Springer Series in Operations Research and Financial Engineering

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

Springer

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.

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: TEX, LATEX, and BibTEX are astonishingly useful and elegant, as is MikTEX.

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

Part II Statistics

Part III Probability