Workshop on

Logical methods in Ramsey Theory and related topics

10-11 July 2023 - Department of Mathematics, Pisa, ITALY

Initial Applications of Alpha Theory in Telecommunications Engineering

Marco Cococcioni, Francesco Fiorini and Michele Pagano

Dipartimento di Ingegneria dell'Informazione
Via Caruso 16 – 56122 Pisa

Università di Pisa

Contents

IN SUPREMÆ DIGNITATIS 1343

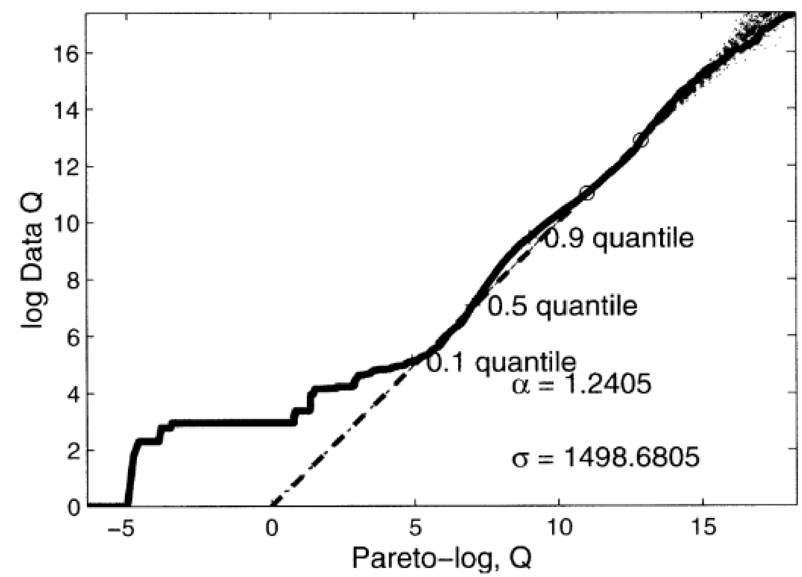
Introduction: What are heavy tailed distributions?

- Heavier tail than the exponential distribution
- Many **outliers** with very high values
- **«Infinite» variance**, or in general, not all their moments finite

→ *High variability* ↔ *Noah Effect*

TLC applications

- File sizes on a web server, uptime and silence times in remote communications, CPU times, connection times.
- Interarrival time of Internet data packets in Ethernet LANs and in WAN
- Interference power in IoT communications
- Variable Bit Rate video streaming traffic



Other fields

- Financial Risk Engineering
- Outputs of machine learning algorithms (e.g. SGD for neural networks)

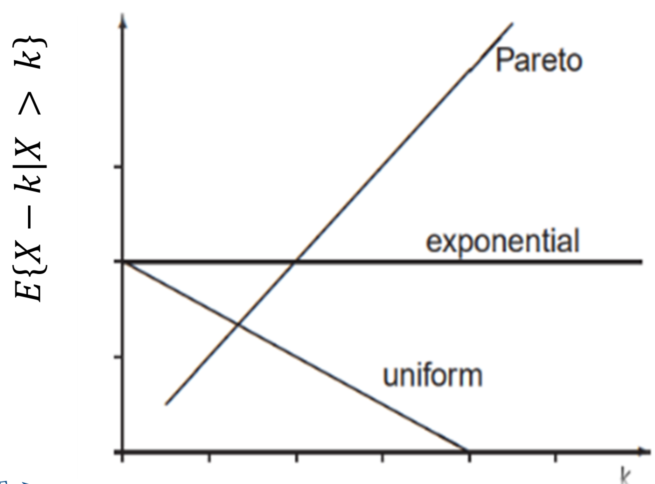
Two important properties

(1) Expectation Paradox

$$E\{X-k|X > k\} \sim k$$

$$= \int_k^{\infty} (x-k)\,\frac{f_X(x)}{\int_k^{\infty} f_X(q)\,dq}\,dx = \frac{\int_k^{\infty} x\,f_X(x)\,dx}{\int_k^{\infty} f_X(q)\,dq} - k$$

"The longer we have waited… the longer we should expect to wait!"



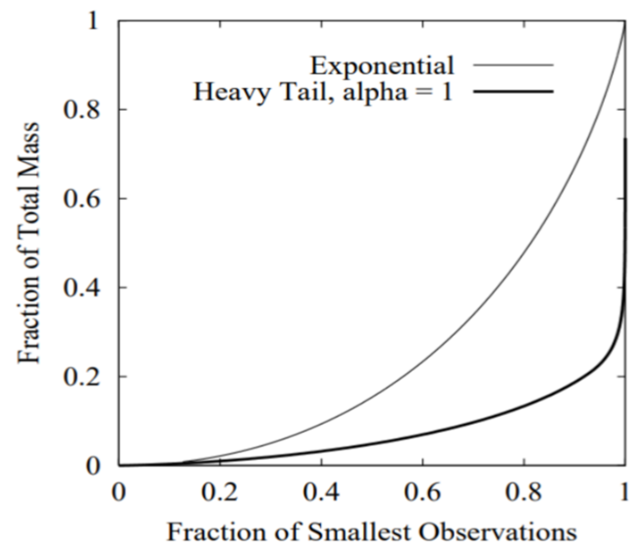
3 Different Examples

(2) Mass Count Disparity

$$\lim_{x\to\infty} \frac{\mathrm{P}[X_1 + \ldots + X_n > x]}{P[\max\,[X_1, \ldots, X_n] > x]} = 1 \quad for\ all\ n \geq 2$$

"A very tiny subset of observations contains the vast bulk of the mass in a set of data"

- **80%** of the smallest observations represent less than **20%** of the total mass



TLC real-world example	*50%-80% of the bytes in FTP transfers are due to the largest 2% of all transfers*

How to exploit heavy tailed models in TLC?

AIM: **Reliable statistical methods** for the purposes of network analysis, network management and design, performances evaluation and protocols optimization, and therefore reduction of **over-provisioning,** without the purchase of additional resources

Some Applications → **Where?** ***Packet Switched Networks***

Eg. Core Network of a cellular network (from LTE IP based!)

Load balancing *in network of queues (e.g. SITA-E, Size Interval Task Assignment with Equal Load)*

- only a small number of tasks need to be relocated due to the mass-count difference
- the expectation paradox means that a task's current lifespan is a good indicator of its predicted future lifetime.

Scheduling in Web servers (e.g. SRTP, shortest-remaining-processing-time)

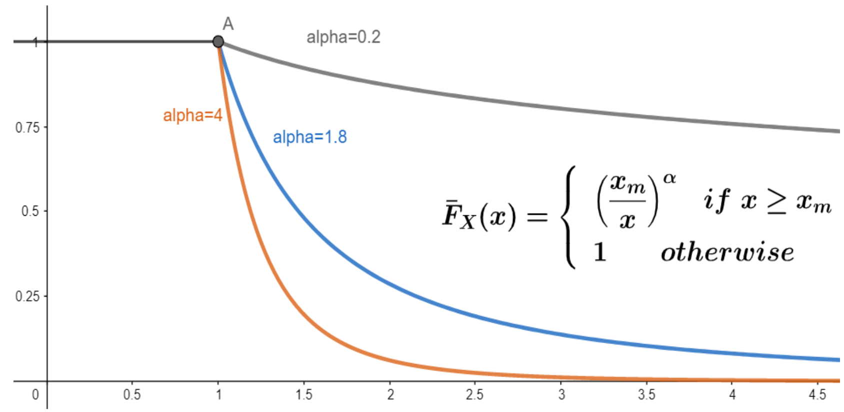
Routing and Switching in Internet

- hardware switching to create temporary circuits (**shortcuts**) for lengthy packet sequences
- setup threshold=number of same-path packets to watch before establishing a switched connection
- majority of bytes may be routed by implementing shortcuts for just a tiny portion of all data flows

Simulations with standard heavy tailed distributions

Some Issues

∞
1. «Infinite» variance
||
«diverging»



Example: The problem of using Pareto distribution and standard analysis

- The sample variance is not able to give a direct information, aligned with the teoretical variance (which is known to be ∞)

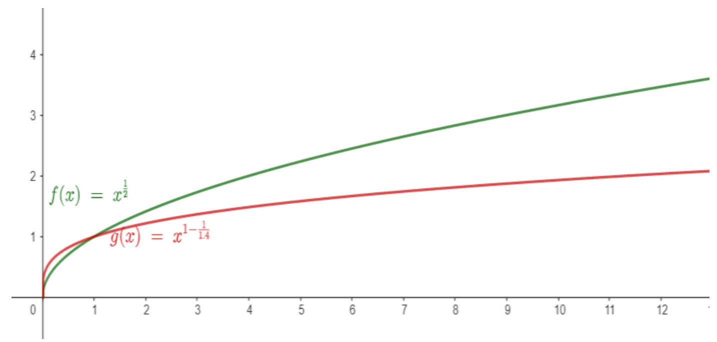
E{X} = 4/3 th_var{X} = 2/9 est_var = 0.222 (the accuracy increases with # of samples)
E{X} = 2.25 th_var{X} = ∞ est_var = 1.2E+4, 2.3E+6, ... **(it increases with the # of samples!)**
E{X} = ∞ th_var{X} = ∄ est_var = 4.5E+10, 7.4E+23, ... **(it increases with the # of samples!)**

→ Not numerically verifiable

Simulations with standard heavy tailed distributions

Some Issues

2. **Slow Convergence &**
   High variability at Steady State



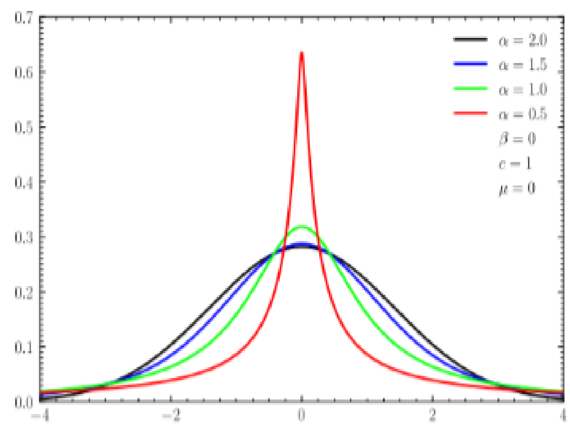
~~CLT~~ → **GCLT**

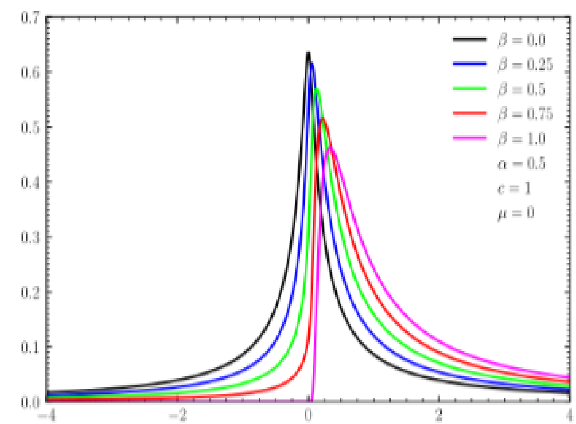
$$A_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$S_n = (A_n - \mu)\, n^{1-\frac{1}{\alpha}} \qquad S_n \to L_\alpha$$

