THE FRANK J. FABOZZI SERIES

PROBABILITY *and* STATISTICS *for* FINANCE

SVETLOZAR T. RACHEV • MARKUS HÖCHSTÖTTER
FRANK J. FABOZZI • SERGIO M. FOCARDI

Probability and Statistics for Finance

SVETLOZAR T. RACHEV
MARKUS HÖCHSTÖTTER
FRANK J. FABOZZI
SERGIO M. FOCARDI

WILEY
John Wiley & Sons, Inc.

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

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 books. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Probability and statistics for finance / Svetlozar T. Rachev ... [et al.].
p. cm. – (Frank J. Fabozzi series ; 176)
Includes index.
ISBN 978-0-470-40093-7 (cloth); 978-0-470-90630-9 (ebk); 978-0-470-90631-6 (ebk); 978-0-470-90632-3 (ebk)
1. Finance–Statistical methods. 2. Statistics. 3. Probability measures. 4. Multivariate analysis. I. Rachev, S. T. (Svetlozar Todorov)
HG176.5.P76 2010
332.01'5195–dc22

2010027030

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

Preface

In this book, we provide an array of topics in probability and statistics that are applied to problems in finance. For example, there are applications to portfolio management, asset pricing, risk management, and credit risk modeling. Not only do we cover the basics found in a typical introductory book in probability and statistics, but we also provide unique coverage of several topics that are of special interest to finance students and finance professionals. Examples are coverage of probability distributions that deal with extreme events and statistical measures, which are particularly useful for portfolio managers and risk managers concerned with extreme events.

The book is divided into four parts. The six chapters in Part One cover descriptive statistics: the different methods for gathering data and presenting them in a more succinct way while still being as informative as possible. The basics of probability theory are covered in the nine chapters in Part Two. After describing the basic concepts of probability, we explain the different types of probability distributions (discrete and continuous), specific types of probability distributions, parameters of a probability distribution, joint probability distributions, conditional probability distributions, and dependence measures for two random variables. Part Three covers statistical inference: the method of drawing information from sample data about unknown parameters of the population from which the sample was drawn. The three chapters in Part Three deal with point estimates of a parameter, confidence intervals of a parameter, and testing hypotheses about the estimates of a parameter. In the last part of the book, Part Four, we provide coverage of the most widely used statistical tool in finance: multivariate regression analysis. In the first of the three chapters in this part, we begin with the assumptions of the multivariate regression model, how to estimate the parameters of the model, and then explain diagnostic checks to evaluate the quality of the estimates. After these basics are provided, we then focus on the design and the building process of multivariate regression models and finally on how to deal with violations of the assumptions of the model.

There are also four appendixes. Important mathematical functions and their features that are needed primarily in the context of Part Two of this book are covered in Appendix A. In Appendix B we explain the basics of matrix operations and concepts needed to aid in understanding the presentation in Part Four. The construction of the binomial and multinomial coefficients used in some discrete probability distributions and an application of the log-normally distributed stock price to derive the price of a certain type of option are provided in Appendix C and D, respectively.

We would like to thank Biliana Bagasheva for her coauthorship of Chapter 15 (Conditional Probability and Bayes' Rule). Anna Serbinenko provided helpful comments on several chapters of the book.

The following students reviewed various chapters and provided us with helpful comments: Kameliya Minova, Diana Trinh, Lindsay Morriss, Marwan ElChamaa, Jens Bürgin, Paul Jana, and Haike Dogendorf.

We also thank Megan Orem for her patience in typesetting this book and giving us the flexibility to significantly restructure the chapters in this book over the past three years.

Svetlozar T. Rachev
Markus Höchstötter
Frank J. Fabozzi
Sergio M. Focardi
April 2010

About the Authors

Svetlozar (Zari) T. Rachev completed his Ph.D. Degree in 1979 from Moscow State (Lomonosov) University, and his Doctor of Science Degree in 1986 from Steklov Mathematical Institute in Moscow. Currently, he is Chair-Professor in Statistics, Econometrics and Mathematical Finance at the University of Karlsruhe in the School of Economics and Business Engineering, and Professor Emeritus at the University of California, Santa Barbara in the Department of Statistics and Applied Probability. Professor Rachev has published 14 monographs, 10 handbooks, and special-edited volumes, and over 300 research articles. His recently coauthored books published by Wiley in mathematical finance and financial econometrics include *Financial Econometrics: From Basics to Advanced Modeling Techniques* (2007), and *Bayesian Methods in Finance* (2008). He is cofounder of Bravo Risk Management Group specializing in financial risk-management software. Bravo Group was acquired by FinAnalytica for which he currently serves as Chief-Scientist.

