

Čížek • Härdle • Weron
Statistical Tools for Finance and Insurance

Pavel Čížek • Wolfgang Härdle • Rafał Weron

Statistical Tools for Finance and Insurance

 Springer

Pavel Čížek
Tilburg University
Dept. of Econometrics & OR
P.O. Box 90153
5000 LE Tilburg, Netherlands
e-mail: P.Cizek@uvt.nl

Rafał Weron
Wrocław University of Technology
Hugo Steinhaus Center
Wyb. Wyspiańskiego 27
50-370 Wrocław, Poland
e-mail: Rafał.Weron@pwr.wroc.pl

Wolfgang Härdle
Humboldt-Universität zu Berlin
CASE – Center for Applied Statistics and Economics
Institut für Statistik und Ökonometrie
Spandauer Straße 1
10178 Berlin, Germany
e-mail: haerdle@wiwi.hu-berlin.de

This book is also available as e-book on www.i-xplore.de.
Use the licence code at the end of the book to download the e-book.

Library of Congress Control Number: 2005920464

Mathematics Subject Classification (2000): 62P05, 91B26, 91B28

ISBN 3-540-22189-1 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law. Springer is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting by the authors
Production: LE- \TeX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover design: design & production GmbH, Heidelberg
Printed on acid-free paper 46/3142YL – 5 4 3 2 1 0

Contents

Contributors	13
Preface	15
I Finance	19
1 Stable Distributions	21
<i>Szymon Borak, Wolfgang Härdle, and Rafał Weron</i>	
1.1 Introduction	21
1.2 Definitions and Basic Characteristic	22
1.2.1 Characteristic Function Representation	24
1.2.2 Stable Density and Distribution Functions	26
1.3 Simulation of α -stable Variables	28
1.4 Estimation of Parameters	30
1.4.1 Tail Exponent Estimation	31
1.4.2 Quantile Estimation	33
1.4.3 Characteristic Function Approaches	34
1.4.4 Maximum Likelihood Method	35
1.5 Financial Applications of Stable Laws	36

2 Extreme Value Analysis and Copulas 45

Krzysztof Jajuga and Daniel Papla

2.1	Introduction	45
2.1.1	Analysis of Distribution of the Extremum	46
2.1.2	Analysis of Conditional Excess Distribution	47
2.1.3	Examples	48
2.2	Multivariate Time Series	53
2.2.1	Copula Approach	53
2.2.2	Examples	56
2.2.3	Multivariate Extreme Value Approach	57
2.2.4	Examples	60
2.2.5	Copula Analysis for Multivariate Time Series	61
2.2.6	Examples	62

3 Tail Dependence 65

Rafael Schmidt

3.1	Introduction	65
3.2	What is Tail Dependence?	66
3.3	Calculation of the Tail-dependence Coefficient	69
3.3.1	Archimedean Copulae	69
3.3.2	Elliptically-contoured Distributions	70
3.3.3	Other Copulae	74
3.4	Estimating the Tail-dependence Coefficient	75
3.5	Comparison of TDC Estimators	78
3.6	Tail Dependence of Asset and FX Returns	81
3.7	Value at Risk – a Simulation Study	84

4 Pricing of Catastrophe Bonds **93**

Krzysztof Burnecki, Grzegorz Kukla, and David Taylor

4.1 Introduction	93
4.1.1 The Emergence of CAT Bonds	94
4.1.2 Insurance Securitization	96
4.1.3 CAT Bond Pricing Methodology	97
4.2 Compound Doubly Stochastic Poisson Pricing Model	99
4.3 Calibration of the Pricing Model	100
4.4 Dynamics of the CAT Bond Price	104

5 Common Functional IV Analysis **115**

Michal Benko and Wolfgang Härdle

5.1 Introduction	115
5.2 Implied Volatility Surface	116
5.3 Functional Data Analysis	118
5.4 Functional Principal Components	121
5.4.1 Basis Expansion	123
5.5 Smoothed Principal Components Analysis	125
5.5.1 Basis Expansion	126
5.6 Common Principal Components Model	127

6 Implied Trinomial Trees **135**

Pavel Čížek and Karel Komorád

6.1 Option Pricing	136
6.2 Trees and Implied Trees	138
6.3 Implied Trinomial Trees	140
6.3.1 Basic Insight	140

6.3.2	State Space	142
6.3.3	Transition Probabilities	144
6.3.4	Possible Pitfalls	145
6.4	Examples	147
6.4.1	Pre-specified Implied Volatility	147
6.4.2	German Stock Index	152
7	Heston's Model and the Smile	161
	<i>Rafał Weron and Uwe Wystup</i>	
7.1	Introduction	161
7.2	Heston's Model	163
7.3	Option Pricing	166
7.3.1	Greeks	168
7.4	Calibration	169
7.4.1	Qualitative Effects of Changing Parameters	171
7.4.2	Calibration Results	173
8	FFT-based Option Pricing	183
	<i>Szymon Borak, Kai Detlefsen, and Wolfgang Härdle</i>	
8.1	Introduction	183
8.2	Modern Pricing Models	183
8.2.1	Merton Model	184
8.2.2	Heston Model	185
8.2.3	Bates Model	187
8.3	Option Pricing with FFT	188
8.4	Applications	192

