

The Fundamentals of Heavy-tails: Properties, Emergence, and Identification

Jayakrishnan Nair
California Institute of
Technology
Pasadena, CA, USA
ujk@caltech.edu

Adam Wierman
California Institute of
Technology
Pasadena, CA, USA
adamw@caltech.edu

Bert Zwart
Centrum voor Wiskunde en
Informatica
Amsterdam, The Netherlands
Bert.Zwart@cwi.nl

Categories and Subject Descriptors

G.3 [Mathematics of computing]: Probability and Statistics

Keywords

Heavy-tailed distributions

Tutorial overview

Heavy-tails are a continual source of excitement and confusion across disciplines as they are repeatedly “discovered” in new contexts. This is especially true within computer systems, where heavy-tails seemingly pop up everywhere – from degree distributions in the internet and social networks to file sizes and interarrival times of workloads. However, despite nearly a decade of work on heavy-tails they are still treated as mysterious, surprising, and even controversial.

The goal of this tutorial is to show that heavy-tailed distributions need not be mysterious and should not be surprising or controversial. In particular, we will demystify heavy-tailed distributions by showing how to reason formally about their counter-intuitive properties; we will highlight that their emergence should be *expected* (not surprising) by showing that a wide variety of general processes lead to heavy-tailed distributions; and we will highlight that most of the controversy surrounding heavy-tails is the result of bad statistics, and can be avoided by using the proper tools.

Intended audience

The tutorial is aimed at students, researchers, and practitioners interested in learning how to rigorously think about heavy-tailed distributions. The assumed background is only a basic understanding of probability and statistics.

References

The organizers of this tutorial are in the process of writing a book on the topic of heavy-tails, and this tutorial is meant to provide a brief overview of the material to be covered in the forthcoming book.

Speaker Biographies

Jayakrishnan Nair received his PhD from California Institute of Technology (Caltech) in 2012. His PhD thesis focused on scheduling for heavy-tailed and light-tailed workloads in queueing systems. He is currently a post-doctoral

scholar at Caltech and will join CWI as a post-doctoral scholar in summer 2013. His research interests include modeling, performance evaluation, and design issues in queueing systems and communication networks. Jayakrishnan was a recipient of the best paper award at IFIP Performance, 2010.

Adam Wierman is a Professor in the Department of Computing and Mathematical Sciences at the California Institute of Technology, where he is a member of the Rigorous Systems Research Group (RSRG). He received his Ph.D., M.Sc. and B.Sc. in Computer Science from Carnegie Mellon University in 2007, 2004, and 2001, respectively. His research interests center around resource allocation and scheduling decisions in computer systems and services. More specifically, his work focuses both on developing analytic techniques in stochastic modeling, queueing theory, scheduling theory, and game theory, and applying these techniques to application domains such as energy-efficient computing, data centers, social networks, and electricity markets. He received the 2011 ACM SIGMETRICS Rising Star award, and has been co-recipient of best paper awards at ACM SIGMETRICS, IEEE INFOCOM, IFIP Performance, IEEE Green Computing Conference, and ACM GREENMETRICS.

Bert Zwart is currently a senior researcher at CWI, where he leads the Probability and Stochastic Networks group. He also holds a full professor position at VU University Amsterdam, is senior fellow at Eurandom, and holds an adjunct professor position at the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology, where he was holding a Coca-Cola Chair until 2008. Bert Zwart is the 2008 recipient of the Erlang prize for outstanding contributions to applied probability by a researcher not older than 35 years old, and an IBM faculty award. His research is concerned with the application of analytic and probabilistic asymptotic methods to applied probability models in computer systems, communication networks, customer contact centers, and manufacturing systems. Dr. Zwart has published more than 70 refereed publications and is council member of the Applied Probability Society of INFORMS. Dr. Zwart has been area editor of Stochastic Models for Operations Research, the flagship journal of his profession, from 2009-2011. In addition, Dr. Zwart is editor-in-chief (with J.K. Lenstra and M. Trick) of the journal Surveys in Operations Research and Management Science, and serves on the editorial board of Mathematics of Operations Research, Mathematical Methods of Operations Research, Operations Research, Queueing Systems and Stochastic Systems.