Markus Höchstötter is lecturer in statistics and econometrics at the Institute of Statistics, Econometrics and Mathematical Finance at the University of Karlsruhe (KIT). Dr. Höchstötter has authored articles on financial econometrics and credit derivatives. He earned a doctorate (Dr. rer. pol.) in Mathematical Finance and Financial Econometrics from the University of Karlsruhe.

Frank J. Fabozzi is Professor in the Practice of Finance in the School of Management and Becton Fellow at Yale University and an Affiliated Professor at the University of Karlsruhe's Institute of Statistics, Econometrics and Mathematical Finance. Prior to joining the Yale faculty, he was a Visiting Professor of Finance in the Sloan School at MIT. Professor Fabozzi is a Fellow of the International Center for Finance at Yale University and on the Advisory Council for the Department of Operations Research and Financial Engineering at Princeton University. He is the editor of the *Journal of Portfolio Management* and an associate editor of *Quantitative Finance*. He is a trustee for the BlackRock family of closed-end funds. In 2002, he was inducted into the Fixed Income Analysts Society's Hall of Fame and is the 2007 recipient of the C. Stewart Sheppard Award given by the CFA Institute. His recently coauthored books published by Wiley include *Institutional Investment Management* (2009), *Quantitative Equity Investing* (2010), *Bayesian Methods in Finance* (2008), *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures* (2008), *Financial Modeling of the Equity Market: From CAPM to Cointegration* (2006), *Robust Portfolio Optimization and Management* (2007), and *Financial Econometrics: From Basics to Advanced Modeling Techniques* (2007). Professor Fabozzi earned a doctorate in economics from the City University of New York in 1972. He earned the designation of Chartered Financial Analyst and Certified Public Accountant.

Sergio M. Focardi is Professor of Finance at the EDHEC Business School in Nice and the founding partner of the Paris-based consulting firm The Intertek Group. He is a member of the editorial board of the *Journal of Portfolio Management*. Professor Focardi has authored numerous articles and books on financial modeling and risk management including the following Wiley books: *Quantitative Equity Investing* (2010), *Financial Econometrics* (2007), *Financial Modeling of the Equity Market* (2006), *The Mathematics of Financial Modeling and Investment Management* (2004), *Risk Management: Framework, Methods and Practice* (1998), and *Modeling the Markets: New Theories and Techniques* (1997). He also authored two monographs published by the CFA Institute's monographs: *Challenges in Quantitative Equity Management* (2008) and *Trends in Quantitative Finance* (2006). Professor Focardi has been appointed as a speaker of the CFA Institute Speaker Retainer Program. His research interests include the econometrics of large equity portfolios and the modeling of regime changes. Professor Focardi holds a degree in Electronic Engineering from the University of Genoa and a Ph.D. in Mathematical Finance and Financial Econometrics from the University of Karlsruhe.

Contents

CHAPTER 1

Introduction

It is no surprise that the natural sciences (chemistry, physics, life sciences/biology, astronomy, earth science, and environmental science) and engineering are fields that rely on advanced quantitative methods. One of the toolsets used by professionals in these fields is from the branch of mathematics known as probability and statistics. The social sciences, such as psychology, sociology, political science, and economics, use probability and statistics to varying degrees.

There are branches within each field of the natural sciences and social sciences that utilize probability and statistics more than others. Specialists in these areas not only apply the tools of probability and statistics, but they have also contributed to the field of statistics by developing techniques to organize, analyze, and test data. Let's look at examples from physics and engineering (the study of natural phenomena in terms of basic laws and physical quantities and the design of physical artifacts) and biology (the study of living organisms) in the natural sciences, and psychology (the study of the human mind) and economics (the study of production, resource allocation, and consumption of goods and services) in the social sciences.

Statistical physics is the branch of physics that applies probability and statistics for handling problems involving large populations of particles. One of the first areas of application was the explanation of thermodynamics laws in terms of statistical mechanics. It was an extraordinary scientific achievement with far-reaching consequences. In the field of *engineering*, the analysis of risk, be it natural or industrial, is another area that makes use of statistical methods. This discipline has contributed important innovations especially in the study of rare extreme events. The engineering of electronic communications applied statistical methods early, contributing to the development of fields such as queue theory (used in communication switching systems) and introduced the fundamental innovation of measuring information.