9 Valuation of Mortgage Backed Securities 201

Nicolas Gaussel and Julien Tamine

9.1	Introduction	201
9.2	Optimally Prepaid Mortgage	204
9.2.1	Financial Characteristics and Cash Flow Analysis	204
9.2.2	Optimal Behavior and Price	204
9.3	Valuation of Mortgage Backed Securities	212
9.3.1	Generic Framework	213
9.3.2	A Parametric Specification of the Prepayment Rate	215
9.3.3	Sensitivity Analysis	218

10 Predicting Bankruptcy with Support Vector Machines 225

Wolfgang Härdle, Rouslan Moro, and Dorothea Schäfer

10.1	Bankruptcy Analysis Methodology	226
10.2	Importance of Risk Classification in Practice	230
10.3	Lagrangian Formulation of the SVM	233
10.4	Description of Data	236
10.5	Computational Results	237
10.6	Conclusions	243

11 Modelling Indonesian Money Demand 249

Noer Azam Achsani, Oliver Holtemöller, and Hizir Sofyan

11.1	Specification of Money Demand Functions	250
11.2	The Econometric Approach to Money Demand	253
11.2.1	Econometric Estimation of Money Demand Functions	253
11.2.2	Econometric Modelling of Indonesian Money Demand	254
11.3	The Fuzzy Approach to Money Demand	260

11.3.1	Fuzzy Clustering	260
11.3.2	The Takagi-Sugeno Approach	261
11.3.3	Model Identification	262
11.3.4	Fuzzy Modelling of Indonesian Money Demand	263
11.4	Conclusions	266
12	Nonparametric Productivity Analysis	271
	<i>Wolfgang Härdle and Seok-Oh Jeong</i>	
12.1	The Basic Concepts	272
12.2	Nonparametric Hull Methods	276
12.2.1	Data Envelopment Analysis	277
12.2.2	Free Disposal Hull	278
12.3	DEA in Practice: Insurance Agencies	279
12.4	FDH in Practice: Manufacturing Industry	281
II	Insurance	287
13	Loss Distributions	289
	<i>Krzysztof Burnecki, Adam Misiorek, and Rafał Weron</i>	
13.1	Introduction	289
13.2	Empirical Distribution Function	290
13.3	Analytical Methods	292
13.3.1	Log-normal Distribution	292
13.3.2	Exponential Distribution	293
13.3.3	Pareto Distribution	295
13.3.4	Burr Distribution	298
13.3.5	Weibull Distribution	298

13.3.6	Gamma Distribution	300
13.3.7	Mixture of Exponential Distributions	302
13.4	Statistical Validation Techniques	303
13.4.1	Mean Excess Function	303
13.4.2	Tests Based on the Empirical Distribution Function	305
13.4.3	Limited Expected Value Function	309
13.5	Applications	311

14 Modeling of the Risk Process 319

Krzysztof Burnecki and Rafał Weron

14.1	Introduction	319
14.2	Claim Arrival Processes	321
14.2.1	Homogeneous Poisson Process	321
14.2.2	Non-homogeneous Poisson Process	323
14.2.3	Mixed Poisson Process	326
14.2.4	Cox Process	327
14.2.5	Renewal Process	328
14.3	Simulation of Risk Processes	329
14.3.1	Catastrophic Losses	329
14.3.2	Danish Fire Losses	334

15 Ruin Probabilities in Finite and Infinite Time 341

Krzysztof Burnecki, Paweł Miśta, and Aleksander Weron

15.1	Introduction	341
15.1.1	Light- and Heavy-tailed Distributions	343
15.2	Exact Ruin Probabilities in Infinite Time	346
15.2.1	No Initial Capital	347

15.2.2	Exponential Claim Amounts	347
15.2.3	Gamma Claim Amounts	347
15.2.4	Mixture of Two Exponentials Claim Amounts	349
15.3	Approximations of the Ruin Probability in Infinite Time	350
15.3.1	Cramér–Lundberg Approximation	351
15.3.2	Exponential Approximation	352
15.3.3	Lundberg Approximation	352
15.3.4	Beekman–Bowers Approximation	353
15.3.5	Renyi Approximation	354
15.3.6	De Vylder Approximation	355
15.3.7	4-moment Gamma De Vylder Approximation	356
15.3.8	Heavy Traffic Approximation	358
15.3.9	Light Traffic Approximation	359
15.3.10	Heavy-light Traffic Approximation	360
15.3.11	Subexponential Approximation	360
15.3.12	Computer Approximation via the Pollaczek-Khinchin Formula	361
15.3.13	Summary of the Approximations	362
15.4	Numerical Comparison of the Infinite Time Approximations	363
15.5	Exact Ruin Probabilities in Finite Time	367
15.5.1	Exponential Claim Amounts	368
15.6	Approximations of the Ruin Probability in Finite Time	368
15.6.1	Monte Carlo Method	369
15.6.2	Segerdahl Normal Approximation	369
15.6.3	Diffusion Approximation	371
15.6.4	Corrected Diffusion Approximation	372
15.6.5	Finite Time De Vylder Approximation	373

15.6.6 Summary of the Approximations 374
 15.7 Numerical Comparison of the Finite Time Approximations . . . 374

16 Stable Diffusion Approximation of the Risk Process 381

Hansjörg Furrer, Zbigniew Michna, and Aleksander Weron

16.1 Introduction 381
 16.2 Brownian Motion and the Risk Model for Small Claims 382
 16.2.1 Weak Convergence of Risk Processes to Brownian Motion 383
 16.2.2 Ruin Probability for the Limit Process 383
 16.2.3 Examples 384
 16.3 Stable Lévy Motion and the Risk Model for Large Claims 386
 16.3.1 Weak Convergence of Risk Processes to α -stable Lévy
 Motion 387
 16.3.2 Ruin Probability in Limit Risk Model for Large Claims 388
 16.3.3 Examples 390