$$G_n = (A_n - \mu)\, n^{1/2}$$

$$G_n \to N\left(0, \sigma^2\right)$$

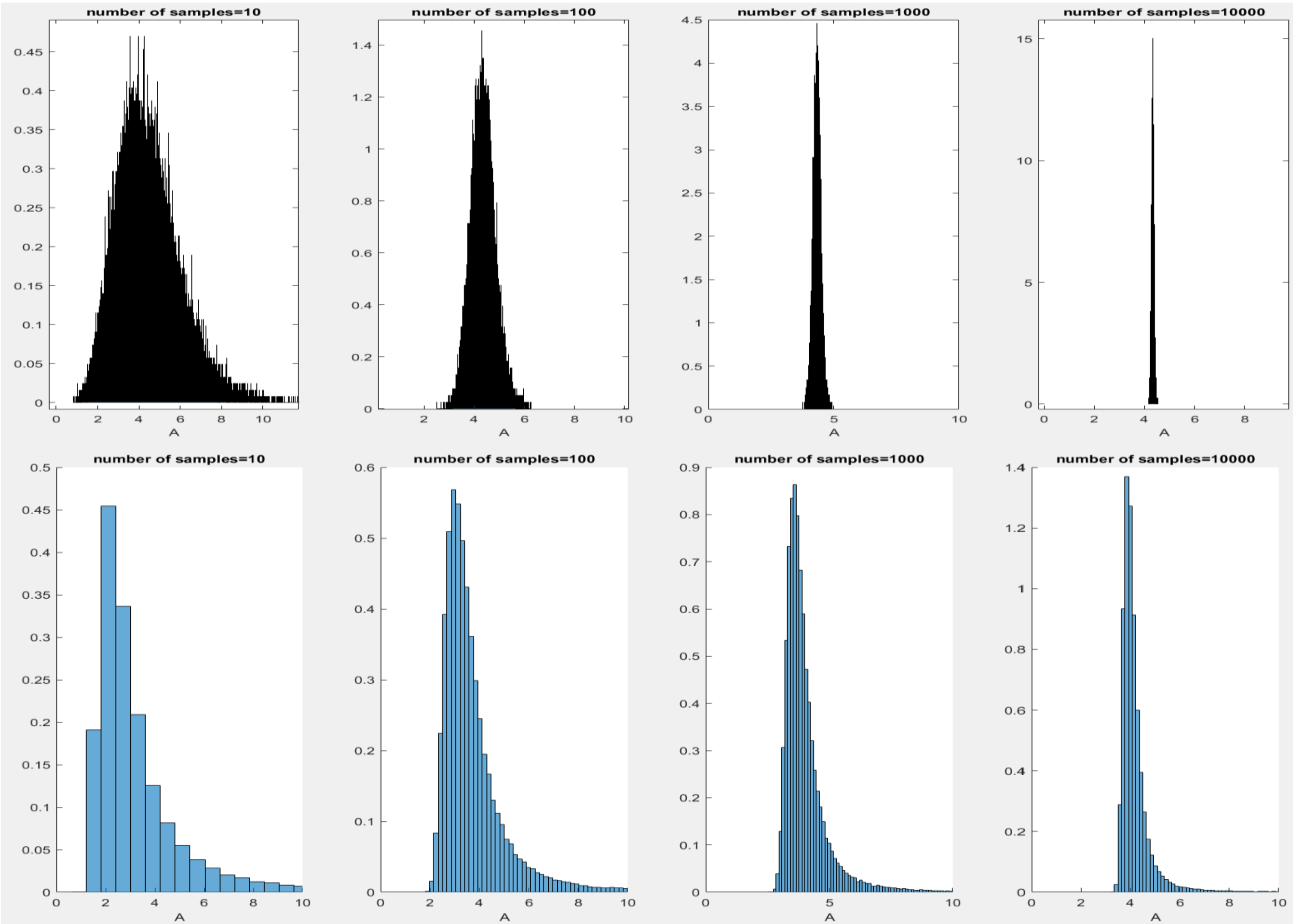




Simulations with standard heavy tailed distributions

Some Issues

2. Slow Convergence & High variability at Steady State



Simulations with standard heavy tailed distributions

Some Issues

2. Slow Convergence & High variability at Steady State

Do you request t digits of accuracy of your sample mean estimator?

$$\frac{|A_n - \mu|}{\mu} \leq 10^{-t}$$

With respect to the true value μ

Recalling the previous formula…

$$|A_n - \mu| = n^{\frac{1}{\alpha}-1} c.$$

$$n^{\frac{1}{\alpha}-1} \leq \frac{1}{v} 10^{-t}$$

$$n^{1-\frac{1}{\alpha}} \geq v\, 10^{t}$$

$$n \geq q\, 10^{\frac{t}{1-1/\alpha}}$$

α	n
2.0	10^6
1.7	$19.3\, 10^7$
1.5	10^9
1.3	10^{13}
1.2	10^{18}
1.1	10^{33}

A practical example with t=3

What can help us?
Euclidean Numbers!

Standard analysis:

uses the set of ***Real** numbers*

$\mathbb{R}$ **(the field of real numbers):**

- $\mathbb{R}$ contains **only finite values**
- Sometimes the set $\overline{\mathbb{R}}$ is introduced, as the union of $\mathbb{R}$ and the symbols $+\infty$ and $-\infty$
- ∞ is just a symbol, not a number
- $\infty - \infty$, $\infty + \infty$, $\frac{\infty}{\infty}$ are not allowed operations between elements of $\overline{\mathbb{R}}$
- Implemented in C++/Matlab as a discrete set, using binary digits, according to the IEEE floating point format

An example of IEEE Float in Matlab

4.15E+7 ⇔ $4.15 \cdot 10^7$

Non-Standard Analysis:

uses the set of ***Euclidean** numbers*

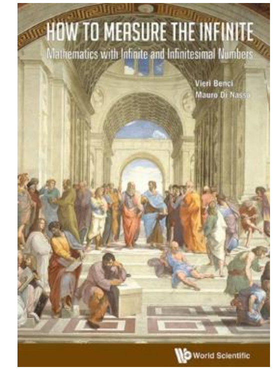
$\mathbb{E}$ **(the Field of Euclidean numbers)**

- $\mathbb{E}$ includes $\mathbb{R}$, hence it ***extends*** $\mathbb{R}$
- $\mathbb{E}$ contains finite, **infinitesimal** and **iperfinite** numbers
- $\eta \in \mathbb{E}$ is the prototypal infinitesimal number
- $\alpha \in \mathbb{E}$ is the prototypal iperfinite values ($\alpha = 1/\eta$ and α= numerosity($\mathbb{N}$))
- Euclidean numbers can be implemented in C/C++/Matlab, using a floating point like-approach: they have been called **Bounded Algorithmic Numbers (BANs)**

An example of BAN in Matlab

(3.5 -2.3 1.2)A+3 ⇔ $(3.5 - 2.3\eta + 1.2\eta^2)\alpha^3$

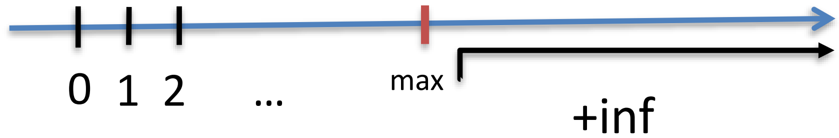
What can help us?
Benci-Di Nasso Alpha Theory!



In MATLAB

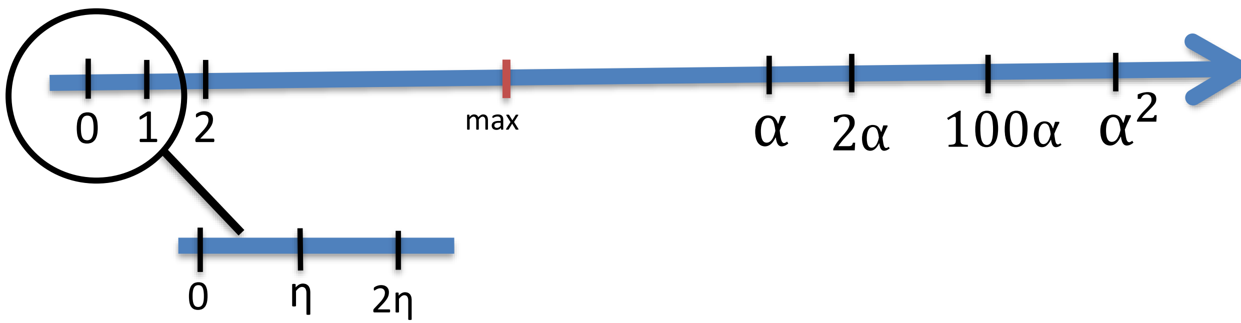
Standard analysis:

FROM THIS…



… TO THIS

Non-Standard Analysis



Thanks to our *implementantion* of a *Matlab object-oriented toolbox*, with the classes Ban and BanArray

In the same computer, with the same memory!

$$\alpha\text{-}3\alpha = -2\alpha \qquad \frac{\alpha+1}{\alpha} = 1 + \eta$$

$$\alpha \cdot (\alpha + 2) = \alpha^2 + 2\alpha$$

$$0 < \frac{1}{\alpha} = \alpha^{-1} < \alpha^0 = 1 < \alpha^1 = \alpha < \alpha + 1$$

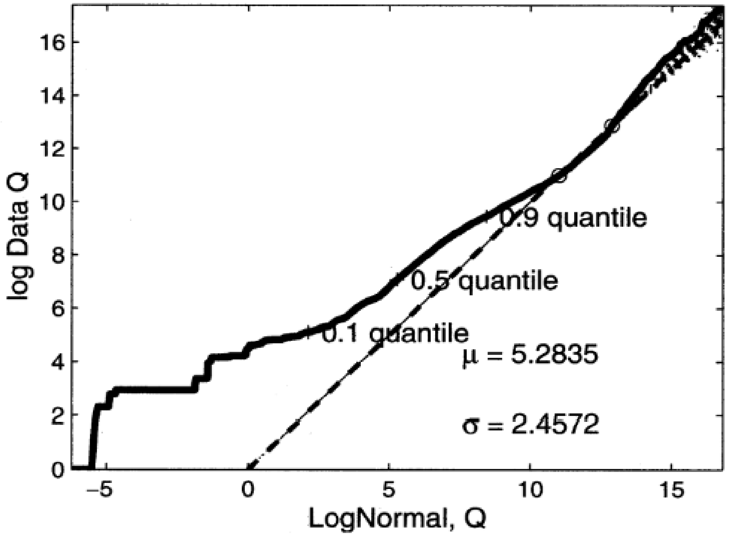
New definition of heavy tailed distributions

New Definition: *A distribution is considered heavy tailed if its variance (or in general any of its moments) diverges towards +∞,* ***or if it assumes a (precise) infinite value***

A practical example?

LogNormal distribution

- Several applications in the telecommunications scenarios: a good statistical model for the amount of Internet traffic per time unit, interference power PDF in device to device communications, shadowing in cellular networks…



… BUT

[The difficulty about using the Log Normal distribution]

«There is evidence that, over their entire range, many of the distributions we study may be well characterized using *lognormal* distributions [19]. However, lognormal distributions do not have infinite variance, and hence are not heavy-tailed. In our work, we are not concerned with distributions over their entire range --- only their tails. As a result we don't use goodness-of-fit tests to determine whether Pareto or lognormal distributions are better at describing our data. However, it is important to verify that our datasets exhibit the infinite variance characteristic of heavy tails. »

From "Explaining World Wide Web Traffic Self-Similarity", M. E. **Crovella** and A. Bestavros, 1995 ,Technical Report TR-95-015

What would alleviate/solve the problem?

The possibility to generate pseudo-random numbers, obeing a Log Normal distribution, <u>but having a finite mean and an infinite variance</u>

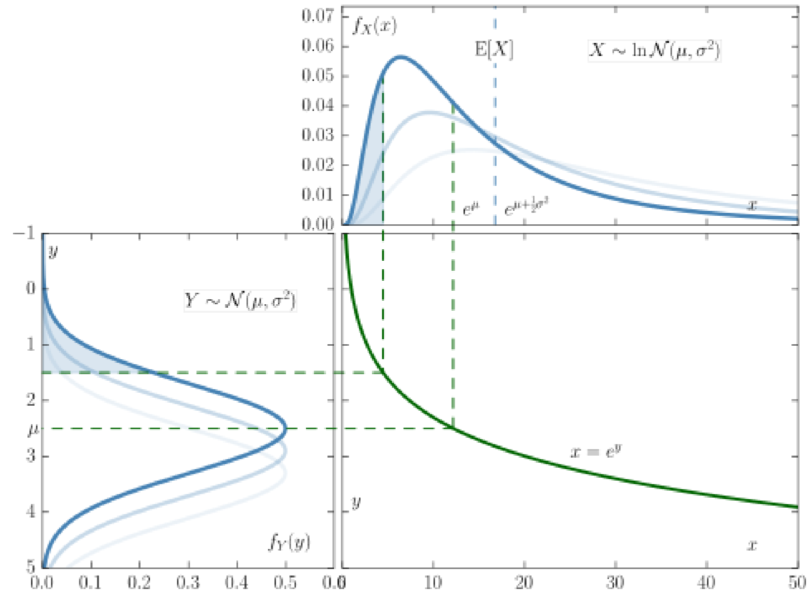
THIS CANNOT BE ACHIEVED USING STANDARD ANALYSIS AND REAL NUMBERS!

