Preface

I developed this textbook while teaching the course *Statistics for Financial Engineering* to master's students in the financial engineering program at Cornell University. These students have already taken courses in portfolio management, fixed income securities, options, and stochastic calculus, so I concentrate on teaching statistics, data analysis, and the use of R, and I cover most sections of Chapters 4–9 and 17–20. These chapters alone are more than enough to fill a one semester course. I do not cover regression (Chapters 12–14 and 21) or the more advanced time series topics in Chapter 10, since these topics are covered in other courses. In the past, I have not covered cointegration (Chapter 15), but I will in the future. The master's students spend much of the third semester working on projects with investment banks or hedge funds. As a faculty adviser for several projects, I have seen the importance of cointegration.

A number of different courses might be based on this book. A two-semester sequence could cover most of the material. A one-semester course with more emphasis on finance would include Chapters 11 and 16 on portfolios and the CAPM and omit some of the chapters on statistics, for instance, Chapters 8, 18, and 20 on copulas, GARCH models, and Bayesian statistics. The book could be used for courses at both the master's and Ph.D. levels.

Readers familiar with my textbook *Statistics and Finance: An Introduction* may wonder how that volume differs from this book. This book is at a somewhat more advanced level and has much broader coverage of topics in statistics compared to the earlier book. As the title of this volume suggests, there is more emphasis on data analysis and this book is intended to be more than just "an introduction." Chapters 8, 15, and 20 on copulas, cointegration, and Bayesian statistics are new. Except for some figures borrowed from *Statistics and Finance*, in this book R is used exclusively for computations, data analysis, and graphing, whereas the earlier book used SAS and MATLAB. Nearly all of the examples in this book use data sets that are available in R, so readers can reproduce the results. In Chapter 20 on Bayesian statistics, WinBUGS is used for Markov chain Monte Carlo and is called from R using

the R2WinBUGS package. There is some overlap between the two books, and, in particular, a substantial amount of the material in Chapters 2, 3, 9, 11–13, and 16, has been taken from the earlier book. Unlike *Statistics and Finance*, this volume does not cover options pricing and behavioral finance.

The prerequisites for reading this book are knowledge of calculus, vectors and matrices; probability including stochastic processes; and statistics typical of third- or fourth-year undergraduates in engineering, mathematics, statistics, and related disciplines. There is an appendix that reviews probability and statistics, but it is intended for reference and is certainly not an introduction for readers with little or no prior exposure to these topics. Also, the reader should have some knowledge of computer programming. Some familiarity with the basic ideas of finance is helpful.

This book does not teach R programming, but each chapter has an "R lab" with data analysis and simulations. Students can learn R from these labs and by using R's help or the manual *An Introduction to R* (available at the CRAN website and R's online help) to learn more about the functions used in the labs. Also, the text does indicate which R functions are used in the examples. Occasionally, R code is given to illustrate some process, for example, in Chapter 11 finding the tangency portfolio by quadratic programming. For readers wishing to use R, the bibliographical notes at the end of each chapter mention books that cover R programming and the book's website contains examples of the R and WinBUGS code used to produce this book. Students enter my course *Statistics for Financial Engineering* with quite disparate knowledge of R. Some are very accomplished R programmers, while others have no experience with R, although all have experience with some programming language. Students with no previous experience with R generally need assistance from the instructor to get started on the R labs. Readers using this book for self-study should learn R first before attempting the R labs.

Ithaca, New York David Ruppert
July 2010

http://www.springer.com/978-1-4419-7787-8

Statistics and Data Analysis for Financial Engineering

Ruppert, D.
2011, XXII, 638 p.,
ISBN: 978-1-4419-7787-8