17 Risk Model of Good and Bad Periods 395

Zbigniew Michna

17.1 Introduction 395
 17.2 Fractional Brownian Motion and Model of Good and Bad Periods 396
 17.3 Ruin Probability in Limit Risk Model of Good and Bad Periods 399
 17.4 Examples 402

18 Premiums in the Individual and Collective Risk Models 407

Jan Iwanik and Joanna Nowicka-Zagrajek

18.1 Premium Calculation Principles 408
 18.2 Individual Risk Model 410
 18.2.1 General Premium Formulae 411

18.2.2	Premiums in the Case of the Normal Approximation . . .	412
18.2.3	Examples	413
18.3	Collective Risk Model	416
18.3.1	General Premium Formulae	417
18.3.2	Premiums in the Case of the Normal and Translated Gamma Approximations	418
18.3.3	Compound Poisson Distribution	420
18.3.4	Compound Negative Binomial Distribution	421
18.3.5	Examples	423

19 Pure Risk Premiums under Deductibles 427

Krzysztof Burnecki, Joanna Nowicka-Zagrajek, and Agnieszka Wylomańska

19.1	Introduction	427
19.2	General Formulae for Premiums Under Deductibles	428
19.2.1	Franchise Deductible	429
19.2.2	Fixed Amount Deductible	431
19.2.3	Proportional Deductible	432
19.2.4	Limited Proportional Deductible	432
19.2.5	Disappearing Deductible	434
19.3	Premiums Under Deductibles for Given Loss Distributions . . .	436
19.3.1	Log-normal Loss Distribution	437
19.3.2	Pareto Loss Distribution	438
19.3.3	Burr Loss Distribution	441
19.3.4	Weibull Loss Distribution	445
19.3.5	Gamma Loss Distribution	447
19.3.6	Mixture of Two Exponentials Loss Distribution	449
19.4	Final Remarks	450

20 Premiums, Investments, and Reinsurance 453

Paweł Miśta and Wojciech Otto

20.1	Introduction	453
20.2	Single-Period Criterion and the Rate of Return on Capital . . .	456
20.2.1	Risk Based Capital Concept	456
20.2.2	How To Choose Parameter Values?	457
20.3	The Top-down Approach to Individual Risks Pricing	459
20.3.1	Approximations of Quantiles	459
20.3.2	Marginal Cost Basis for Individual Risk Pricing	460
20.3.3	Balancing Problem	461
20.3.4	A Solution for the Balancing Problem	462
20.3.5	Applications	462
20.4	Rate of Return and Reinsurance Under the Short Term Criterion	463
20.4.1	General Considerations	464
20.4.2	Illustrative Example	465
20.4.3	Interpretation of Numerical Calculations in Example 2 .	467
20.5	Ruin Probability Criterion when the Initial Capital is Given . .	469
20.5.1	Approximation Based on Lundberg Inequality	469
20.5.2	“Zero” Approximation	471
20.5.3	Cramér–Lundberg Approximation	471
20.5.4	Beekman–Bowers Approximation	472
20.5.5	Diffusion Approximation	473
20.5.6	De Vylder Approximation	474
20.5.7	Subexponential Approximation	475
20.5.8	Panjer Approximation	475
20.6	Ruin Probability Criterion and the Rate of Return	477
20.6.1	Fixed Dividends	477

20.6.2 Flexible Dividends	479
20.7 Ruin Probability, Rate of Return and Reinsurance	481
20.7.1 Fixed Dividends	481
20.7.2 Interpretation of Solutions Obtained in Example 5	482
20.7.3 Flexible Dividends	484
20.7.4 Interpretation of Solutions Obtained in Example 6	485
20.8 Final remarks	487
III General	489
21 Working with the XQC	491
<i>Szymon Borak, Wolfgang Härdle, and Heiko Lehmann</i>	
21.1 Introduction	491
21.2 The XploRe Quantlet Client	492
21.2.1 Configuration	492
21.2.2 Getting Connected	493
21.3 Desktop	494
21.3.1 XploRe Quantlet Editor	495
21.3.2 Data Editor	496
21.3.3 Method Tree	501
21.3.4 Graphical Output	503
Index	507

Contributors

Noer Azam Achsani Department of Economics, University of Potsdam

Michal Benko Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin

Szymon Borak Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin

Krzysztof Burnecki Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology

Pavel Čížek Center for Economic Research, Tilburg University

Kai Detlefsen Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin

Hansjörg Furrer Swiss Life, Zürich

Nicolas Gaussel Société Générale Asset Management, Paris

Wolfgang Härdle Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin

Oliver Holtemöller Department of Economics, RWTH Aachen University

Jan Iwanik Concordia Capital S.A., Poznań

Krzysztof Jajuga Department of Financial Investments and Insurance, Wrocław University of Economics

Seok-Oh Jeong Institut de statistique, Université catholique de Louvain

Karel Komorád Komerční Banka, Praha

Grzegorz Kukla Towarzystwo Ubezpieczeniowe EUROPA S.A., Wrocław

Heiko Lehmann SAP AG, Walldorf

Zbigniew Michna Department of Mathematics, Wrocław University of Economics

Adam Misiorek Institute of Power Systems Automation, Wrocław

Paweł Miśta Institute of Mathematics, Wrocław University of Technology

Rouslan Moro Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin

Joanna Nowicka-Zagrajek Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology

