

Springer Series in Operations Research and Financial Engineering

Series Editors:

Thomas V. Mikosch

Sidney I. Resnick

Stephen M. Robinson

Springer Series in Operations Research and Financial Engineering

- Altiok*: Performance Analysis of Manufacturing Systems
Birge and Louveaux: Introduction to Stochastic Programming
Bonnans and Shapiro: Perturbation Analysis of Optimization Problems
Bramel, Chen, and Simchi-Levi: The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management (second edition)
Dantzig and Thapa: Linear Programming 1: Introduction
Dantzig and Thapa: Linear Programming 2: Theory and Extensions
de Haan and Ferreira: Extreme Value Theory: An Introduction
Drezner (Editor): Facility Location: A Survey of Applications and Methods
Facchinei and Pang: Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I
Facchinei and Pang: Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II
Fishman: Discrete-Event Simulation: Modeling, Programming, and Analysis
Fishman: Monte Carlo: Concepts, Algorithms, and Applications
Haas: Stochastic Petri Nets: Modeling, Stability, Simulation
Klamroth: Single-Facility Location Problems with Barriers
Muckstadt: Analysis and Algorithms for Service Parts Supply Chains
Nocedal and Wright: Numerical Optimization
Olson: Decision Aids for Selection Problems
Pinedo: Planning and Scheduling in Manufacturing and Services
Pochet and Wolsey: Production Planning by Mixed Integer Programming
Resnick: Heavy Tail Phenomena: Probabilistic and Statistical Modeling
Whitt: Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues
Yao (Editor): Stochastic Modeling and Analysis of Manufacturing Systems
Yao and Zheng: Dynamic Control of Quality in Production-Inventory Systems: Coordination and Optimization
Yeung and Petrosyan: Cooperative Stochastic Differential Games

Forthcoming

- Muckstadt and Sapra*: Models and Solutions in Inventory Management

Sidney I. Resnick

Heavy-Tail Phenomena

Probabilistic and Statistical Modeling

Sidney I. Resnick
Cornell University
School of Operations Research and
Industrial Engineering
Ithaca, NY 14853
U.S.A.
sirl@cornell.edu

Series Editors:

Thomas V. Mikosch
University of Copenhagen
Laboratory of Actuarial Mathematics
DK-1017 Copenhagen
Denmark
mikosh@act.ku.dk

Stephen M. Robinson
University of Wisconsin-Madison
Department of Industrial Engineering
Madison, WI 53706
U.S.A.
smrobins@facstaff.wise.edu

Sidney I. Resnick
Cornell University
School of Operations Research and
Industrial Engineering
Ithaca, NY 14853
U.S.A.
sirl@cornell.edu

Mathematics Subject Classification (2000): 60F17, 60F05, 60K25, 60K10, 60K30, 60G18, 60G51, 60G52, 60G55, 60G57, 60G70, 62G32, 62M10, 62E20, 62F10, 62F12, 62G30, 62G09, 62H20, 62P30, 62P05, 90B18, 90B15, 90B22, 68M20, 91B28, 91B30, 91B70

Library of Congress Control Number: 2006934621

ISBN-10: 0-387-24272-4 e-ISBN-10: 0-387-45024-6
ISBN-13: 978-0-387-24272-9 e-ISBN-13: 978-0-0387-45024-7

Printed on acid-free paper.

© 2007 Springer Science+Business Media, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media LLC, 233 Spring Street, New York, NY 10013, U.S.A.), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

9 8 7 6 5 4 3 2 1

springer.com

(JLS/Ham)

Preface

“... was it heavy? Did it achieve total heaviosity?”

—Alvie (Woody Allen) to Annie (Diane Keaton) in *Annie Hall*, 1977.

Heavy-tail analysis is a branch of extreme-value theory devoted to studying phenomena governed by large movements rather than gradual ones. It encompasses both probability modeling as well as statistical inference. Its mathematical tools are based on regular variation, weak convergence of probability measures and random measures and point processes. Its applications are diverse, including the following:

- data networks, where the presence of heavy-tailed file sizes on network servers leads to long range dependence in the traffic rates;
- finance, where financial returns are heavy tailed and thus risk management calculations of *value-at-risk* require heavy-tailed methods;
- insurance, where the field of reinsurance is, by its nature, obsessed with very large values.

The structure of the book

There is an introductory chapter to describe the flavor and applicability of the subject. Then there are two chapters termed *crash courses*: one on regular variation and the other on weak convergence. These chapters contain essential material that could have been relegated to appendices; however, you should go through them where they are placed in the book. If you know the material, move quickly. Otherwise, pay some attention to style and notation. In particular, note what goes on in Sections 3.4–3.6. Such chapters are, inevitably, a compromise between wanting the book to be self-contained and not wanting to duplicate at length what is standard in other excellent references.