Euclidean LogNormal distribution

$$X = e^{\mu+\sigma Z} \quad , \quad Z \in N(0,1)$$

$$\mathrm{var}\{X\} = \left(e^{\sigma^2} - 1\right) e^{2\mu+\sigma^2}$$

$$E\{X\} = e^{\mu+\frac{\sigma^2}{2}}$$



Thanks to the two parameters, it is possible to obtain finite mean and hyperfinite variance: e.g. fixing:

$\sigma = \alpha$

$\mu = -\frac{\alpha^2}{2}$

$$E\{X\} = e^{\mu+\frac{\sigma^2}{2}} = e^{-\frac{\alpha^2}{2}+\frac{\alpha^2}{2}} = 1$$

$$\mathrm{VAR}\{X\} = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = = e^{-\alpha^2+2\alpha^2} - e^{-\alpha^2+\alpha^2} = e^{\alpha^2} - 1.$$

- We filled the **gap**: LogNormal and heavy tailness are **consistent** with each other

Thanks to our Matlab toolbox implementation, we have been able to generate pseudo-random samples following an Euclidean LogNormal distribution, with finite mean and well-defined infinite variance

Euclidean Gaussian distribution

```
(-0.5 -0.204753η +2.40271η^{2} +1.76278η^{3})α^{2}
(-0.5 -0.196902η +0.513552η^{2} +1.33282η^{3})α^{2}
(-0.5 -0.638722η -0.707328η^{2} -1.10714η^{3})α^{2}
(-0.5 +0.168241η -1.15482η^{2} -0.404674η^{3})α^{2}
(-0.5 -1.86823η +0.33322η^{2} +0.209238η^{3})α^{2}
(-0.5 -1.59852η +1.47004η^{2} -1.50524η^{3})α^{2}
(-0.5 -0.515042η -1.69168η^{2} +0.643354η^{3})α^{2}
(-0.5 +0.222991η -1.65619η^{2} -1.25562η^{3})α^{2}
(-0.5 -2.31859η -1.78236η^{2} -0.391342η^{3})α^{2}
(-0.5 +0.895515η -0.461908η^{2} +1.33434η^{3})α^{2}
(-0.5 -0.494549η -0.610375η^{2} +2.3183η^{3})α^{2}
(-0.5 +1.38613η -1.452η^{2} -1.35357η^{3})α^{2}
(-0.5 +0.196699η -0.61907η^{2} +0.750757η^{3})α^{2}
(-0.5 -1.2165η +0.0805883η^{2} +1.30813η^{3})α^{2}
(-0.5 +0.607305η -2.0783η^{2} -1.02289η^{3})α^{2}
(-0.5 -2.57167η -1.58652η^{2} +0.292704η^{3})α^{2}
(-0.5 -0.505779η -2.10522η^{2} -0.453094η^{3})α^{2}
(-0.5 +0.102648η -0.731497η^{2} +2.763η^{3})α^{2}
(-0.5 -1.22966η +0.770481η^{2} +0.293396η^{3})α^{2}
(-0.5 -1.9979η +2.14819η^{2} +1.27009η^{3})α^{2}
(-0.5 +1.01019η +1.50467η^{2} -0.684386η^{3})α^{2}
(-0.5 +0.895845η -0.306128η^{2} +1.51678η^{3})α^{2}
(-0.5 -0.655424η -1.09089η^{2} -0.844247η^{3})α^{2}
(-0.5 -0.0809119η +0.116024η^{2} -0.681505η^{3})α^{2}
(-0.5 +0.518093η +0.34763η^{2} +0.166315η^{3})α^{2}
(-0.5 -1.10611η +0.543505η^{2} +1.37694η^{3})α^{2}
(-0.5 -0.0336609η -0.563575η^{2} -0.0551417η^{3})α^{2}
(-0.5 +1.52928η -0.548954η^{2} -0.775672η^{3})α^{2}
(-0.5 +0.794615η +1.20346η^{2} -1.77628η^{3})α^{2}
(-0.5 +0.355161η +1.02907η^{2} +0.733149η^{3})α^{2}
(-0.5 -0.509023η -1.70048η^{2} -1.93089η^{3})α^{2}
(-0.5 -0.203379η +0.529956η^{2} -0.249475η^{3})α^{2}
(-0.5 -1.51332η -0.00464971η^{2} -0.431933η^{3})α^{2}
(-0.5 -0.667532η -0.725943η^{2} +1.025η^{3})α^{2}
(-0.5 -2.49426η -0.339739η^{2} +0.591365η^{3})α^{2}
(-0.5 +0.431631η +0.0148419η^{2} +0.574158η^{3})α^{2}
```

Underlying Euclidean Gaussian distribution

An example of 36 pseudo-random samples generated in Matlab

… How ?

1. By defining the two parameters μ and σ like Bans

```
true_MU_of_EG = Ban([-0.5 0 0 0 ], 2);
true_SIGMA_of_EG = Ban([1 0 0 0 ], 1);
```

2. By exploting the well-known Gaussian Displacement method

```
bArr = randn(10^N(n_), 1, 'like', BanArray);
g = bArr*true_SIGMA_of_EG + true_MU_of_EG;
```

Euclidean LogNormal distribution

... And to numerically check the coherence between the sample values of mean and variance with the correspective theoretical values

$\mu = -0.5\alpha^2 + 9.54681$

$\sigma = \alpha$

$$E\{X\} = e^{\mu+\frac{\sigma^2}{2}} = e^{-0.5\alpha^2 + 9.54681 + 0.5\alpha^2} = e^{9.54681}$$

$$\mathrm{VAR}\{X\} = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = e^{(1+19.0936\eta^2)\alpha^2} - e^{19.0936}$$

#	$E\{X\}$ (Sample Mean)	$VAR\{X\}$ (Sample Variance)
10^3	$e^{(\mathbf{0.0260943}+9.55455\eta^2)\alpha^2}$	$e^{(1.10438+19.1091\eta^2)\alpha^2} - e^{(\mathbf{0.0521886}+19.1091\eta^2)\alpha^2}$
10^4	$e^{(\mathbf{-0.0126293}\ \mathbf{+0.0190467}\eta+9.56013\eta^2)\alpha^2}$	$e^{(0.949483+\mathbf{0.0380934}\eta+19.1203\eta^2)\alpha^2} - e^{(\mathbf{-0.0252587}\ \mathbf{+0.0380934}\eta+19.1203\eta^2)\alpha^2}$
10^5	$e^{(\mathbf{-0.00351095}+9.54501\eta)\alpha}$	$e^{(1.00631\mathbf{-0.0070219}\eta+19.09\eta^2)\alpha^2} - e^{(\mathbf{0.00315714}\ \mathbf{-0.0070219}\eta+19.09\eta^2)\alpha^2}$
10^6	$e^{9.54617}$	$e^{(0.995845+19.0923\eta^2)\alpha^2} - e^{19.0923}$
10^7	$e^{9.54666}$	$e^{(1.00065+19.0933\eta^2)\alpha^2} - e^{19.0933}$

(Threshold: $2.5\ 10^{-3}$)

Fitting Phase

```
est_mean = mean(x);
est_var = var(x);
```

Reference

Journal Article

Section: *Engineering Mathematics*

https://doi.org/10.3390/math11071758

Article

Modelling Heavy Tailed Phenomena Using a LogNormal Distribution Having a *Numerically Verifiable* Infinite Variance

Marco Cococcioni *,†, Francesco Fiorini † and Michele Pagano †

Department of Information Engineering, L.go Lucio Lazzarino, 1-56122 Pisa, Italy
* Correspondence: marco.cococcioni@unipi.it
† These authors contributed equally to this work.

Abstract: One-sided heavy tailed distributions have been used in many engineering applications, ranging from teletraffic modelling to financial engineering. In practice, the most interesting heavy tailed distributions are those having a finite mean and a diverging variance. The LogNormal distribution is sometimes discarded from modelling heavy tailed phenomena because it has a finite variance, even when it seems the most appropriate one to fit the data. In this work we provide for the first time a LogNormal distribution having a finite mean and a variance which converges to a well-defined *infinite* value. This is possible thanks to the use of Non-Standard Analysis. In particular, we have been able to obtain a Non-Standard LogNormal distribution, for which it is possible to *numerically* and *experimentally* verify whether the expected mean and variance of a set of generated pseudo-random numbers agree with the theoretical ones. Moreover, such a check would be much more cumbersome (and sometimes even impossible) when considering heavy tailed distributions in the traditional framework of standard analysis.

Keywords: non-standard analysis; alpha-theory; algorithmic numbers; non-archimedean scientific computing; heavy tailed distributions

MSC: 03H10; 60E05; 65C20

«[...] The theorem and current results provided in this manuscript look very promising and **I am very impressed by it**. [...] » ← opinion of one of the reviewers (*boldface added by us*)

Another possible application in TLC

➢ **Queueing Theory**

M/G/1 ← Kendall's Notation

PK formula

$$E\{T_W\} = \frac{\lambda\, E\{T_s{}^2\}}{2\,(1-\rho)}$$

$$\rho = \frac{E\{n_s\}}{n_s} = \frac{\lambda}{\mu}$$

What if

$E\{T_s{}^2\} \to \infty$?

This could happen…
With an heavy tailed Service distribution!

Alpha Theory and Pareto distribution

Standard Pareto (infinite support)

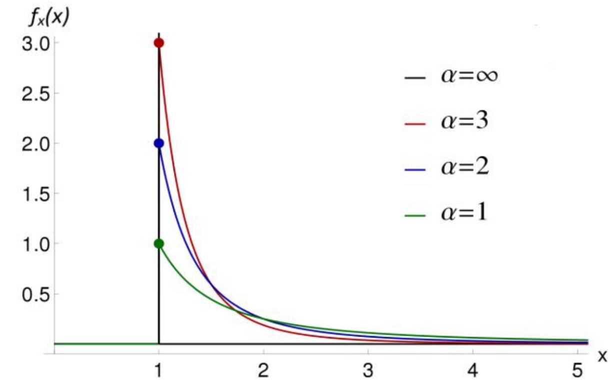
$$\mathrm{E}(X) = \begin{cases} \infty & \alpha \leq 1, \\ \frac{\alpha x_m}{\alpha - 1} & \alpha > 1. \end{cases}$$

$$\mathrm{Var}(X) = \begin{cases} \infty & \alpha \in (1, 2], \\ \left(\frac{x_m}{\alpha - 1}\right)^2 \frac{\alpha}{\alpha - 2} & \alpha > 2. \end{cases}$$

"Wild Behaviour": only diverging variance

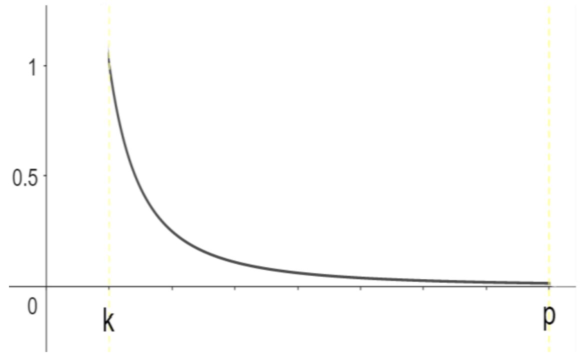
retained in the Euclidean framework

few degrees of freedom



Difficult/meaningless Euclidean extension

Truncated Pareto (finite support)



- More suitable for real dataset, with **an upper limit**
- **More degrees of freedom** in the mean and the variance

Meaningful for a future Euclidean extension

Conclusions

1. We have analysed some **peculiar properties** of heavy tailed distributions and **possible approaches** to exploit them in telecommunication scenarios
2. We have highlighted and proved in Matlab some **issues in simulations** with standard heavy tailed distributions
3. We have implemented in Matlab a mini **object-oriented toolbox (Ban & BanArray classes)** that supports **Euclidean numbers** and their operations, to reduce simulations troubles with ht distributions
4. We have proposed (and simulated in Matlab) a new **Euclidean ht LogNormal distribution** and **numerically checked** the correctcness of the sample mean and variance with respect to their theoretical values

Future Works

Application of Euclidean heavy tailed models to TLC Queues Scheduling algorithms, Load balancing, Routing decisions, etc.

Grazie per l'attenzione!

Francesco Fiorini

francefiorini14@gmail.com

backup slides

Another possible application in TLC

➤ **Queueing Theory** **M/G/1**

Table 5.1 Measures of Effectiveness for the $M/G/1$ Queue

$$L_q = \frac{1+C_B^2}{2}\cdot\frac{\rho^2}{1-\rho} = \frac{\lambda^2 \mathrm{E}[S^2]}{2(1-\rho)} = \frac{\rho^2+\lambda^2\sigma_B^2}{2(1-\rho)}$$

$$W_q = \frac{1+C_B^2}{2}\cdot\frac{\rho}{\mu-\lambda} = \frac{\lambda \mathrm{E}[S^2]}{2(1-\rho)} = \frac{\rho^2/\lambda+\lambda\sigma_B^2}{2(1-\rho)}$$

$$W = \frac{1+C_B^2}{2}\cdot\frac{\rho}{\mu-\lambda}+\frac{1}{\mu} = \frac{\lambda \mathrm{E}[S^2]}{2(1-\rho)}+\frac{1}{\mu} = \frac{\rho^2/\lambda+\lambda\sigma_B^2}{2(1-\rho)}+\frac{1}{\mu}$$

$$L = \frac{1+C_B^2}{2}\cdot\frac{\rho^2}{1-\rho}+\rho = \frac{\lambda^2 \mathrm{E}[S^2]}{2(1-\rho)}+\rho = \frac{\rho^2+\lambda^2\sigma_B^2}{2(1-\rho)}+\rho$$

The product of two LogNormals… is still LogNormal

$$Y = \underbrace{e^{\overbrace{\mu_1+\sigma_1 Z}^{Z_1}}}_{X_1} \cdot \underbrace{e^{\overbrace{\mu_2+\sigma_2 Z}^{Z_2}}}_{X_2} = e^{G}$$

$Z \in N(0,1)$

$G \in N(\mu_{Tot}, \sigma_{Tot}^2)$

$$\mu_{Tot} = \mu_1 + \mu_2$$

$$\sigma_{Tot}^2 = \sigma_1^2 + \sigma_2^2 + 2\,\underbrace{cov(Z_1, Z_2)}_{\rho\,\sigma_1\,\sigma_2}$$

… And is this true even for BANs?

