23.41
3M	62/7/3	10446	0.044	1.469	−0.69	20.06	−30.08	10.92
Microsoft	86/3/14	4493	0.126	2.518	−0.73	13.23	−35.83	17.87
Citi-Group	86/10/30	4333	0.084	2.289	−0.21	7.47	−24.51	18.86
<i>Monthly Simple Returns (%)</i>								
SP	62/1	936	0.64	5.63	−0.35	9.26	−29.94	42.22
VW	26/1	936	0.95	5.49	−0.18	7.52	−28.98	38.27
EW	26/1	936	1.31	7.49	−1.54	14.46	−31.18	65.51
IBM	26/1	936	1.42	7.11	−0.27	2.15	−26.19	35.38
Intel	73/1	372	2.71	13.42	−0.26	2.43	−44.87	62.50
3M	46/2	695	1.37	6.53	−0.24	0.96	−27.83	25.80
Microsoft	86/4	213	3.37	11.95	−0.53	1.40	−34.35	51.55
Citi-Group	86/11	206	2.20	9.52	−0.18	0.87	−34.48	26.08
<i>Monthly Log Returns (%)</i>								
SP	26/1	936	0.48	5.62	−0.50	7.77	−35.58	35.22
VW	26/1	936	0.79	5.48	−0.54	6.72	−34.22	32.41
EW	26/1	936	1.04	7.21	−0.29	8.40	−37.37	50.38
IBM	26/1	936	1.16	7.02	−0.15	2.04	−30.37	30.29
Intel	73/1	372	1.80	13.37	−0.60	2.90	−59.54	48.55
3M	46/2	695	1.16	6.43	−0.06	1.25	−32.61	22.95
Microsoft	86/4	213	2.66	11.48	−0.01	1.19	−42.09	41.58
Citi-Group	86/11	206	1.73	9.55	−0.65	2.08	−42.28	23.18

^aReturns are in percentages and the sample period ends on December 31, 2003. The statistics are defined in eqs. (1.10)–(1.13). VW, EW, and SP denote value-weighted, equal-weighted, and S&P composite index.

SCA Demonstration

% denotes explanation.

```
input date, ibm, vw, ew, sp. file 'd-ibmvewwsp6203.txt'
```

```
% Load data into SCA and name the columns date,
```

```
% ibm, vw, ew, and sp.
```

```
--
```

```
ibm=ibm*100 % Compute percentage returns
```

```
--
```

```
desc ibm % Obtain descriptive statistics of ibm
```

```
VARIABLE      NAME      IS      IBM
NUMBER OF OBSERVATIONS      10446
NUMBER OF MISSING VALUES      0
```

	STATISTIC	STD. ERROR	STATISTIC/S.E.
MEAN	0.0523	0.0161	3.2457
VARIANCE	2.7163		
STD DEVIATION	1.6481		
C.V.	31.4900		
SKEWNESS	0.0775	0.0240	
KURTOSIS	10.2144	0.0479	

	QUARTILE
MINIMUM	-22.9630
1ST QUARTILE	-0.8380
MEDIAN	0.0000
3RD QUARTILE	0.8805
MAXIMUM	13.1640

	RANGE
MAX - MIN	36.1270
Q3 - Q1	1.7185

S-Plus Demonstration

> is the prompt character and % marks explanation.

```
> module(finmetrics) % Load the Finmetrics module.
```

```
> x=matrix(scan(file='d-ibmvewwsp6203.txt'),5) % Load data
```

```
> ibm=x[2,]*100 % compute percentage returns
```

```
> summaryStats(ibm) % obtain summary statistics
```