Wojciech Otto Faculty of Economic Sciences, Warsaw University

Daniel Papla Department of Financial Investments and Insurance, Wrocław University of Economics

Dorothea Schäfer Deutsches Institut für Wirtschaftsforschung e.V., Berlin

Rafael Schmidt Department of Statistics, London School of Economics

Hizir Sofyan Mathematics Department, Syiah Kuala University

Julien Tamine Société Générale Asset Management, Paris

David Taylor School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg

Aleksander Weron Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology

Rafał Weron Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology

Agnieszka Wyłomańska Institute of Mathematics, Wrocław University of Technology

Uwe Wystup MathFinance AG, Waldems

Preface

This book is designed for students, researchers and practitioners who want to be introduced to modern statistical tools applied in finance and insurance. It is the result of a joint effort of the Center for Economic Research (CentER), Center for Applied Statistics and Economics (C.A.S.E.) and Hugo Steinhaus Center for Stochastic Methods (HSC). All three institutions brought in their specific profiles and created with this book a wide-angle view on and solutions to up-to-date practical problems.

The text is comprehensible for a graduate student in financial engineering as well as for an inexperienced newcomer to quantitative finance and insurance who wants to get a grip on advanced statistical tools applied in these fields. An experienced reader with a bright knowledge of financial and actuarial mathematics will probably skip some sections but will hopefully enjoy the various computational tools. Finally, a practitioner might be familiar with some of the methods. However, the statistical techniques related to modern financial products, like MBS or CAT bonds, will certainly attract him.

“Statistical Tools for Finance and Insurance” consists naturally of two main parts. Each part contains chapters with high focus on practical applications. The book starts with an introduction to *stable distributions*, which are the standard model for heavy tailed phenomena. Their numerical implementation is thoroughly discussed and applications to finance are given. The second chapter presents the ideas of *extreme value and copula analysis* as applied to multivariate financial data. This topic is extended in the subsequent chapter which deals with *tail dependence*, a concept describing the limiting proportion that one margin exceeds a certain threshold given that the other margin has already exceeded that threshold. The fourth chapter reviews the market in *catastrophe insurance* risk, which emerged in order to facilitate the direct transfer of reinsurance risk associated with natural catastrophes from corporations, insurers, and reinsurers to capital market investors. The next contribution employs *functional data analysis* for the estimation of smooth implied volatility sur-

faces. These surfaces are a result of using an oversimplified market benchmark model – the Black-Scholes formula – to real data. An attractive approach to overcome this problem is discussed in chapter six, where *implied trinomial trees* are applied to modeling implied volatilities and the corresponding state-price densities. An alternative route to tackling the implied volatility smile has led researchers to develop stochastic volatility models. The relative simplicity and the direct link of model parameters to the market makes *Heston's model* very attractive to front office users. Its application to FX option markets is covered in chapter seven. The following chapter shows how the computational complexity of stochastic volatility models can be overcome with the help of the *Fast Fourier Transform*. In chapter nine the valuation of *Mortgage Backed Securities* is discussed. The optimal prepayment policy is obtained via optimal stopping techniques. It is followed by a very innovative topic of predicting corporate bankruptcy with *Support Vector Machines*. Chapter eleven presents a novel approach to *money-demand modeling* using fuzzy clustering techniques. The first part of the book closes with *productivity analysis* for cost and frontier estimation. The nonparametric Data Envelopment Analysis is applied to efficiency issues of insurance agencies.

The insurance part of the book starts with a chapter on *loss distributions*. The basic models for claim severities are introduced and their statistical properties are thoroughly explained. In chapter fourteen, the methods of simulating and visualizing the *risk process* are discussed. This topic is followed by an overview of the approaches to *approximating the ruin probability* of an insurer. Both finite and infinite time approximations are presented. Some of these methods are extended in chapters sixteen and seventeen, where classical and anomalous *diffusion approximations* to ruin probability are discussed and extended to cases when the risk process exhibits *good and bad periods*. The last three chapters are related to one of the most important aspects of the insurance business – *premium calculation*. Chapter eighteen introduces the basic concepts including the pure risk premium and various safety loadings under different loss distributions. Calculation of a joint premium for a portfolio of insurance policies in the individual and collective risk models is discussed as well. The inclusion of *deductibles* into premium calculation is the topic of the following contribution. The last chapter of the insurance part deals with setting the appropriate level of insurance premium within a broader context of business decisions, including risk transfer through *reinsurance* and the rate of return on capital required to ensure solvability.

Our e-book offers a complete PDF version of this text and the corresponding HTML files with links to algorithms and quantlets. The reader of this book

may therefore easily reconfigure and recalculate all the presented examples and methods via the enclosed XploRe Quantlet Server (XQS), which is also available from www.xploRe-stat.de and www.quantlet.com. A tutorial chapter explaining how to setup and use XQS can be found in the third and final part of the book.

We gratefully acknowledge the support of Deutsche Forschungsgemeinschaft (SFB 373 Quantifikation und Simulation Ökonomischer Prozesse, SFB 649 Ökonomisches Risiko) and Komitet Badań Naukowych (PBZ-KBN 016/P03/99 Mathematical models in analysis of financial instruments and markets in Poland). A book of this kind would not have been possible without the help of many friends, colleagues, and students. For the technical production of the e-book platform and quantlets we would like to thank Zdeněk Hlávka, Sigbert Klinke, Heiko Lehmann, Adam Misiorek, Piotr Uniejewski, Qingwei Wang, and Rodrigo Witzel. Special thanks for careful proofreading and supervision of the insurance part go to Krzysztof Burnecki.