Chapter 4 gets you into the heart of inference issues fairly quickly. The approach to inference is semiparametric and asymptotic in nature. This leads to a statistical theory that is different from classical contexts. We assume there is some structure out there at *asymptopia* and we are trying to infer what it is using a pitiful finite sample whose true model has not yet converged to the asymptotic model. Thus, maximum likelihood methods are not really available unless we simply assume from some threshold onwards that the asymptotic model holds. We give some diagnostics that help decide on values of parameters and when a heavy-tail model is appropriate.

Chapter 5 begins the probability treatment which is geared towards a dimensionless theory. It focuses on the Poisson process and stochastic processes derived from the Poisson process, including Lévy and extremal processes. We also give an introduction to data network modeling. Chapter 6 gives the dimensionless treatment of regular variation and its probabilistic equivalents. We survey weak convergence techniques and discuss why it is difficult to bootstrap heavy-tail phenomena. Chapter 7 exploits the weak convergence technology to discuss weak convergence of extremes to extremal processes and weak convergence of summation processes to Lévy limits. Special cases include sums of heavy-tailed iid random variables converging to α -stable Lévy motion. We close the chapter with a unit on how weak convergence techniques can be used to study various transformations of regularly varying random vectors. We include Tauberian theory for Laplace transforms in this discussion.

Applied probability takes center stage in Chapter 8 which uses heavy-tail techniques to learn about the properties of three models. Two of the models are for data networks and the last one is a more traditional queueing model. We return to statistical issues in Chapter 9, discussing asymptotic normality for estimators and then moving to inference for multivariate heavy-tailed models. We include examples of analysis of exchange rate data, Internet data, telephone network data and insurance data. Finally, we close the chapter with a discussion of the much praised and vilified sample correlation function. There are some appendices devoted to notational conventions and a list of symbols and also a section which timidly discusses some useful software.

Each chapter contains exercises. Ignoring the exercises guarantees voyeur status.

Acknowledgments

Several institutions provided support and allowed me to harrass their students with heavy lectures that eventually became the basis for this book. I gave a semester-long seminar at Cornell in the spring of 2002 which I thought would create the momentum for a quick completion of the writing task. However, having assumed the directorship of Cornell's superb School of Operations Research and Industrial Engineering, it was impossible to find the time to focus on the pleasures of writing. Upon finishing my term as Director in 2004, I spent a month at the University of Bern, giving a short course on heavy tails, and subsequently, during the summer of 2005, while serving as Eurandom Professor, I gave a modified set of lectures in Eindhoven, Netherlands.

Finally, in the fall of 2005, while visiting Columbia University I was able to obsessively focus on assembling all of my thoughts. I much appreciated Columbia's fine atmosphere and hospitality as well as the support and office space provided by Columbia's Departments of Statistics, Department of Industrial Engineering and Operations Research, and Columbia's Business School. While at Columbia I lectured from material written for the book and Soumik Pal and Frank Isaacs proved to be prize-winning fault finders.

The final polishing, indexing, and debugging was done while housed and supported by the University of North Carolina's Department of Statistics and Operations Research and also by SAMSI/NISS at Research Triangle Park. During this period, Bikramjit Das at Cornell supplied me with additional corrections and comments via Skype.

Some additional acknowledgments in random order:

For many years, the National Science Foundation and the National Security Agency have provided the research support necessary for ideas to flourish.

Springer/Birkhäuser continue to be a pleasure to work with, and Ann Kostant should be declared a national treasure. Finally, I want to express a big thank you to John Spiegelman, for working with me, tirelessly, often suggesting things I overlooked; he

was altogether wonderful as well as amusing. And, as part of the Springer/Birkhäuser team, Elizabeth Loew continues to inspire total trust in handling my books.

Being on sabbatical is to be truly in a state of grace and I appreciate Cornell's support for the institution of sabbaticals. The opportunity to become a working member of Cornell in 1987 was a turning point in my adult professional life.

I have been blessed and supported by excellent colleagues, coworkers and students over the years who have helped me develop ideas about heavy tails. A partial list includes Laurens de Haan, Richard Davis, Holger Rootzen, Charles Goldie, Paul Feigin, Gennady Samorodnitsky, Catalin Stărică, Eric van den Berg, Krishanu Maulik, and Jan Heffernan. Two anonymous referees provided reviews chock full of useful, constructive criticism and suggestions. . . . And then there is Thomas Mikosch, who undertook to read the whole manuscript and made an uncountable number of a.e. useful and wise comments. Thomas displayed a genius for identifying inconsistencies and was only mildly sarcastic about my inept grammar, hyphenation, misplaced parentheses and creative spelling. ("Is that American spelling?") I, like, totally owe this guy a Coke! This is the second time [259] in my career I have been blessed by an effort whose helpfulness went way beyond what one has a right to expect from a colleague and friend.