Example 1: product of 2 BANs LogNormals

- Same means
- StdDev with the same real finite part and **different** first order **infinitesimal** part
- Completely anticorrelated: ρ=-1

$$\mu_1 = \mu_2 = 1 + 1\eta$$

$$\sigma_1 = 10^{-3} + 0.8\eta,\ \sigma_2 = 10^{-3} + 0.2\eta$$

By ordering the two samples in *opposite ways*

$$\sigma_1 = a + b\eta + 0\eta^2, \quad \sigma_2 = a + c\eta + 0\eta^2;$$

$$\sigma_1^2 = a^2 + 2ab\eta + b^2\eta^2, \quad \sigma_2^2 = a^2 + 2ac\eta + c^2\eta^2, \quad \sigma_1\sigma_2 = a^2 + ab\eta + ac\eta + bc\eta^2;$$

$$\sigma_{Tot}^2 = a^2 + 2ab\eta + b^2\eta^2 + a^2 + 2ac\eta + c^2\eta^2 - 2(a^2 + ab\eta + ac\eta + bc\eta^2) = (b-c)^2\eta^2$$

$$E\{Y\} =$$
$$= 7.3890 + 14.7781\eta + 16.1081\eta^2$$

$$Var\{Y\} =$$
$$= 0 + 0\eta + 19.6553\eta^2$$

NumerosityOfSample	ComputedMean	ComputedVariance
10^3	$7.38915 + 14.8773\eta + 16.6294\eta^2$	$2.12328\ 10^{-9} + 0.00259586\eta + 20.6137\eta^2$
10^4	$7.38905 + 14.8353\eta + 16.2424\eta^2$	$1.07988\ 10^{-8} - 7.06736\ 10^{-5}\eta + 19.4986\eta^2$
10^5	$7.38902 + 14.7417\eta + 16.0663\eta^2$	$1.62759\ 10^{-9} - 6.42634\ 10^{-5}\eta + 19.7008\eta^2$

Example 1: product of 2 BANs LogNormals

```
% Number of elements of the two samples of random generated numbers
NumerosityOfSample = 1e5;

% Means of the two samples
mu1 = Ban([1 1], 0);
mu2 = Ban([1 1], 0);

% Standard Deviations of the two samples
sig1 = Ban([1e-3 0.8], 0);
sig2 = Ban([1e-3 0.2], 0);

% Correlation coefficient of the two samples
rho = -1;

% Generation of the Gaussian sample with the chosen mean and standard ...
    deviation
X1 = BanArray(randn(NumerosityOfSample,3)) * sig1 + mu1;

% Generation of a Log Normal sample based on the Gaussian one
Z1 = exp(X1);

% Confront theoretical and computed means
TheoreticalSampleMeanZ1 = exp(mu1 + sig1*sig1/2)
ComputedSampleMeanZ1 = mean(Z1)

% Confront theoritcal and computed variances
TheoreticalSampleVarianceZ1 = exp(mu1*2 + sig1*sig1*2) - exp(mu1*2 + ...
    sig1*sig1)
ComputedSampleVarianceZ1 = var(Z1)
```

Example 1: product of 2 BANs LogNormals

```
% Sorting the first sample by ascending order
SortedZ1 = sort(Z1,"ascend");

% Generation of the Gaussian sample with the chosen mean and standard ...
    deviation
X2 = BanArray(randn(NumerosityOfSample,3)) * sig2 + mu2;

% Generation of a Log Normal sample based on the Gaussian one
Z2 = exp(X2);

% Confront theoretical and computed means
TheoreticalSampleMeanZ2 = exp(mu2 + sig2*sig2/2)
ComputedSampleMeanZ2 = mean(Z2)

% Confront theoritcal and computed variances
TheoreticalSampleVarianceZ2 = exp(mu2*2 + sig2*sig2*2) - exp(mu2*2 + ...
    sig2*sig2)
ComputedSampleVarianceZ2 = var(Z2)

% Sorting the second sample by descending order
SortedZ2 = sort(Z2,"descend");

% Computing the product of the two samples
Z = SortedZ1*SortedZ2;

% Computing theoretical mean and variance of the product of the underlying
% Gaussians
VarGauss = sig1*sig1 + sig2*sig2 + sig1*sig2*rho*2;
MeanGauss = mu1 + mu2;
% Confront theoretical and computed means
TheoreticalSampleMeanZ = exp(MeanGauss + VarGauss/2)
ComputedSampleMeanZ = mean(Z)

% Confront theoretical and computed variances
TheoreticalSampleVarianceZ = exp(MeanGauss*2 + VarGauss*2) - ...
    exp(MeanGauss*2 + VarGauss)
ComputedSampleVarianceZ = var(Z)
```

Example 2: sum of 2 BANs Gaussians

- Same **infinite** means!
- StdDev with the **same infinite** part and different η coefficients
- Completely anticorrelated: ρ=-1 → Same previous method

$$\mu_1 = \mu_2 = (100 + 100\eta)\alpha$$

$$\sigma_1 = (1 + 80\eta)\alpha, \ \sigma_2 = (1 + 20\eta)\alpha$$

$$\sigma_1 = a\alpha + b + 0\eta, \quad \sigma_2 = a\alpha + c + 0\eta;$$

$$\sigma_1^2 = a^2\alpha^2 + 2ab\alpha + b^2, \quad \sigma_2^2 = a^2\alpha^2 + 2ac\alpha + c^2, \quad \sigma_1\sigma_2 = a^2\alpha^2 + ab\alpha + ac\alpha + bc;$$

$$\sigma_{Tot}^2 = \sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho = \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2 = (b-c)^2$$

$$\mu_{Tot} = (200 + 200\eta + 0\eta^2)\alpha$$

$$\sigma_{Tot}^2 = 3600 + 0\eta + 0\eta^2$$

NumerosityOfSample	ComputedMean	ComputedVariance
10^3	$(200.053 + 201.899\eta - 2.48293\eta^2)\alpha$	$(0.00447257 + 1.32653\eta + 3798.93\eta^2)\alpha^2$
10^4	$(200.013 + 200.406\eta + 0.035365\eta^2)\alpha$	$(0.000229686 + 0.821994\eta + 3672.7\eta^2)\alpha^2$
10^5	$(200.009 + 200.233\eta + 0.218476\eta^2)\alpha$	$(5.72684\ 10^{-5} - 0.42596\eta + 3572.24\eta^2)\alpha^2$

Example 2: sum of 2 BANs Gaussians

```
% Number of elements of the two samples of random generated numbers
NumerosityOfSample = 1e3;

% Means of the two samples
mu1 = Ban([100 100], 1);
mu2 = Ban([100 100], 1);

% Standard Deviations of the two samples
sig1 = Ban([1 80], 1);
sig2 = Ban([1 20], 1);

% Correlation coefficient of the two samples
rho = -1;

% Generation of the Gaussian sample with the chosen mean and standard ...
    deviation
X1 = BanArray(randn(NumerosityOfSample,3)) * sig1 + mu1;

% Sorting the first sample by ascending order
SortedX1 = sort(X1,"ascend");

% Generation of the Gaussian sample with the chosen mean and standard ...
    deviation
X2 = BanArray(randn(NumerosityOfSample,3)) * sig2 + mu2;

% Sorting the second sample by descending order
SortedX2 = sort(X2,"descend");

% Computing the sum of the two samples (this time the Gaussians are
% considered, so the sum is the one to look at)
X = SortedX1 + SortedX2;

% Confront theoretical and computed means
TheoreticalSampleMeanX = mu1 + mu2
ComputedSampleMeanX = mean(X)
% Confront theoretical and computed variances
TheoreticalSampleVarianceX = sig1*sig1 + sig2*sig2 + sig1*sig2*rho*2
ComputedSampleVarianceX = var(X)
```

Exp of a BAN

```
function eb = exp(b) % exponentiation
            c = b.coeff(1);
            b = b - Ban(b.coeff(1));
            b_square = b*b;
            b_cube = b_square*b;
            eb = (b + b_square/2.0 +  b_cube/6.0 + Ban(1.0)) * exp(c);
        end
```

It extracts the finite part (c) and compute the exp by hand for it; for the remaining BAN (that is infinitesimal: first coefficient, that is the finite part, is now 0), it uses the Mc-Laurin expansion around x_0=0 (the formula below)

Taylor: $$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \cdots + \frac{f^n(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n)$$

Mc-Laurin: $x_0 = 0$ $$f(x) = f(0) + f'(0)\, x + \frac{f''(0)}{2}\, x^2 + \cdots + \frac{f^n(0)}{n!}\, (x)^n + o(x^n)$$

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n)$$

Simulations with standard heavy tailed distributions

Some Issues

2. Slow Convergence & High variability at Steady State ? ➡ **Standard Case**

Do you request t digits of accuracy of your sample mean estimator? ⟷ $$\frac{|A_n - \mu|}{\mu} \leq 10^{-t}$$

With respect to the true value μ

Recalling the CLT formula…

$$|A_n - \mu| = n^{-1/2} d$$

$$n^{-1/2} \leq \frac{1}{v} 10^{-t}$$

$$n^{1/2} \geq v 10^{t}$$

$$n \geq g 10^{2t}$$

α	n
2.0	10^6
1.7	$19.3\ 10^7$
1.5	10^9
1.3	10^{13}
1.2	10^{18}
1.1	10^{33}

n
10^6
10^6
10^6
10^6
10^6
10^6

A practical example with t=3 & Pareto distribution on the left, and standard case on the right

Simulations with standard heavy tailed distributions

Some Issues

2. Slow Convergence & High variability at Steady State

Do you request t digits of accuracy of your sample mean estimator?

With respect to the true value **μ**

Why can it be useful?

In a **queue system simulation**: if you want that:

Measured utilisation coefficient

$$\lambda \bar{x} \to \rho$$

Desired utilisation

Sample Mean of the service times

$\bar{x}$ must be ***close*** to its desired mean value μ, to have **stability**

Therefore the «accuracy» is an important factor

GCLT - α-stable Distributions

Обобщённая предельная теорема

Пусть случайные величины X_i независимы, одинаково распределены и удовлетворяют условиям

$$\mathbb{P}(X > x) \sim a_+ x^{-\alpha} \qquad x \to \infty \qquad a_+ \geq 0$$

$$\mathbb{P}(X < -x) \sim a_- x^{-\alpha} \qquad x \to \infty \qquad a_- \geq 0$$

где $0 < \alpha \leq 2$ и $a_+ + a_- > 0$. Тогда существуют такие последовательности A_n и $B_n > 0$, что при $n \to \infty$

$$\left(\sum_{i=1}^{n} X_i - A_n\right) / B_n \xrightarrow{(d)} S_\alpha(1, \beta, 0) \qquad \text{где} \quad \beta = \frac{a_+ - a_-}{a_+ + a_-}$$

A_n и B_n могут быть определены следующим образом

при $\alpha = 2$	$A_n = na$	$B_n = \sqrt{cn \ln n}$
при $\alpha \in (1, 2)$	$A_n = na$	$B_n = [\pi nc/(2\Gamma(\alpha)\sin(\alpha\pi/2))]^{1/\alpha}$
при $\alpha = 1$	$A_n = \beta cn \ln n$	$B_n = \pi nc/2$
при $\alpha \in (0, 1)$	$A_n = 0$	$B_n = [\pi nc/(2\Gamma(\alpha)\sin(\alpha\pi/2))]^{1/\alpha}$

где $a = \mathbb{E}X$ и $c = a_+ + a_-$

α-stable Distributions

α stable distributions

☞ A random variable X is said to have a stable distribution (indicated as $X \sim S_\alpha(\sigma, \beta, \mu)$) if there are parameters $0 < \alpha \leq 2, \sigma \geq 0, -1 \leq \beta \leq 1$, and $\mu \in \mathbb{R}$ such that its characteristic function has the following form

$$\mathbb{E}e^{(i\lambda X)} = \begin{cases} \exp\{-\sigma^\alpha|\lambda|^\alpha \left[1 - i\beta \tan(\frac{\pi\alpha}{2}\mathrm{sgn}(\lambda))\right] + i\mu\lambda\} & \alpha \neq 1 \\ \exp\{-\sigma|\lambda| \left[1 + i\beta\frac{2}{\pi}\mathrm{sgn}(\lambda) \log|\lambda|\right] + i\mu\lambda\} & \alpha = 1 \end{cases}$$