Sample Quantiles:

min	1Q	median	3Q	max
-22.96	-0.838	0	0.8807	13.16

Sample Moments:

mean	std	skewness	kurtosis
0.05234	1.648	0.0775	13.22

Number of Observations: 10446

1.2.2 Distributions of Returns

The most general model for the log returns $\{r_{it}; i = 1, \dots, N; t = 1, \dots, T\}$ is its joint distribution function:

$$F_r(r_{11}, \dots, r_{N1}; r_{12}, \dots, r_{N2}; \dots; r_{1T}, \dots, r_{NT}; \mathbf{Y}; \boldsymbol{\theta}), \quad (1.14)$$

where $\mathbf{Y}$ is a state vector consisting of variables that summarize the environment in which asset returns are determined and $\boldsymbol{\theta}$ is a vector of parameters that uniquely determine the distribution function $F_r(\cdot)$. The probability distribution $F_r(\cdot)$ governs the stochastic behavior of the returns r_{it} and $\mathbf{Y}$. In many financial studies, the state vector $\mathbf{Y}$ is treated as given and the main concern is the conditional distribution of $\{r_{it}\}$ given $\mathbf{Y}$. Empirical analysis of asset returns is then to estimate the unknown parameter $\boldsymbol{\theta}$ and to draw statistical inference about the behavior of $\{r_{it}\}$ given some past log returns.

The model in Eq. (1.14) is too general to be of practical value. However, it provides a general framework with respect to which an econometric model for asset returns r_{it} can be put in a proper perspective.

Some financial theories such as the capital asset pricing model (CAPM) of Sharpe (1964) focus on the joint distribution of N returns at a single time index t (i.e., the distribution of $\{r_{1t}, \dots, r_{Nt}\}$). Other theories emphasize the dynamic structure of individual asset returns (i.e., the distribution of $\{r_{i1}, \dots, r_{iT}\}$ for a given asset i). In this book, we focus on both. In the univariate analysis of Chapters 2–7, our main concern is the joint distribution of $\{r_{it}\}_{t=1}^T$ for asset i . To this end, it is useful to partition the joint distribution as

$$\begin{aligned} F(r_{i1}, \dots, r_{iT}; \boldsymbol{\theta}) &= F(r_{i1})F(r_{i2}|r_{i1}) \cdots F(r_{iT}|r_{i,T-1}, \dots, r_{i1}) \\ &= F(r_{i1}) \prod_{t=2}^T F(r_{it}|r_{i,t-1}, \dots, r_{i1}), \end{aligned} \quad (1.15)$$

where, for simplicity, the parameter $\boldsymbol{\theta}$ is omitted. This partition highlights the temporal dependencies of the log return r_{it} . The main issue then is the specification of the conditional distribution $F(r_{it}|r_{i,t-1}, \dots, r_{i1})$, in particular, how the conditional distribution evolves over time. In finance, different distributional specifications lead to different theories. For instance, one version of the random-walk hypothesis is that the conditional distribution $F(r_{it}|r_{i,t-1}, \dots, r_{i1})$ is equal to the marginal distribution $F(r_{it})$. In this case, returns are temporally independent and, hence, not predictable.

It is customary to treat asset returns as continuous random variables, especially for index returns or stock returns calculated at a low frequency, and use their probability density functions. In this case, using the identity in Eq. (1.9), we can write the partition in Eq. (1.15) as

$$f(r_{i1}, \dots, r_{iT}; \boldsymbol{\theta}) = f(r_{i1}; \boldsymbol{\theta}) \prod_{t=2}^T f(r_{it}|r_{i,t-1}, \dots, r_{i1}, \boldsymbol{\theta}). \quad (1.16)$$

For high-frequency asset returns, discreteness becomes an issue. For example, stock prices change in multiples of a tick size on the New York Stock Exchange (NYSE). The tick size was one-eighth of a dollar before July 1997 and was one-sixteenth of a dollar from July 1997 to January 2001. Therefore, the tick-by-tick return of an individual stock listed on the NYSE is not continuous. We discuss high-frequency stock price changes and time durations between price changes later in Chapter 5.