Reminder to the blasé: $\text{T}_{\text{E}}\text{X}$, $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$, and $\text{BibT}_{\text{E}}\text{X}$ are astonishingly useful and elegant, as is $\text{MikT}_{\text{E}}\text{X}$.

Family history as described in literature:

- (1987) I . . . thank Minna, Nathan, and Rachel Resnick for a cheery, happy family life. Minna and Rachel bought me the mechanical pencil that made this project possible, and Rachel generously shared her erasers with me as well as providing a backup mechanical pencil from her stockpile when the original died after 400 manuscript pages. I appreciate the fact that Nathan was only moderately aggressive about attacking my Springer-Verlag correspondence with a hole puncher [260].
- (1992) Minna, Rachel, and Nathan Resnick provided a warm, loving family life and generously shared the home computer with me. They were also very consoling as I coped with two hard disk crashes and a monitor meltdown [262].
- (1998) Rachel, who grew into a terrific adult, no longer needs to share her mechanical pencils with me. Nathan has stopped attacking my manuscripts with a hole puncher and gives ample evidence of the fine adult he will soon be. Minna is the ideal companion on the random path of life [264].

Time marches on and it is 2006. Nathan graduated Cornell, moved to New York City, and found a job, and Rachel (now a Director!) has married Randy. Nathan and Randy gang up on poor, defenseless me, which Minna claims I deserve. Minna and I,

the artist and the ninja math geek, continue to explore a wonderful life together. As for mechanical pencils and home computers, I wrote this whole book on a laptop, the wonderous IBM (now Lenovo) ThinkPad. It is remarkable what one can accomplish while moving around with a good laptop and Internet connection.

Ithaca, NY

Sidney Resnick
June 2006

Contents

Preface	v
Acknowledgments	vii
1 Introduction	1
1.1 Welcome	1
1.2 Survey	1
1.3 Context and examples	3
1.3.1 Data networks	3
1.3.2 Finance	5
Value-at-risk	9
1.3.3 Insurance and reinsurance	13

Part I Crash Courses

2 Crash Course I: Regular Variation	17
2.1 Preliminaries from analysis	17
2.1.1 Uniform convergence	17
2.1.2 Inverses of monotone functions	18
2.1.3 Convergence of monotone functions	19
2.1.4 Cauchy's functional equation	20
2.2 Regular variation: Definition and first properties	20
2.2.1 A maximal domain of attraction	23
2.3 Regular variation: deeper results; Karamata's theorem	24
2.3.1 Uniform convergence	24

2.3.2	Integration and Karamata's theorem	25
2.3.3	Karamata's representation	29
2.3.4	Differentiation	30
2.4	Regular variation: Further properties	32
2.5	Problems	35
3	Crash Course II: Weak Convergence; Implications for Heavy-Tail	
	Analysis	39
3.1	Definitions	39
3.2	Basic properties of weak convergence	40
3.2.1	Portmanteau theorem	40
3.2.2	Skorohod's theorem	41
3.2.3	Continuous mapping theorem	42
3.2.4	Subsequences and Prohorov's theorem	43
3.3	Some useful metric spaces	44
3.3.1	\mathbb{R}^d , finite-dimensional Euclidean space	44
3.3.2	\mathbb{R}^∞ , sequence space	45
3.3.3	$C[0, 1]$ and $C[0, \infty)$, continuous functions	45
3.3.4	$D[0, 1]$ and $D[0, \infty)$	46
3.3.5	Radon measures and point measures; vague convergence	48
	Spaces of measures	48
	Convergence concept	49
	The vague topology; more on $M_+(\mathbb{E})$ (and hence, more on $M_p(\mathbb{E})$)	51
3.4	How to prove weak convergence	53
3.4.1	Methods in spaces useful for heavy-tail analysis	53
3.4.2	Donsker's theorem	54
3.5	New convergences from old	55
3.5.1	Slutsky approximations	55
3.5.2	Combining convergences	57
3.5.3	Inversion techniques	58
	Inverses	58
	Vervaat's lemma	59
3.6	Vague convergence and regular variation	61
	Vague convergence on $(0, \infty]$	62
3.7	Problems	64