Pavel Čížek, Wolfgang Härdle, and Rafał Weron

Tilburg, Berlin, and Wrocław, February 2005

Part I

Finance

1 Stable Distributions

Szymon Borak, Wolfgang Härdle, and Rafał Weron

1.1 Introduction

Many of the concepts in theoretical and empirical finance developed over the past decades – including the classical portfolio theory, the Black-Scholes-Merton option pricing model and the RiskMetrics variance-covariance approach to Value at Risk (VaR) – rest upon the assumption that asset returns follow a normal distribution. However, it has been long known that asset returns are not normally distributed. Rather, the empirical observations exhibit fat tails. This heavy tailed or leptokurtic character of the distribution of price changes has been repeatedly observed in various markets and may be quantitatively measured by the kurtosis in excess of 3, a value obtained for the normal distribution (Bouchaud and Potters, 2000; Carr et al., 2002; Guillaume et al., 1997; Mantegna and Stanley, 1995; Rachev, 2003; Weron, 2004).

It is often argued that financial asset returns are the cumulative outcome of a vast number of pieces of information and individual decisions arriving almost continuously in time (McCulloch, 1996; Rachev and Mittnik, 2000). As such, since the pioneering work of Louis Bachelier in 1900, they have been modeled by the Gaussian distribution. The strongest statistical argument for it is based on the Central Limit Theorem, which states that the sum of a large number of independent, identically distributed variables from a finite-variance distribution will tend to be normally distributed. However, as we have already mentioned, financial asset returns usually have heavier tails.

In response to the empirical evidence Mandelbrot (1963) and Fama (1965) proposed the stable distribution as an alternative model. Although there are other heavy-tailed alternatives to the Gaussian law – like Student's t , hyperbolic, normal inverse Gaussian, or truncated stable – there is at least one good reason

for modeling financial variables using stable distributions. Namely, they are supported by the generalized Central Limit Theorem, which states that stable laws are the only possible limit distributions for properly normalized and centered sums of independent, identically distributed random variables.

Since stable distributions can accommodate the fat tails and asymmetry, they often give a very good fit to empirical data. In particular, they are valuable models for data sets covering extreme events, like market crashes or natural catastrophes. Even though they are not universal, they are a useful tool in the hands of an analyst working in finance or insurance. Hence, we devote this chapter to a thorough presentation of the computational aspects related to stable laws. In Section 1.2 we review the analytical concepts and basic characteristics. In the following two sections we discuss practical simulation and estimation approaches. Finally, in Section 1.5 we present financial applications of stable laws.

1.2 Definitions and Basic Characteristics

Stable laws – also called α -stable, stable Paretian or Lévy stable – were introduced by Levy (1925) during his investigations of the behavior of sums of independent random variables. A sum of two independent random variables having an α -stable distribution with index α is again α -stable with the same index α . This invariance property, however, does not hold for different α 's.

The α -stable distribution requires four parameters for complete description: an index of stability $\alpha \in (0, 2]$ also called the tail index, tail exponent or characteristic exponent, a skewness parameter $\beta \in [-1, 1]$, a scale parameter $\sigma > 0$ and a location parameter $\mu \in \mathbb{R}$. The tail exponent α determines the rate at which the tails of the distribution taper off, see the left panel in Figure 1.1. When $\alpha = 2$, the Gaussian distribution results. When $\alpha < 2$, the variance is infinite and the tails are asymptotically equivalent to a Pareto law, i.e. they exhibit a power-law behavior. More precisely, using a central limit theorem type argument it can be shown that (Janicki and Weron, 1994; Samorodnitsky and Taqqu, 1994):

$$\begin{cases} \lim_{x \rightarrow \infty} x^\alpha \mathbb{P}(X > x) = C_\alpha(1 + \beta)\sigma^\alpha, \\ \lim_{x \rightarrow \infty} x^\alpha \mathbb{P}(X < -x) = C_\alpha(1 - \beta)\sigma^\alpha, \end{cases} \quad (1.1)$$

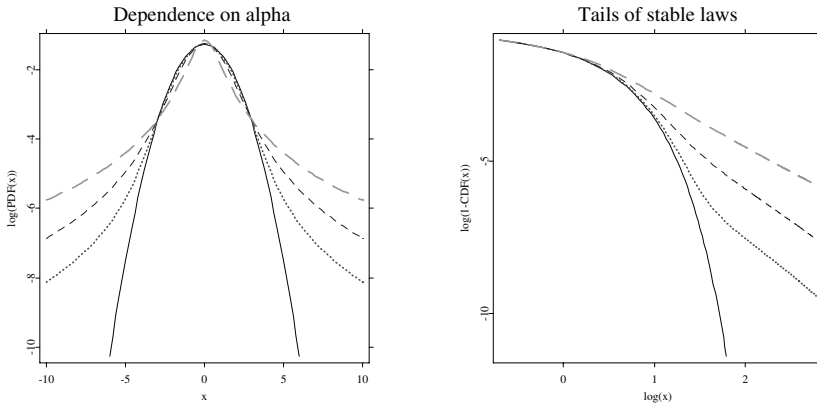



Figure 1.1: *Left panel:* A semilog plot of symmetric ($\beta = \mu = 0$) α -stable probability density functions (pdfs) for $\alpha = 2$ (black solid line), 1.8 (red dotted line), 1.5 (blue dashed line) and 1 (green long-dashed line). The Gaussian ($\alpha = 2$) density forms a parabola and is the only α -stable density with exponential tails. *Right panel:* Right tails of symmetric α -stable cumulative distribution functions (cdfs) for $\alpha = 2$ (black solid line), 1.95 (red dotted line), 1.8 (blue dashed line) and 1.5 (green long-dashed line) on a double logarithmic paper. For $\alpha < 2$ the tails form straight lines with slope $-\alpha$.