Remark. On August 28, 2000, the NYSE began a pilot program with seven stocks priced in decimals and the American Stock Exchange (AMEX) began a pilot program with six stocks and two options classes. The NYSE added 57 stocks and 94 stocks to the program on September 25 and December 4, 2000, respectively. All NYSE and AMEX stocks started trading in decimals on January 29, 2001. $\square$

Equation (1.16) suggests that conditional distributions are more relevant than marginal distributions in studying asset returns. However, the marginal distributions may still be of some interest. In particular, it is easier to estimate marginal distributions than conditional distributions using past returns. In addition, in some cases, asset returns have weak empirical serial correlations, and, hence, their marginal distributions are close to their conditional distributions.

Several statistical distributions have been proposed in the literature for the marginal distributions of asset returns, including normal distribution, lognormal distribution, stable distribution, and scale-mixture of normal distributions. We briefly discuss these distributions.

Normal Distribution

A traditional assumption made in financial study is that the simple returns $\{R_{it}|t = 1, \dots, T\}$ are independently and identically distributed as normal with fixed mean and variance. This assumption makes statistical properties of asset returns tractable. But it encounters several difficulties. First, the lower bound of a simple return is -1 . Yet the normal distribution may assume any value in the real line and, hence, has no lower bound. Second, if R_{it} is normally distributed, then the multiperiod simple return $R_{it}[k]$ is not normally distributed because it is a product of one-period returns. Third, the normality assumption is not supported by many empirical asset returns, which tend to have a positive excess kurtosis.

Lognormal Distribution

Another commonly used assumption is that the log returns r_t of an asset are independent and identically distributed (iid) as normal with mean μ and variance σ^2 . The simple returns are then iid lognormal random variables with mean and variance given by

$$E(R_t) = \exp\left(\mu + \frac{\sigma^2}{2}\right) - 1, \quad \text{Var}(R_t) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]. \quad (1.17)$$

These two equations are useful in studying asset returns (e.g., in forecasting using models built for log returns). Alternatively, let m_1 and m_2 be the mean and variance of the simple return R_t , which is lognormally distributed. Then the mean and variance of the corresponding log return r_t are

$$E(r_t) = \ln \left(\frac{m_1 + 1}{\sqrt{1 + m_2/(1 + m_1)^2}} \right), \quad \text{Var}(r_t) = \ln \left(1 + \frac{m_2}{(1 + m_1)^2} \right).$$

Because the sum of a finite number of iid normal random variables is normal, $r_t[k]$ is also normally distributed under the normal assumption for $\{r_t\}$. In addition, there is no lower bound for r_t , and the lower bound for R_t is satisfied using $1 + R_t = \exp(r_t)$. However, the lognormal assumption is not consistent with all the properties of historical stock returns. In particular, many stock returns exhibit a positive excess kurtosis.

Stable Distribution

The stable distributions are a natural generalization of normal in that they are stable under addition, which meets the need of continuously compounded returns r_t . Furthermore, stable distributions are capable of capturing excess kurtosis shown by historical stock returns. However, non-normal stable distributions do not have a finite variance, which is in conflict with most finance theories. In addition, statistical modeling using non-normal stable distributions is difficult. An example of non-normal stable distributions is the Cauchy distribution, which is symmetric with respect to its median but has infinite variance.

Scale Mixture of Normal Distributions

Recent studies of stock returns tend to use scale mixture or finite mixture of normal distributions. Under the assumption of scale mixture of normal distributions, the log return r_t is normally distributed with mean μ and variance σ^2 [i.e., $r_t \sim N(\mu, \sigma^2)$]. However, σ^2 is a random variable that follows a positive distribution (e.g., σ^{-2} follows a gamma distribution). An example of finite mixture of normal distributions is

$$r_t \sim (1 - X)N(\mu, \sigma_1^2) + XN(\mu, \sigma_2^2),$$

where X is a Bernoulli random variable such that $P(X = 1) = \alpha$ and $P(X = 0) = 1 - \alpha$ with $0 < \alpha < 1$, σ_1^2 is small, and σ_2^2 is relatively large. For instance, with $\alpha = 0.05$, the finite mixture says that 95% of the returns follow $N(\mu, \sigma_1^2)$ and 5% follow $N(\mu, \sigma_2^2)$. The large value of σ_2^2 enables the mixture to put more mass at the tails of its distribution. The low percentage of returns that are from $N(\mu, \sigma_2^2)$ says that the majority of the returns follow a simple normal distribution. Advantages of mixtures of normal include that they maintain the tractability of normal, have finite higher order moments, and can capture the excess kurtosis. Yet it is hard to estimate the mixture parameters (e.g., the α in the finite-mixture case).