Part II Statistics

4	Dipping a Toe in the Statistical Water	73
4.1	Statistical inference for heavy tails: This is a song about α	73
4.2	Exceedances, thresholds, and the POT method	74
4.2.1	Exceedances	75
4.2.2	Exceedance times	75
	Subsequence principle	76
4.2.3	Peaks over threshold	77
4.3	The tail empirical measure	78
4.4	The Hill estimator	80
4.4.1	Random measures and the consistency of the Hill estimator	81
4.4.2	The Hill estimator in practice	85
4.4.3	Variants of the Hill plot	89
	The smooHill plot	89
	Changing the scale, Alt plotting	90
4.5	Alternative estimators I: The Pickands estimator	90
4.5.1	Extreme-value theory	91
4.5.2	The Pickands estimator	93
4.6	Alternative estimators II: QQ plotting and the QQ estimator	97
4.6.1	Quantile-quantile or QQ plots: Preliminaries	97
4.6.2	QQ plots: The method	98
4.6.3	QQ plots and location-scale families	100
4.6.4	Adaptation to the heavy-tailed case: Are the data heavy tailed?	101
4.6.5	Additional remarks and related plots	102
	Diagnosing deviations from the line in the QQ plot	102
	A related plot: The PP plot	104
	Another variant: The tail plot for heavy tails	104
4.6.6	The QQ estimator	106
	Consistency of the QQ estimator	108
4.7	How to compute value-at-risk	111
4.8	Problems	114

Part III Probability

5	The Poisson Process	119
5.1	The Poisson process as a random measure	119
5.1.1	Definition and first properties	119
5.1.2	Point transformations	120
5.1.3	Augmentation or marking	122
5.2	Models for data transmission	123
5.2.1	Background	124
5.2.2	Probability models	125
5.2.3	Long-range dependence	126
	Simple minded detection of long-range dependence using the sample acf plot	127
5.2.4	The infinite-node Poisson model	127
5.2.5	Connection between heavy tails and long-range dependence	130
5.3	The Laplace functional	132
5.3.1	Definition and first properties	132
5.3.2	The Laplace functional of the Poisson process	134
5.4	See the Laplace functional flex its muscles!	137
5.4.1	The Laplace functional and weak convergence	137
	Convergence of empirical measures	138
	Preservation of weak convergence under mappings of the state space	141
5.4.2	A general construction of the Poisson process	143
5.4.3	Augmentation, location-dependent marking	144
5.5	Lévy processes	146
5.5.1	Itô's construction of Lévy processes	146
	Lévy measure	146
	Compound Poisson representations	147
	Variance calculations	148
	Process definition	149
5.5.2	Basic properties of Lévy processes	150
	The characteristic function of $X(t)$	151
	Independent increment property of $X(t)$	151
	Stationary increment property	152
	Stochastic continuity of $X(\cdot)$	152
	Subordinators	153
	Stable Lévy motion	154

	Symmetric α -stable Lévy motion	154
5.5.3	Basic path properties of Lévy processes	155
5.6	Extremal processes	160
5.6.1	Construction	161
5.6.2	Discussion	161
5.7	Problems	162
6	Multivariate Regular Variation and the Poisson Transform	167
6.1	Multivariate regular variation: Basics	167
6.1.1	Multivariate regularly varying functions	167
6.1.2	The polar coordinate transformation	168
6.1.3	The one-point uncompactification	170
6.1.4	Multivariate regular variation of measures	172
6.2	The Poisson transform	179
6.3	Multivariate peaks over threshold	183
6.4	Why bootstrapping heavy-tailed phenomena is difficult	184
6.4.1	An example to fix ideas	184
6.4.2	Why the bootstrap sample size must be carefully chosen	186
	The bootstrap procedure	186
	What exactly is the bootstrap procedure?	187
	When bootstrap asymptotics work	188
	When bootstrap asymptotics do not work	189
6.5	Multivariate regular variation: Examples, comments, amplification	191
6.5.1	Two examples	192
	Independence and asymptotic independence	192
	Repeated components and asymptotic full dependence	195
6.5.2	A general representation for the limiting measure ν	196
6.5.3	A general construction of a multivariate regularly varying distribution	197
6.5.4	Regularly varying densities	199
6.5.5	Beyond the nonnegative orthant	201
6.5.6	Standard vs. nonstandard regular variation	203
6.6	Problems	206
7	Weak Convergence and the Poisson Process	211
7.1	Extremes	211
7.1.1	Weak convergence of multivariate extremes: The timeless result	211