Biostatistics and *biomathematics* within the field of biology include many areas of great scientific interest such as public health, epidemiology, demography, and genetics, in addition to designing biological experiments (such as clinical experiments in medicine) and analyzing the results of those experiments. The study of the dynamics of populations and the study of evolutionary phenomena are two important fields in biomathematics. *Biometry* and *biometrics* apply statistical methods to identify quantities that characterize living objects.

Psychometrics, a branch of psychology, is concerned with designing tests and analyzing the results of those tests in an attempt to measure or quantify some human characteristic. Psychometrics has its origins in personality testing, intelligence testing, and vocational testing, but is now applied to measuring attitudes and beliefs and health-related tests.

Econometrics is the branch of economics that draws heavily on statistics for testing and analyzing economic relationships. Within econometrics, there are theoretical econometricians who analyze statistical properties of estimators of models. Several recipients of the Nobel Prize in Economic Sciences received the award as a result of their lifetime contribution to this branch of economics. To appreciate the importance of econometrics to the discipline of economics, when the first Nobel Prize in Economic Sciences was awarded in 1969, the corecipients were two econometricians, Jan Tinbergen and Ragnar Frisch (who is credited for first using the term econometrics in the sense that it is known today). Further specialization within econometrics, and the area that directly relates to this book, is *financial econometrics*. As Jianqing Fan (2004) writes, financial econometrics

> uses statistical techniques and economic theory to address a variety of problems from finance. These include building financial models, estimation and inferences of financial models, volatility estimation, risk management, testing financial economics theory, capital asset pricing, derivative pricing, portfolio allocation, risk-adjusted returns, simulating financial systems, hedging strategies, among others.

Robert Engle and Clive Granger, two econometricians who shared the 2003 Nobel Prize in Economics Sciences, have contributed greatly to the field of financial econometrics.

Historically, the core probability and statistics course offered at the university level to undergraduates has covered the fundamental principles and applied these principles across a wide variety of fields in the natural sciences and social sciences. Universities typically offered specialized courses within these fields to accommodate students who sought more focused applications. The exceptions were the schools of business administration that early on provided a course in probability and statistics with applications to business decision making. The applications cut across finance, marketing, management, and accounting. However, today, each of these areas in business requires specialized tools for dealing with real-world problems in their respective disciplines.

This brings us to the focus of this book. Finance is an area that relies heavily on probability and statistics. The quotation above by Jianqing Fan basically covers the wide range of applications within finance and identifies some of the unique applications. Two examples may help make this clear. First, in standard books on statistics, there is coverage of what one might refer to as "probability distributions with appealing properties." A distribution called the "normal distribution," referred to in the popular press as a "bell-shaped curve," is an example. Considerable space is devoted to this distribution and its application in standard textbooks. Yet, the overwhelming historical evidence suggests that real-world financial data commonly used in financial applications are not normally distributed. Instead, more focus should be on distributions that deal with extreme events, or, in other words, what are known as the "tails" of a distribution. In fact, many market commentators and regulators view the failure of financial institutions and major players in the financial markets to understand non-normal distributions as a major reason for the recent financial debacles throughout the world. This is one of the reasons that, in certain areas in finance, extreme event distributions (which draw from extreme value theory) have supplanted the normal distribution as the focus of attention. The recent financial crisis has clearly demonstrated that because of the highly leveraged position (i.e., large amount of borrowing relative to the value of equity) of financial institutions throughout the world, these entities are very sensitive to extreme events. This means that the management of these financial institutions must be aware of the nature of the tails of distributions, that is, the probability associated with extreme events.

As a second example, the statistical measure of correlation that measures a certain type of association between two random variables may make sense when the two random variables are normally distributed. However, correlation may be inadequate in describing the link between two random variables when a portfolio manager or risk manager is concerned with extreme events that can have disastrous outcomes for a portfolio or a financial institution. Typically models that are correlation based will underestimate the likelihood of extreme events occurring simultaneously. Alternative statistical measures that would be more helpful, the copula measure and the tail dependence, are typically not discussed in probability and statistics books.

It is safe to say that the global financial system has been transformed since the mid-1970s due to the development of models that can be used to value derivative instruments. Complex derivative instruments such as options, caps, floors, and swaptions can only be valued (i.e., priced) using tools from probability and statistical theory. While the model for such pric-