 STFstab01.xpl

where:

$$C_\alpha = \left(2 \int_0^\infty x^{-\alpha} \sin(x) dx \right)^{-1} = \frac{1}{\pi} \Gamma(\alpha) \sin \frac{\pi\alpha}{2}.$$

The convergence to a power-law tail varies for different α 's and, as can be seen in the right panel of Figure 1.1, is slower for larger values of the tail index. Moreover, the tails of α -stable distribution functions exhibit a crossover from an approximate power decay with exponent $\alpha > 2$ to the true tail with exponent α . This phenomenon is more visible for large α 's (Weron, 2001).

When $\alpha > 1$, the mean of the distribution exists and is equal to μ . In general, the p th moment of a stable random variable is finite if and only if $p < \alpha$. When the skewness parameter β is positive, the distribution is skewed to the right,

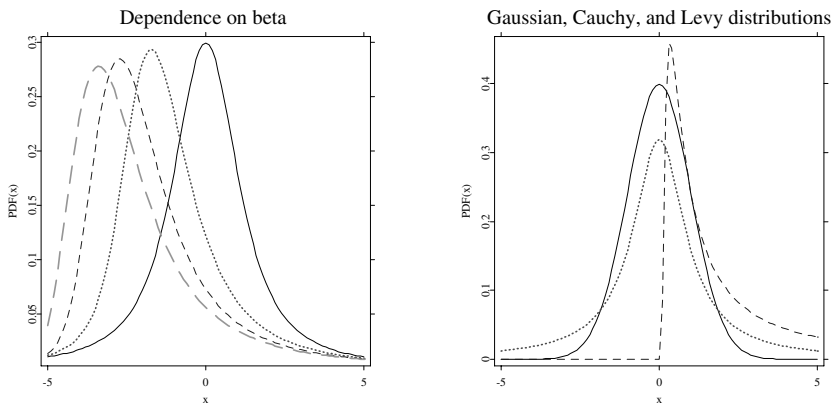



Figure 1.2: *Left panel*: Stable pdfs for $\alpha = 1.2$ and $\beta = 0$ (black solid line), 0.5 (red dotted line), 0.8 (blue dashed line) and 1 (green long-dashed line). *Right panel*: Closed form formulas for densities are known only for three distributions – Gaussian ($\alpha = 2$; black solid line), Cauchy ($\alpha = 1$; red dotted line) and Levy ($\alpha = 0.5, \beta = 1$; blue dashed line). The latter is a totally skewed distribution, i.e. its support is \mathbb{R}_+ . In general, for $\alpha < 1$ and $\beta = 1$ (-1) the distribution is totally skewed to the right (left).

 STFstab02.xpl

i.e. the right tail is thicker, see the left panel of Figure 1.2. When it is negative, it is skewed to the left. When $\beta = 0$, the distribution is symmetric about μ . As α approaches 2, β loses its effect and the distribution approaches the Gaussian distribution regardless of β . The last two parameters, σ and μ , are the usual scale and location parameters, i.e. σ determines the width and μ the shift of the mode (the peak) of the density. For $\sigma = 1$ and $\mu = 0$ the distribution is called standard stable.

1.2.1 Characteristic Function Representation

Due to the lack of closed form formulas for densities for all but three distributions (see the right panel in Figure 1.2), the α -stable law can be most

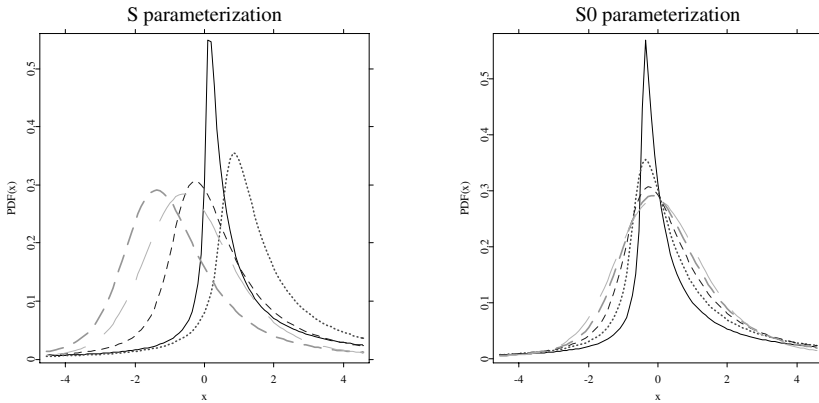



Figure 1.3: Comparison of S and S^0 parameterizations: α -stable pdfs for $\beta = 0.5$ and $\alpha = 0.5$ (black solid line), 0.75 (red dotted line), 1 (blue short-dashed line), 1.25 (green dashed line) and 1.5 (cyan long-dashed line).