- α is the index of stability, related to the weight of the tails of the distribution function
  - Gaussian distribution if $\alpha = 2$: $X \sim S_2(\sigma, \beta, \mu)$ is equivalent to $X \sim N(\mu, 2\sigma^2)$
  - Cauchy distribution if $\alpha = 1$; if $X \sim S_1(\sigma, 0, \mu) \quad \Rightarrow$

$$f(x) = \frac{1}{\pi} \frac{\sigma}{\sigma^2 + (x - \mu)^2}$$

- β is the skewness parameter, related to the shape of the tails of the distribution function
  - $\beta = 0$ in the case of symmetric distribution
- σ is the dispersion parameter, related to the spread of the distribution around its location parameter μ, similar to the variance of the Gaussian distribution
- μ is the location parameter
  - mean if $1 < \alpha \leq 2$
  - median if $0 < \alpha \leq 1$

α-stable Distributions

Important properties of α-stable distributions

- Let $X \sim S_\alpha(\sigma, \beta, \mu)$ with $0 < \alpha < 2$. Then

$$\begin{cases} \lim_{x\to\infty} x^\alpha \mathbb{P}(X > x) = C_\alpha \dfrac{1+\beta}{2}\sigma^\alpha \\ \lim_{x\to\infty} x^\alpha \mathbb{P}(X < -x) = C_\alpha \dfrac{1-\beta}{2}\sigma^\alpha \end{cases}$$

where

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx\right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & \alpha \neq 1 \\ 2/\pi & \alpha = 1 \end{cases}$$

- Let $X \sim S_\alpha(\sigma, \beta, \mu)$ with $0 < \alpha < 2$. Then

$$\mathbb{E}|X|^p < \infty \quad \text{for} \quad 0 < p < \alpha$$
$$\mathbb{E}|X|^p = \infty \quad \text{for} \quad p \geq \alpha$$

- A symmetric α-stable distribution behaves approximately like a Gaussian around its origin
- but for $\alpha < 2$ the α-stable distribution is heavy-tailed (more precisely power-law tailed)
- The lower the characteristic exponent α the heavier the tails of the α-stable distribution
- α-stable distributions are used to model phenomenons which are impulsive in nature

α-stable Distributions

☞ Central Limit Theorem

Let $X_1,\ X_2,\ldots$ *be a sequence of IID RVs with finite mean* a *and variance* σ^2*. Then*

$$Z_n \triangleq \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} X_i - n\,a\right) \stackrel{(d)}{\rightarrow} \mathcal{N}\left(0,\sigma^2\right)$$

☞ Generalized Central Limit Theorem: the family of stable distributions contains all limiting distributions of sums of IID random variables

$$Z_n \triangleq \frac{1}{n^{1/\alpha}}\left(\sum_{i=1}^{n} X_i - n\,a\right) \stackrel{(d)}{\rightarrow} L_\alpha$$

- L_α is an α–stable distribution, with index of stability α
- α is the critical order of convergence of the moments of X, i.e., $\forall\, q \geq \alpha\ \mathbb{E}X^q = \infty$

Stable Distribution: *def*

Let X_1 and X_2 be independent realizations of a random variable X. Then X is said to be **stable** if for any constants $a > 0$ and $b > 0$ the random variable $aX_1 + bX_2$ has the same distribution as $cX + d$ for some constants $c > 0$ and d. The distribution is said to be ***strictly stable*** if this holds with $d = 0$.[7]

α-stable case

$$\varphi(t;\alpha,\beta,c,\mu) = \exp\left(it\mu - |ct|^\alpha \left(1 - i\beta\,\mathrm{sgn}(t)\Phi\right)\right)$$

where $\mathrm{sgn}(t)$ is just the sign of t and

$$\Phi = \begin{cases} \tan\left(\frac{\pi\alpha}{2}\right) & \alpha \neq 1 \\ -\frac{2}{\pi}\log|t| & \alpha = 1 \end{cases}$$

The reason this gives a stable distribution is that the characteristic function for the sum of two independent random variables equals the product of the two corresponding characteristic functions. Adding two random variables from a stable distribution gives something with the same values of α and β, but possibly different values of μ and c.

GCLT

A generalization due to Gnedenko and Kolmogorov states that the sum of a number of random variables with symmetric distributions having power-law tails (Paretian tails), decreasing as $|x|^{-\alpha-1}$ where $0 < \alpha \leqslant 2$ (and therefore having infinite variance), will tend to a stable distribution $f(x;\alpha,0,c,0)$ as the number of summands grows.[14] If $\alpha > 2$ then the sum converges to a stable distribution with stability parameter equal to 2, i.e. a Gaussian distribution.[15]

Truncated Pareto

$$F_X(x;k,p,\alpha) = \frac{1-\left(\frac{k}{x}\right)^\alpha}{1-\left(\frac{k}{p}\right)^\alpha}$$

$$f_X(x;k,p,\alpha) = \frac{\alpha\, k^\alpha}{1-\left(\frac{k}{p}\right)^\alpha}\, x^{-\alpha-1}$$

α≠1

$$E\{X\} = \frac{\alpha k}{\alpha-1}\,\frac{1-\left(\frac{k}{p}\right)^{\alpha-1}}{1-\left(\frac{k}{p}\right)^\alpha}$$

$$E\{X\} = \frac{k}{1-\left(\frac{k}{p}\right)} \int_k^p \frac{1}{x}\, dx = \frac{k}{1-\left(\frac{k}{p}\right)} \ln\frac{p}{k}$$

α≠2

$$Var\{X\} = E\left\{(X-E(X))^2\right\} = E\{X^2\} - E\{X\}^2 = \frac{\alpha\, k^2}{\alpha-2}\,\frac{1-\left(\frac{k}{p}\right)^{\alpha-2}}{1-\left(\frac{k}{p}\right)^\alpha} - E\{X\}^2$$

$$E\{X^2\} = \frac{2\,k^2}{1-\left(\frac{k}{p}\right)^2} \int_k^p x^{-1}\, dx = \frac{2\,k^2}{1-\left(\frac{k}{p}\right)^2} \ln\frac{p}{k}$$

Another possible application in TLC

- **Queueing Theory** M/G/1

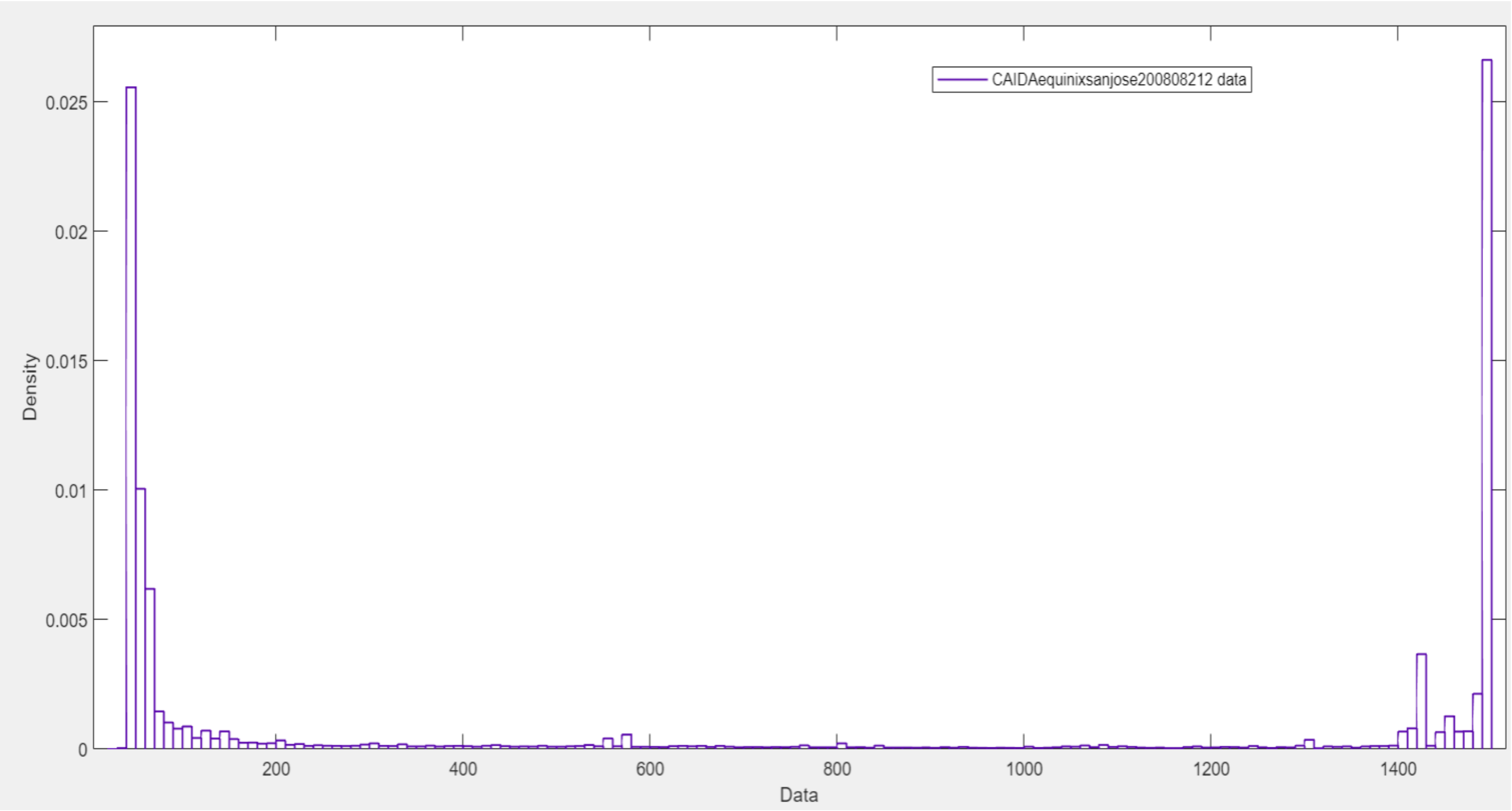
Table 5.1 Measures of Effectiveness for the $M/G/1$ Queue

$$L_q = \frac{1+C_B^2}{2} \cdot \frac{\rho^2}{1-\rho} = \frac{\lambda^2 \mathrm{E}[S^2]}{2(1-\rho)} = \frac{\rho^2 + \lambda^2\sigma_B^2}{2(1-\rho)}$$

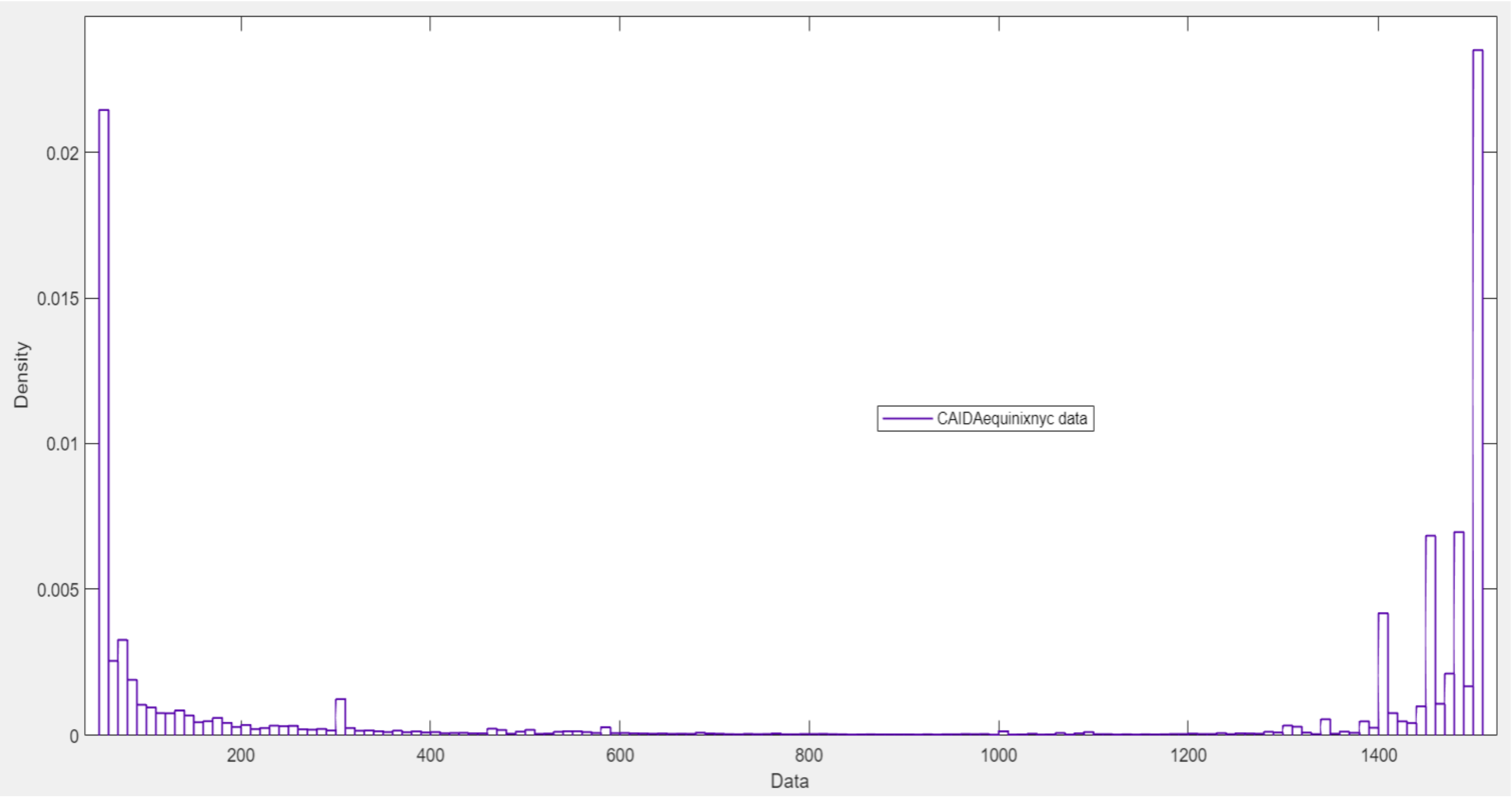
$$W_q = \frac{1+C_B^2}{2} \cdot \frac{\rho}{\mu-\lambda} = \frac{\lambda \mathrm{E}[S^2]}{2(1-\rho)} = \frac{\rho^2/\lambda + \lambda\sigma_B^2}{2(1-\rho)}$$