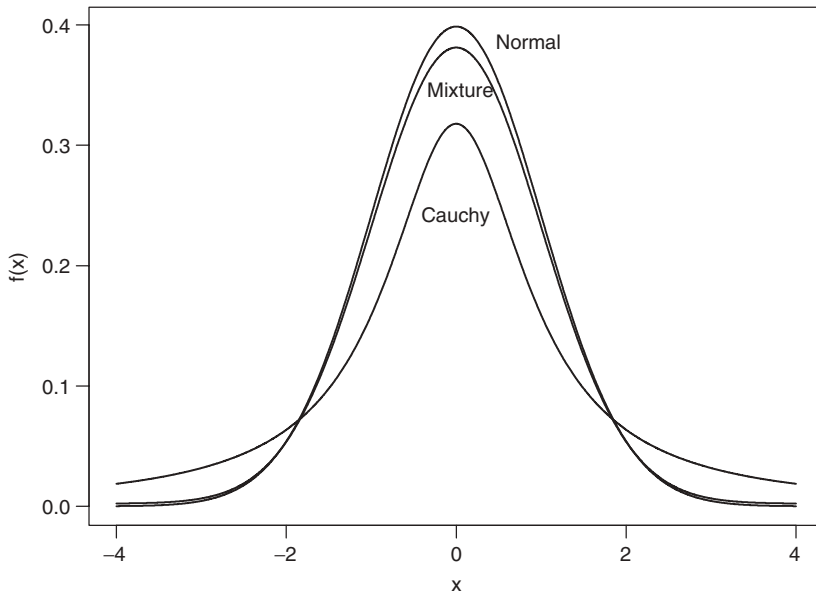


Figure 1.1. Comparison of finite mixture, stable, and standard normal density functions.

Figure 1.1 shows the probability density functions of a finite mixture of normal, Cauchy, and standard normal random variable. The finite mixture of normal is $(1 - X)N(0, 1) + X \times N(0, 16)$ with X being Bernoulli such that $P(X = 1) = 0.05$, and the density function of Cauchy is

$$f(x) = \frac{1}{\pi(1 + x^2)}, \quad -\infty < x < \infty.$$

It is seen that the Cauchy distribution has fatter tails than the finite mixture of normal, which, in turn, has fatter tails than the standard normal.

1.2.3 Multivariate Returns

Let $\mathbf{r}_t = (r_{1t}, \dots, r_{Nt})'$ be the log returns of N assets at time t . The multivariate analyses of Chapters 8 and 10 are concerned with the joint distribution of $\{\mathbf{r}_t\}_{t=1}^T$. This joint distribution can be partitioned in the same way as that of Eq. (1.15). The analysis is then focused on the specification of the conditional distribution function $F(\mathbf{r}_t | \mathbf{r}_{t-1}, \dots, \mathbf{r}_1, \boldsymbol{\theta})$. In particular, how the conditional expectation and conditional covariance matrix of $\mathbf{r}_t$ evolve over time constitute the main subjects of Chapters 8 and 10.

The mean vector and covariance matrix of a random vector $\mathbf{X} = (X_1, \dots, X_p)$ are defined as

$$\begin{aligned} E(\mathbf{X}) &= \boldsymbol{\mu}_x = [E(X_1), \dots, E(X_p)]', \\ \text{Cov}(\mathbf{X}) &= \boldsymbol{\Sigma}_x = E[(\mathbf{X} - \boldsymbol{\mu}_x)(\mathbf{X} - \boldsymbol{\mu}_x)'], \end{aligned}$$