 STFstab03.xpl

conveniently described by its characteristic function $\phi(t)$ – the inverse Fourier transform of the probability density function. However, there are multiple parameterizations for α -stable laws and much confusion has been caused by these different representations, see Figure 1.3. The variety of formulas is caused by a combination of historical evolution and the numerous problems that have been analyzed using specialized forms of the stable distributions. The most popular parameterization of the characteristic function of $X \sim S_\alpha(\sigma, \beta, \mu)$, i.e. an α -stable random variable with parameters α , σ , β , and μ , is given by (Samorodnitsky and Taqqu, 1994; Weron, 2004):

$$\ln \phi(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \left\{ 1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2} \right\} + i\mu t, & \alpha \neq 1, \\ -\sigma |t| \left\{ 1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \ln |t| \right\} + i\mu t, & \alpha = 1. \end{cases} \quad (1.2)$$

For numerical purposes, it is often advisable to use Nolan's (1997) parameterization:

$$\ln \phi_0(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \{1 + i\beta \text{sign}(t) \tan \frac{\pi\alpha}{2} [(\sigma|t|)^{1-\alpha} - 1]\} + i\mu_0 t, & \alpha \neq 1, \\ -\sigma |t| \{1 + i\beta \text{sign}(t) \frac{2}{\pi} \ln(\sigma|t|)\} + i\mu_0 t, & \alpha = 1. \end{cases} \quad (1.3)$$

The $S_\alpha^0(\sigma, \beta, \mu_0)$ parameterization is a variant of Zolotariev's (M)-parameterization (Zolotarev, 1986), with the characteristic function and hence the density and the distribution function jointly continuous in all four parameters, see the right panel in Figure 1.3. In particular, percentiles and convergence to the power-law tail vary in a continuous way as α and β vary. The location parameters of the two representations are related by $\mu = \mu_0 - \beta\sigma \tan \frac{\pi\alpha}{2}$ for $\alpha \neq 1$ and $\mu = \mu_0 - \beta\sigma \frac{2}{\pi} \ln \sigma$ for $\alpha = 1$. Note also, that the traditional scale parameter σ_G of the Gaussian distribution defined by:

$$f_G(x) = \frac{1}{\sqrt{2\pi}\sigma_G} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma_G^2} \right\}, \quad (1.4)$$

is not the same as σ in formulas (1.2) or (1.3). Namely, $\sigma_G = \sqrt{2}\sigma$.

1.2.2 Stable Density and Distribution Functions

The lack of closed form formulas for most stable densities and distribution functions has negative consequences. For example, during maximum likelihood estimation computationally burdensome numerical approximations have to be used. There generally are two approaches to this problem. Either the fast Fourier transform (FFT) has to be applied to the characteristic function (Mittnik, Doganoglu, and Chenyao, 1999) or direct numerical integration has to be utilized (Nolan, 1997, 1999).

For data points falling between the equally spaced FFT grid nodes an interpolation technique has to be used. Taking a larger number of grid points increases accuracy, however, at the expense of higher computational burden. The FFT based approach is faster for large samples, whereas the direct integration method favors small data sets since it can be computed at any arbitrarily chosen point. Mittnik, Doganoglu, and Chenyao (1999) report that for $N = 2^{13}$ the FFT based method is faster for samples exceeding 100 observations and slower for smaller data sets. Moreover, the FFT based approach is less universal – it is efficient only for large α 's and only for pdf calculations. When

computing the cdf the density must be numerically integrated. In contrast, in the direct integration method Zolotarev's (1986) formulas either for the density or the distribution function are numerically integrated.

Set $\zeta = -\beta \tan \frac{\pi\alpha}{2}$. Then the density $f(x; \alpha, \beta)$ of a standard α -stable random variable in representation S^0 , i.e. $X \sim S^0_\alpha(1, \beta, 0)$, can be expressed as (note, that Zolotarev (1986, Section 2.2) used yet another parametrization):

- when $\alpha \neq 1$ and $x > \zeta$:

$$f(x; \alpha, \beta) = \frac{\alpha(x - \zeta)^{\frac{1}{\alpha-1}}}{\pi |\alpha - 1|} \int_{-\xi}^{\frac{\pi}{2}} V(\theta; \alpha, \beta) \exp \left\{ -(x - \zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta) \right\} d\theta, \quad (1.5)$$

- when $\alpha \neq 1$ and $x = \zeta$:

$$f(x; \alpha, \beta) = \frac{\Gamma(1 + \frac{1}{\alpha}) \cos(\xi)}{\pi(1 + \zeta^2)^{\frac{1}{2\alpha}}},$$

- when $\alpha \neq 1$ and $x < \zeta$:

$$f(x; \alpha, \beta) = f(-x; \alpha, -\beta),$$

- when $\alpha = 1$:

$$f(x; 1, \beta) = \begin{cases} \frac{1}{2|\beta|} e^{-\frac{\pi x}{2\beta}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta; 1, \beta) \exp \left\{ -e^{-\frac{\pi x}{2\beta}} V(\theta; 1, \beta) \right\} d\theta, & \beta \neq 0, \\ \frac{1}{\pi(1+x^2)}, & \beta = 0, \end{cases}$$

where

$$\xi = \begin{cases} \frac{1}{\alpha} \arctan(-\zeta), & \alpha \neq 1, \\ \frac{\pi}{2}, & \alpha = 1, \end{cases}$$

and

$$V(\theta; \alpha, \beta) = \begin{cases} (\cos \alpha \xi)^{\frac{1}{\alpha-1}} \left(\frac{\cos \theta}{\sin \alpha(\xi+\theta)} \right)^{\frac{\alpha}{\alpha-1}} \frac{\cos\{\alpha\xi + (\alpha-1)\theta\}}{\cos \theta}, & \alpha \neq 1, \\ \frac{2}{\pi} \left(\frac{\frac{\pi}{2} + \beta\theta}{\cos \theta} \right) \exp \left\{ \frac{1}{\beta} \left(\frac{\pi}{2} + \beta\theta \right) \tan \theta \right\}, & \alpha = 1, \beta \neq 0. \end{cases}$$