$$W = \frac{1+C_B^2}{2} \cdot \frac{\rho}{\mu-\lambda} + \frac{1}{\mu} = \frac{\lambda \mathrm{E}[S^2]}{2(1-\rho)} + \frac{1}{\mu} = \frac{\rho^2/\lambda + \lambda\sigma_B^2}{2(1-\rho)} + \frac{1}{\mu}$$

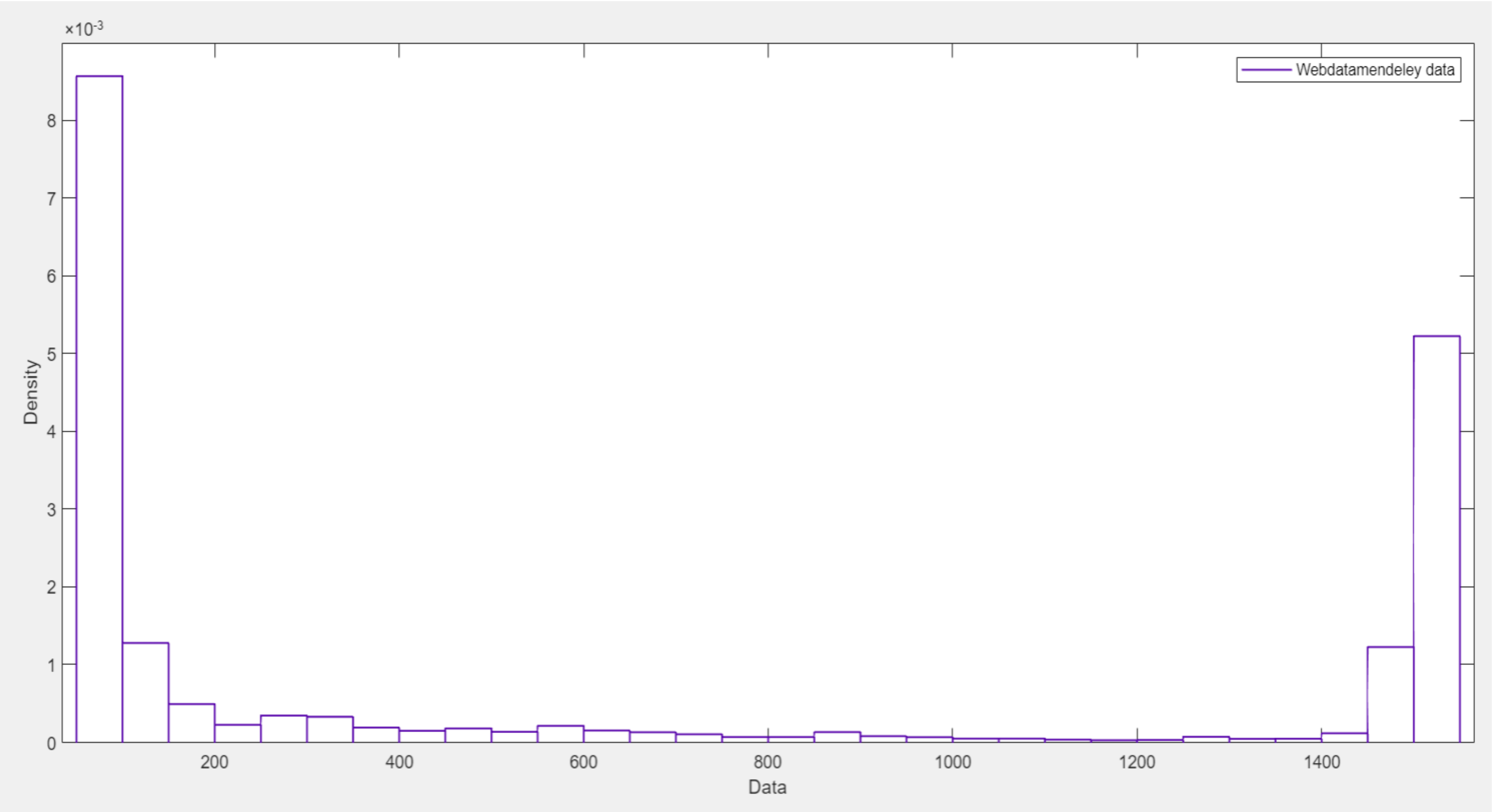
$$L = \frac{1+C_B^2}{2} \cdot \frac{\rho^2}{1-\rho} + \rho = \frac{\lambda^2 \mathrm{E}[S^2]}{2(1-\rho)} + \rho = \frac{\rho^2 + \lambda^2\sigma_B^2}{2(1-\rho)} + \rho$$



CAIDAequinixsanjose200808212 data
Density
0.025
0.02
0.015
0.01
0.005
0
200
400
600
800
1000
1200
1400
Data



0.02
0.015
0.01
0.005
0
Density
CAIDAequinixnyc data
200
400
600
800
1000
1200
1400
Data


×10-3
Webdatamendeley data
Density
Data
0
1
2
3
4
5
6
7
8
200
400
600
800
1000
1200
1400

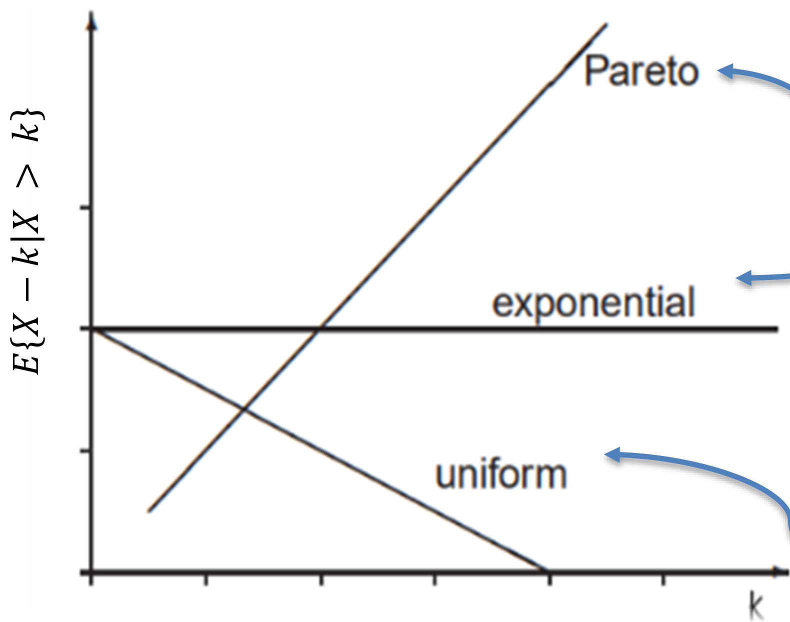
Two important properties: (1) Expectation Paradox

$$E\{X - k | X > k\} \sim k$$

$$= \int_k^\infty (x-k)\,\frac{f_X(x)}{\int_k^\infty f_X(q)\,dq}\,dx = \frac{\int_k^\infty x\, f_X(x)\,dx}{\int_k^\infty f_X(q)\,dq} - k$$

"The longer we have waited…
the longer we should expect to wait!"

3 Different Examples



$$E\{X - k|X > k\} = \frac{\int_k^\infty \frac{\alpha x}{x^{\alpha+1}}dx}{\int_k^\infty \frac{\alpha}{q^{\alpha+1}}dq} - k = \frac{\frac{1}{(\alpha-1)k^{\alpha-1}}}{\frac{1}{\alpha k^{\alpha}}} - k = \frac{k}{\alpha - 1}$$

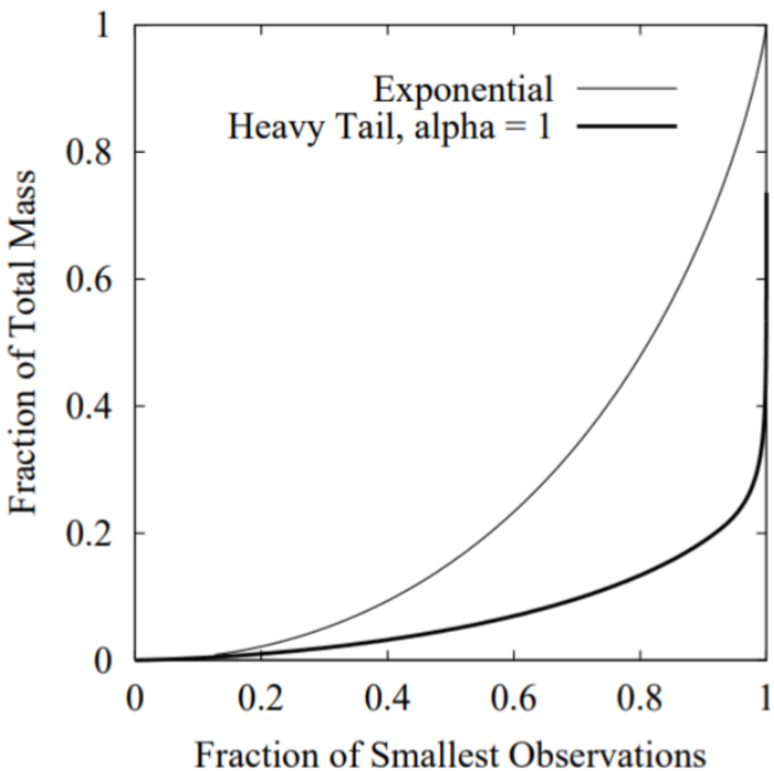
$$E\{X - k|X > k\} = \frac{\int_k^\infty x\, e^{-\frac{x}{\beta}}dx}{\int_k^\infty e^{-\frac{q}{\beta}}\,dq} - k = \frac{\beta e^{-\frac{k}{\beta}}(k+\beta)}{\beta e^{-\frac{k}{\beta}}} - k = \beta$$

$$E\{X - k|X > k\} = \frac{\int_k^{2b} x\,\frac{1}{2b}\,dx}{\int_k^{2b} \frac{1}{2b}\,dq} - k = \frac{\frac{1}{2}\left(4b^2 - k^2\right)}{2b - k} - k = \frac{1}{2}\,(2b - k)$$

Two important properties: (2) Mass Count Disparity

$$\lim_{x \to \infty} \frac{P[X_1 + \ldots + X_n > x]}{P[\max[X_1, \ldots, X_n] > x]} = 1 \quad for\ all\ n \geq 2$$

"A very tiny subset of observations contains the vast bulk of the mass in a set of data"



- 60% of the mass is contained in the top 1% of the observations, which is completely out of proportion to the fraction of observations taken into account!
- 80% of the smallest observations represent less than 20% of the total mass