7.1.2	Weak convergence of multivariate extremes: Functional convergence to extremal processes	212
7.2	Partial sums	214
7.2.1	Weak onvergence of partial sum processes to Lévy processes . . .	214
7.2.2	Weak convergence to stable Lévy motion	218
7.2.3	Continuity of the summation functional	221
7.3	Transformations	226
7.3.1	Addition	227
	Linear combinations of components of a random vector	227
	Adding independent vectors	228
7.3.2	Products	231
	Breiman’s theorem: A factor has a relatively thin tail	231
	Products of heavy-tailed random variables which are jointly regularly varying	236
	Internet data	238
7.3.3	Laplace transforms	239
	Special case for $d = 1$: Karamata’s Tauberian theorem	245
	Renewal theory	245
7.4	Problems	247
8	Applied Probability Models and Heavy Tails	253
8.1	A network model for cumulative traffic on large time scales	253
8.1.1	Model review	253
8.1.2	The critical input rate	255
8.1.3	Why stable Lévy motion can approximate cumulative input under slow growth	257
	The basic decomposition	257
	One-dimensional convergence	259
	Finite-dimensional convergence	263
8.2	A model for network activity rates	264
8.2.1	Mean value analysis when $\alpha, \beta < 1$	265
8.2.2	Behavior of $N(t)$, the renewal counting function when $0 < \alpha < 1$	265
8.2.3	Activity rates when $\alpha, \beta < 1$ and tails are comparable	266
	Counting function of $\{(S_k, T_k), k \geq 0\}$	266
	Number of active sources when tails are comparable	267
8.2.4	Activity rates when $0 < \alpha, \beta < 1$, and F_{on} has a heavier tail . . .	269
	Number of active sources when \bar{F}_{on} is heavier	271
8.3	Heavy traffic and heavy tails	272

8.3.1	Crash course on waiting-time processes	273
8.3.2	Heavy-traffic approximation for queues with heavy-tailed services	275
8.3.3	Approximation to a negative-drift random walk	279
8.3.4	Approximation to the supremum of a negative-drift random walk	281
8.3.5	Proof of the heavy-traffic approximation	283
8.4	Problems	286

Part IV More Statistics

9	Additional Statistics Topics	291
9.1	Asymptotic normality	291
9.1.1	Asymptotic normality of the tail empirical measure	291
9.1.2	Asymptotic normality of the Hill estimator	296
	Blood and guts	298
	Removing the random centering	299
	Centering by $1/\alpha$	302
	Conclusions	303
9.2	Estimation for multivariate heavy-tailed variables	304
9.2.1	Dependence among extreme events	304
	Example: Modeling of exchange rates	305
9.2.2	Estimation in the standard case	307
9.2.3	Estimation in the nonstandard case	309
	Live with diversity	309
	Be crude!	310
	The ranks method	310
	Estimation of the angular measure	313
9.2.4	How to choose k ; the Stărică plot	314
9.3	Examples	316
9.3.1	Internet data	316
	Boston University data	316
	Internet HTTP response data	317
9.3.2	Exchange rates	318
9.3.3	Insurance	319
9.4	The coefficient of tail dependence and hidden regular variation	322
9.4.1	Hidden regular variation	323
	Definition of hidden regular variation	324
	Topology is destiny	324

9.4.2	A simple characterization	325
9.4.3	Two examples	330
9.4.4	Detection of hidden regular variation	332
	A first step	332
	But wait! Why does the rank transform preserve hidden regular variation?	332
	Estimating the hidden angular measure	337
9.5	The sample correlation function	340
9.5.1	Overview	340
9.5.2	Limit theory	342
	Preliminaries	342
	Point process limits	343
	Summing the points	346
9.5.3	The heavy-tailed sample acf; $\alpha < 1$	346
9.5.4	The classical sample acf: $1 < \alpha < 2$	347
9.5.5	Suggestions to use	349
9.6	Problems	350

Part V Appendices

10	Notation and Conventions	359
10.1	Vector notation	359
10.2	Symbol shock	360
11	Software	363
11.1	One dimension	364
11.1.1	Hill estimation	364
	Hillalpha	364
	altHillalpha	364
	smooHillalpha	365
11.1.2	QQ plotting	366
	pppareto	366
	parfit	366
	QQ estimator plot	367
11.1.3	Estimators from extreme-value theory	368
	The Pickands estimator	368
	The moment estimator	369
11.2	Multivariate heavy tails	370
11.2.1	Estimation of the angular distribution	370

Rank transform	371
Estimate the angular density using ranks	371
Estimate the angular density using power transforms	371
Estimate the angular distribution using the rank transform	372
11.2.2 The Stărică plot	372
Norms	373
Stărică plot using the power transform	373
Stărică plot using the rank transform	374
Allowing the Stărică plot to choose k	374
References	377
Index	397