The distribution $F(x; \alpha, \beta)$ of a standard α -stable random variable in representation S^0 can be expressed as:

- when $\alpha \neq 1$ and $x > \zeta$:

$$F(x; \alpha, \beta) = c_1(\alpha, \beta) + \frac{\text{sign}(1 - \alpha)}{\pi} \int_{-\xi}^{\frac{\pi}{2}} \exp\{-(x - \zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta)\} d\theta,$$

where

$$c_1(\alpha, \beta) = \begin{cases} \frac{1}{\pi} \left(\frac{\pi}{2} - \xi\right), & \alpha < 1, \\ 1, & \alpha > 1, \end{cases}$$

- when $\alpha \neq 1$ and $x = \zeta$:

$$F(x; \alpha, \beta) = \frac{1}{\pi} \left(\frac{\pi}{2} - \xi\right),$$

- when $\alpha \neq 1$ and $x < \zeta$:

$$F(x; \alpha, \beta) = 1 - F(-x; \alpha, -\beta),$$

- when $\alpha = 1$:

$$F(x; 1, \beta) = \begin{cases} \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\left\{-e^{-\frac{\pi x}{2\beta}} V(\theta; 1, \beta)\right\} d\theta, & \beta > 0, \\ \frac{1}{2} + \frac{1}{\pi} \arctan x, & \beta = 0, \\ 1 - F(x, 1, -\beta), & \beta < 0. \end{cases}$$

Formula (1.5) requires numerical integration of the function $g(\cdot) \exp\{-g(\cdot)\}$, where $g(\theta; x, \alpha, \beta) = (x - \zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta)$. The integrand is 0 at $-\xi$, increases monotonically to a maximum of $\frac{1}{e}$ at point θ^* for which $g(\theta^*; x, \alpha, \beta) = 1$, and then decreases monotonically to 0 at $\frac{\pi}{2}$ (Nolan, 1997). However, in some cases the integrand becomes very peaked and numerical algorithms can miss the spike and underestimate the integral. To avoid this problem we need to find the argument θ^* of the peak numerically and compute the integral as a sum of two integrals: one from $-\xi$ to θ^* and the other from θ^* to $\frac{\pi}{2}$.

1.3 Simulation of α -stable Variables

The complexity of the problem of simulating sequences of α -stable random variables results from the fact that there are no analytic expressions for the

inverse F^{-1} of the cumulative distribution function. The first breakthrough was made by Kanter (1975), who gave a direct method for simulating $S_\alpha(1, 1, 0)$ random variables, for $\alpha < 1$. It turned out that this method could be easily adapted to the general case. Chambers, Mallows, and Stuck (1976) were the first to give the formulas.

The algorithm for constructing a standard stable random variable $X \sim S_\alpha(1, \beta, 0)$, in representation (1.2), is the following (Weron, 1996):

- generate a random variable V uniformly distributed on $(-\frac{\pi}{2}, \frac{\pi}{2})$ and an independent exponential random variable W with mean 1;
- for $\alpha \neq 1$ compute:

$$X = S_{\alpha,\beta} \cdot \frac{\sin\{\alpha(V + B_{\alpha,\beta})\}}{\{\cos(V)\}^{1/\alpha}} \cdot \left[\frac{\cos\{V - \alpha(V + B_{\alpha,\beta})\}}{W} \right]^{(1-\alpha)/\alpha}, \quad (1.6)$$

where

$$\begin{aligned} B_{\alpha,\beta} &= \frac{\arctan(\beta \tan \frac{\pi\alpha}{2})}{\alpha}, \\ S_{\alpha,\beta} &= \left\{ 1 + \beta^2 \tan^2 \left(\frac{\pi\alpha}{2} \right) \right\}^{1/(2\alpha)}; \end{aligned}$$

- for $\alpha = 1$ compute:

$$X = \frac{2}{\pi} \left\{ \left(\frac{\pi}{2} + \beta V \right) \tan V - \beta \ln \left(\frac{\frac{\pi}{2} W \cos V}{\frac{\pi}{2} + \beta V} \right) \right\}. \quad (1.7)$$

Given the formulas for simulation of a standard α -stable random variable, we can easily simulate a stable random variable for all admissible values of the parameters α , σ , β and μ using the following property: if $X \sim S_\alpha(1, \beta, 0)$ then

$$Y = \begin{cases} \sigma X + \mu, & \alpha \neq 1, \\ \sigma X + \frac{2}{\pi} \beta \sigma \ln \sigma + \mu, & \alpha = 1, \end{cases} \quad (1.8)$$

is $S_\alpha(\sigma, \beta, \mu)$. It is interesting to note that for $\alpha = 2$ (and $\beta = 0$) the Chambers-Mallows-Stuck method reduces to the well known Box-Muller algorithm for generating Gaussian random variables (Janicki and Weron, 1994). Although many other approaches have been proposed in the literature, this method is regarded as the fastest and the most accurate (Weron, 2004).