TLC real-world example

50-80% of the bytes in FTP transfers are due to the largest 2% of all transfers

BANs operations

Sum (assuming $p \geq q$):

$$\begin{aligned} \xi + \zeta &= \alpha^p P(\eta) + \alpha^p \left(Q(\eta)\, \eta^{p-q} \right) \\ &= \alpha^p \left(P(\eta) + \mathrm{tr}_n \left[Q(\eta)\, \eta^{p-q} \right] \right). \end{aligned}$$

Product:

$$\xi\zeta = \alpha^{p+q}\, \mathrm{tr}_n \left[P(\eta) \cdot Q(\eta) \right].$$

Division:

After having rewritten ζ as:

$$\zeta = \alpha^q \left(q_0 + \sum_{k=1}^{n} q_k \eta^k \right) = q_0 \alpha^q \left(1 - \varepsilon \right),$$

where $\varepsilon = -\sum_{k=1}^{n} \frac{q_k}{q_0} \eta^k$, the definition of the division becomes:

$$\begin{aligned} \frac{\xi}{\zeta} &= \alpha^{p-q}\, \mathrm{tr}_n \left[\frac{P(\eta)}{q_0} \left(1 + \varepsilon + \varepsilon^2 + \ldots + \varepsilon^n \right) \right] \\ &= \alpha^{p-q} \left(\frac{P(\eta)}{q_0} + \mathrm{tr}_n \left[\varepsilon \frac{P(\eta)}{q_0} \right] + \ldots + \mathrm{tr}_n \left[\varepsilon^n \frac{P(\eta)}{q_0} \right] \right). \end{aligned}$$

Backup4

```
function avg = mean(bArr)
    if size(bArr,2) > 1
        error('To Be Implemented!');
    end
    sum = bArr(1).bArr;
    for i = 2:size(bArr,1)
        sum = sum + bArr(i).bArr;
    end
    avg = sum/length(bArr);
end % mean

function v = var(bArr)
    if size(bArr,2) > 1
        error('To Be Implemented!');
    end
    avg = mean(bArr);
    slacks = bArr-avg;
    squared_slacks = slacks*slacks;
    v = mean(squared_slacks);
end % var
```

```
function abs_bArr = abs(bArr1)
            if size(bArr1,2) > 1
                error('To Be
Implemented!');
            end
            abs_bArr = bArr1;
            for i = 1:size(bArr1,1)
                if (
bArr1(i).bArr.coeff(1) < 0 )
                    abs_bArr(i).bArr =
Ban(-
bArr1(i).bArr.coeff,bArr1(i).bArr.lexp);
                end
            end
end % abs
```

Backup6

Heavy tail definition

Considering its moment-generating function:

$$M_X(t) = E\left\{e^{tX}\right\}$$

a distribution is said to be heavy tailed if, for all $t > 0$,

$$M_X(t) = E\left\{e^{tX}\right\} \to \infty.$$

More details on this definition of heavy tailedness can be found in Konstantinides [2018]; Bianchi et al. [2019].
An implication of this is that:

$$\lim_{x \to \infty} e^{tx} \bar{F}_X(x) = \infty \qquad \forall t > 0.$$

Backup7

Heavy tail definition

Another possible definition of heavy tailed distributions that can be found in the literature, that actually refers to asymptotically power-law distributions, is the following: Let X be a random variable with CDF $F_X(x)$ and survival function $\overline{F}_X(x)$. A heavy tailed distribution is said to exist if:

$$\overline{F}_X(x) \sim L(x)x^{-\epsilon}$$

where $f(x) \sim a(x)$ means that $\lim_{x\to\infty} f(x)/a(x) = c$ for some positive constant c and with $L(x)$ a slowly varying function at infinity, i.e. for all positive x, $\lim_{\tau\to\infty} L(\tau x)/L(\tau) = 1$.

The case

$$1 < \epsilon \leq 2$$

is of special interest and concerns heavy tailed distributions with finite mean but infinite variance. If

$$\epsilon \leq 1,$$

X has infinite mean.

Heavy tail definition – HILL'S ESTIMATOR

An alternative, more rigorous, analytical method to estimate the intensity of the Noah effect, i.e. the value of the parameter ϵ of a heavy tailed distribution (according to the definition presented in section 5), is to use the Hill estimator. Let

$$U_1, U_2, ..., U_n$$

denote, for example, the observed data; let them then be written in

$$U_{(1)} \geq U_{(2)} \geq ... \geq U_{(n)}$$

form, i.e. through the ordered statistics; the Hill estimator of ϵ is given by:

$$\hat{\epsilon}_k = \left(\frac{1}{k} \sum_{i=1}^{i=k} \left(\log\ U_{(i)} \ - \ \log\ U_{(k)} \right) \right)^{-1},$$

where the choice of

$$1 < k \leq n$$

indicates how many of the largest observations enter into the calculation of the formula. In practice, we plot the estimator as a function of k, for a certain range of values. In the presence of tail behaviour in the data, the plot will vary considerably for small values of k, since only a small fraction of the largest data is so considered, but will become more stable as more points in the distribution are included.

Heavy tail definition - LOGNORMAL

$$\mathbb{E}\left(X^t\right) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$$

The skewness and kurtosis of X are

1. $\text{skew}(X) = \left(e^{\sigma^2} + 2\right)\sqrt{e^{\sigma^2} - 1}$
2. $\text{kurt}(X) = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 3$

$\mathbb{E}\left(e^{tX}\right) = \infty$ for every $t > 0$.

Proof

By definition, $X = e^Y$ where Y has the normal distribution with mean μ and standard deviation σ. Using the change of variables formula for expected value we have

$$\mathbb{E}\left(e^{tX}\right) = \mathbb{E}\left(e^{te^Y}\right) = \int_{-\infty}^{\infty} \exp(te^y)\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]dy = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp\left[te^y - \frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]dy \tag{5.12.11}$$

If $t > 0$ the integrand in the last integral diverges to ∞ as $y \to \infty$, so there is no hope that the integral converges.

Heavy tail definition - HAZARD RATE -LogNormal

Suppose that X represents the life of some item, with the distribution function $F_X(x)$. The function defined by

$$\boxed{h_X(t) = \frac{f_X(t)}{S_X(t)} = \frac{f_X(t)}{1 - F_X(t)}} \tag{6.36}$$

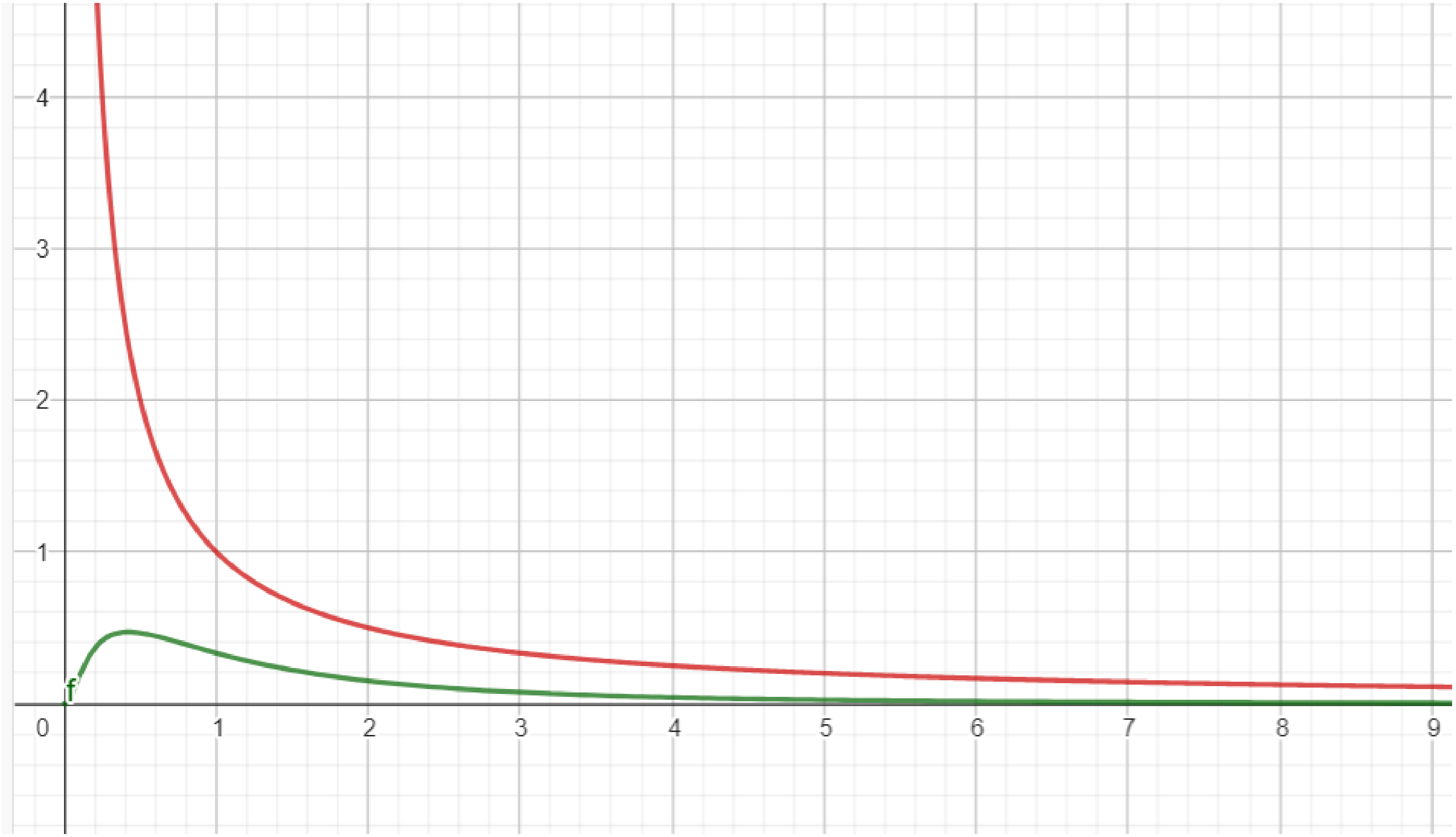
is called the **hazard function** or the **failure rate**, because $h_X(t)\,dt$ represents the probability that the life will end in the interval $(t, t + dt]$, given that X has survived up to age t; i.e., $X \geq t$. If X represents the service time of a customer, as in queueing theory, $h_X(t)$ is called the **completion rate function**.

The hazard functions of the exponential, **Weibull**, **Pareto**, and **log-normal** distributions are given as follows:

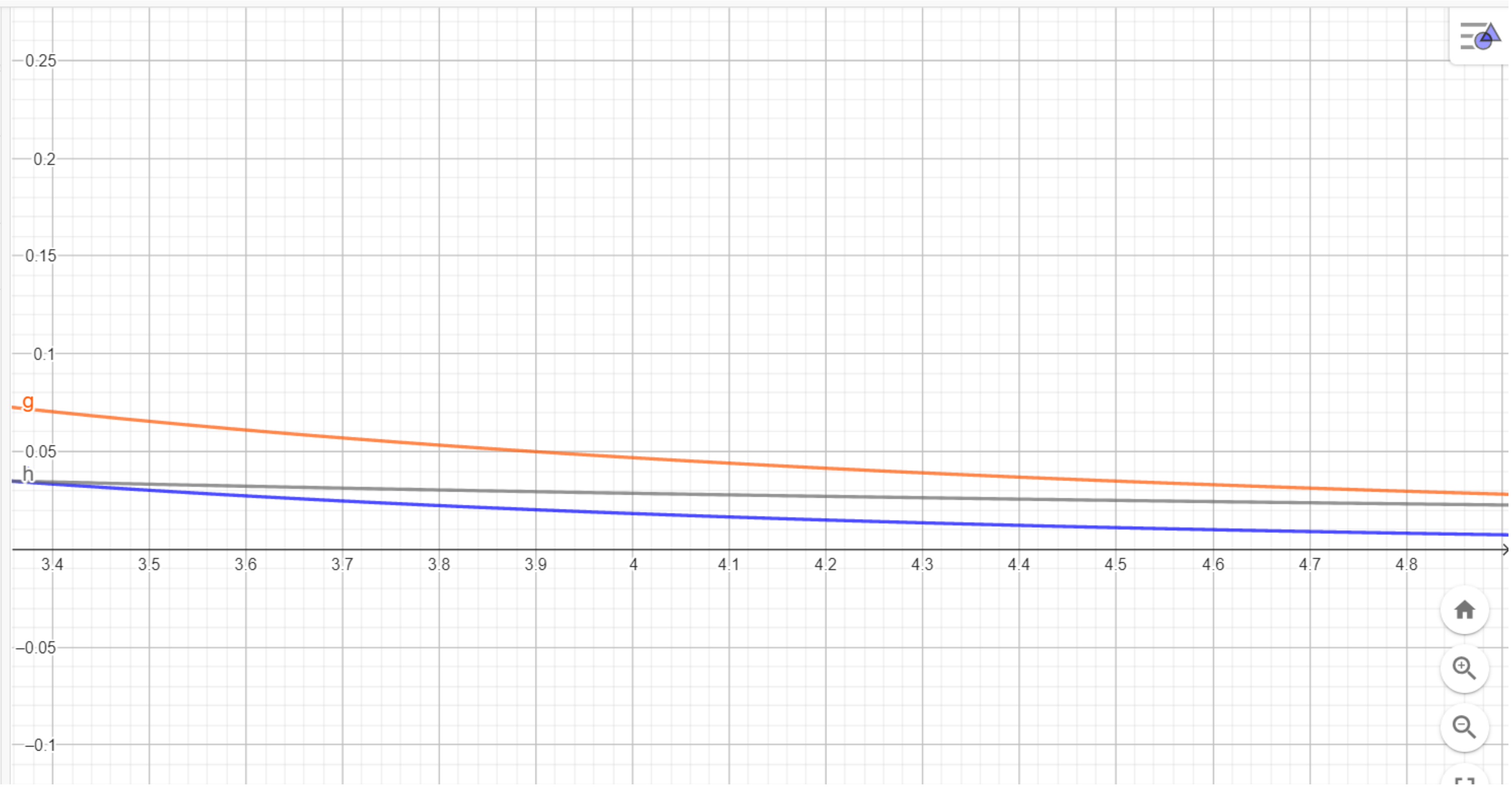
$$h_X(t) = \begin{cases} \lambda, & t \geq 0, \quad \text{for exponential,} \\ \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1}, & t \geq 0, \quad \text{for Weibull,} \\ \frac{\alpha}{t}, & t \geq \beta, \quad \text{for Pareto,} \\ \dfrac{t^{-1}\exp\left[-\frac{(\log t - \mu_Y)^2}{2\sigma_Y^2}\right]}{\int_{\log t}^{\infty}\exp\left[-\frac{(u-\mu_Y)^2}{2\sigma_Y^2}\right]du}, & t > 0, \quad \text{for log-normal,} \end{cases} \tag{6.37}$$

Backup7

Heavy tail definition - HAZARD RATE – LogNormal(green) & Pareto(red)



Heavy tail definition - HEAVY TAIL? PDF tails in order: Pareto, LogNormal and Exp



Backup7

Heavy tail definition - LOGNORMAL shadowing

Now let us discuss how the log-normal RV appears in the signal propagation in a radio channel. Consider the signal power (or signal strength) at the receiver. It should be the signal power sent from the transmitter divided by the attenuation or loss factor L (>1) due to propagation loss. If the propagation is in free space, then $L = 4\pi d^2$, where d is the distance between the transmitter and the receiver. In practice, there are additional components such as absorption of signals in trees, buildings, and other objects, and these lossy components will vary. Thus, it is proper to treat L as a random variable. Furthermore, if we divide the path between the transmitter and receiver into contiguous and disjoint segments, then the overall loss L is the product of the loss within each segment:

$$L = \prod_{i=1}^{n} L_i$$

It is reasonable to assume that in most cases these RVs' L_i are statistically independent. Of course, the mean values of the L_i may be commonly affected by such factors as the temperature, precipitation, and so forth, but the variation of L_i from its mean should be unrelated to that of L_j; hence, L_i and L_j are statistically independent for $j \neq i$. Taking the logarithm of (7.54), we have

$$Y = \sum_{i=1}^{n} Y_i, \tag{7.55}$$

where we set

$$Y = \ln L \text{ and } Y_i = \ln L_i, \ \text{ for } i = 1, 2, \ldots, n.$$

The transformed RVs $Y_1, Y_2, \ldots, Y_n$ are statistically independent because L_i are independent. We do not require the assumption that they are statistically identical to each other, because a generalized version of the CLT, as stated in Theorem 11.23 of Section 11.3.4, does not require the identical distribution assumption. Assume that the Y_i have finite mean μ_i and variance σ_i^2. Then, from the CLT, we can show that Y is asymptotically (i.e., as $n \to \infty$) normally distributed according to $N(\mu_Y, \sigma_Y^2)$, where $\mu_Y = \sum_{i=1}^{n} \mu_i$ and $\sigma_Y^2 = \sum_{i=1}^{n} \sigma_i^2$, as long as none of the σ_i^2 represent a significant portion of their sum σ_Y^2. Therefore, the overall attenuation factor is log-normally distributed.

SITA-E - Why Finite Mean needed

More precisely, let $F(x) = \Pr\{X \leq x\}$ denote the cumulative distribution function of task sizes with finite mean M. Let k denote the smallest task size, p (possibly equal to infinity) denote the largest task size, and h be the number of hosts. Then we determine "cutoff points" x_i, $i = 0 \ldots h$ where $k = x_0 < x_1 < x_2 < \ldots < x_{h-1} < x_h = p$, such that

$$\int_{x_0=k}^{x_1} x \cdot dF(x) = \int_{x_1}^{x_2} x \cdot dF(x) = \cdots = \int_{x_{h-1}}^{x_h=p} x \cdot dF(x) = \frac{M}{h} = \frac{\int_k^p x \cdot dF(x)}{h}$$

and assign to the ith host all tasks ranging in size from x_{i-1} to x_i.

SITA-E as defined can be applied to *any* task size distribution with finite mean. In the remainder of this case study we will always assume the task size distribution is the Bounded Pareto distribution, $B(k, p, \alpha)$.

Random : an incoming task is sent to host i with probability $1/h$. This policy equalizes the expected number of tasks at each host.

Round-Robin : tasks are assigned to hosts in cyclical fashion with the ith task being assigned to host i mod h. This policy also equalizes the expected number of tasks at each host, and typically has less variability in interarrival times than Random.

Size-Based : Each host serves tasks whose service demand falls in a designated range. This policy attempts to keep small tasks from getting "stuck" behind large tasks.

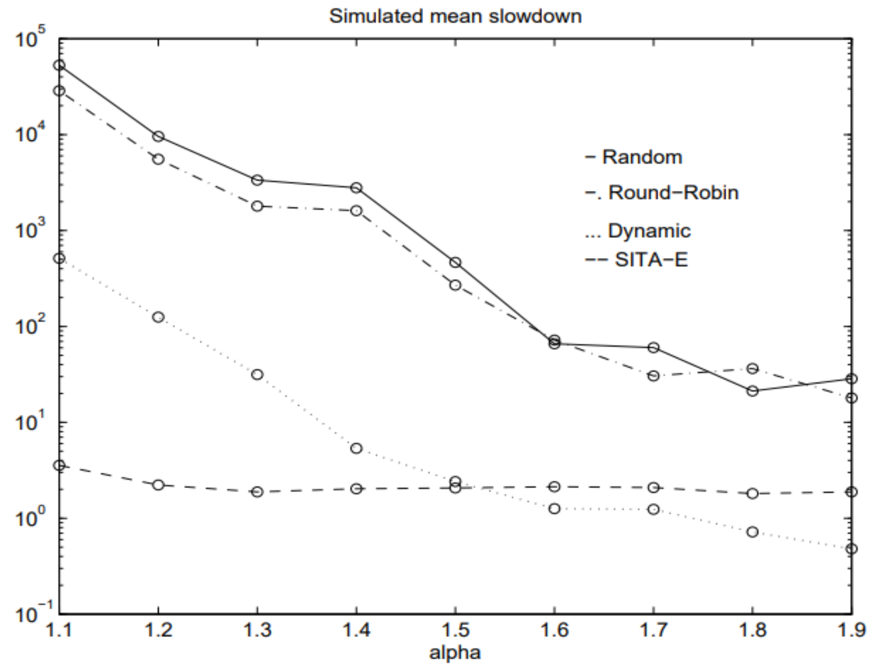
Dynamic : Each incoming task is assigned to the host with the smallest amount of outstanding work, which is the sum of the sizes of the tasks in the host's queue plus the work remaining on that task currently being served. This policy is optimal from the standpoint of an individual task, and from a system standpoint attempts to achieve instantaneous load balance.

DEFINITION 1.1. *For any given policy, the slowdown, S, is defined as response time divided by job size, namely,* $S = \frac{T(X)}{X}$. *The slowdown for a job of size x, S(x), is thus given by*

$$S(x) = \frac{T(x)}{x}.$$

The expected slowdown for a job of size x, E[S(x)], is given by

$$E[S(x)] = \frac{E[T(x)]}{x}.$$



SRPT

SRPT: Starvation of Large Jobs? No!

Consider a job in the 99th percentile of the job size distribution (i.e. a very large job). It turns out that such a job has lower expected slowdown when the scheduling policy is SRPT-like than under a fair scheduling (Processor-Sharing) type of policy. To see this, recall from Section 2 that the sizes of requests arriving at a Web server have been shown to have a heavy-tailed distribution. **Now consider a job j in the 99th percentile of the job size distribution**. By the heavy-tailed property (see Section 2), more than half the total workload is contained in jobs of size greater than j. Thus job j is preempted by less than half the total workload, which in turn implies (see [8]) that j's expected response time is actually better under SRPT-like scheduling than under a Processor-Sharing type of scheduling where job j would have to share the resource with the total workload. By contrast**, in the case of an exponential distribution only 5% of the total workload is contained in jobs of size greater than j.** Thus under an exponential workload, job j would be held up by over 95% of the workload and would in fact have signicantly worse performance under an SRPT-like scheduling policy than under a processor-sharing-like scheduling policy. Thus for an exponential workload, SRPT-like scheduling is not a good idea

Backup

- **Definition 1. Ordinary Set** A set A is an ordinary set ⇐==⇒ A satisfies one of the following properties:
  - i. A = N
  - ii. A = P(B), where B is an ordinary set
  - iii. A = $\bigcup_{i \in I} B_i$, where I and all B_i are ordinary sets
  - iv. A = f(B), where B is an ordinary set and f any function
- **Definition: Ordered Field K**. ∀ a,b,c ∈ K, if a ≤ b, then a+c ≤ b+c and if 0 ≤a, 0 ≤b, then 0 ≤a.b
- **Pareto distribution**:

$$E(X) = \mu = \begin{cases} \infty & if\ \alpha \le 1 \\ \frac{\alpha x_m}{\alpha - 1} & if\ \alpha > 1 \end{cases} \qquad \bar{F}_X(x) = \begin{cases} \left(\frac{x_m}{x}\right)^{\alpha} & if\ x \ge x_m \\ 1 & if\ x < x_m \end{cases}$$

$$\operatorname{var}(X) = E\left\{(X-\mu)^2\right\} = \int_{xm}^{\infty} (x-\mu)^2 f_X(x)\,dx = \begin{cases} \infty & if\ \alpha \in (1,2] \\ \left(\frac{x_m}{\alpha-1}\right)^2 \frac{\alpha}{\alpha-2} & if\ \alpha > 2 \end{cases}$$

ition, it's easy to find the CDF (cumulative di

$$F_X(x) = \begin{cases} 1 - \left(\frac{x_m}{x}\right)^{\alpha} & if\ x \ge x_m \\ 0 & if\ x < x_m \end{cases}$$

$$\mu_n = E\{X^n\} = \begin{cases} \infty & if\ \alpha \le n \\ \frac{\alpha {x_m}^n}{\alpha - n} & if\ \alpha > n \end{cases}$$

ast function, it follows that the probability den

$$\log f_X(x) = \log\left(\alpha \frac{{x_m}^{\alpha}}{x^{\alpha+1}}\right) = \log\left(\alpha {x_m}^{\alpha}\right) - (\alpha+1)\log x$$

$$f_X(x) = \begin{cases} \frac{\alpha {x_m}^{\alpha}}{x^{\alpha+1}} & if\ x \ge x_m \\ 0 & otherwise \end{cases}$$

Backup3

- The 2.37 first-order infinitesimal as aBan3: γ−1(2.37 + 0η+ 0η2+ 0η3) =γ−1[2.37 0 0 0] =(-1)[2.37 0 0 0] =“2.37Ban3infmal” Observe how an equivalent representation would be this one: γ0(0 + 2.37η+ 0η2+ 0η3) =γ0[0 2.37 0 0] =(0)[0 2.37 0] but the latter is not in normal Ban representation.
- The 4.38 first-order infinite as aBan3: γ1(4.38 + 0η+ 0η2+ 0η3) =γ1[4.38 0 0 0] =(1)[4.38 0 0 0] =“4.38Ban3INFTY
- **Log Normal:**

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Pr(X \le x) = \frac{\mathrm{d}}{\mathrm{d}x}\Pr(\ln X \le \ln x) = \frac{\mathrm{d}}{\mathrm{d}x}\Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$

$$F_X(x) = \Phi\left(\frac{(\ln x) - \mu}{\sigma}\right)$$

$$= \varphi\left(\frac{\ln x - \mu}{\sigma}\right)\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\ln x - \mu}{\sigma}\right) = \varphi\left(\frac{\ln x - \mu}{\sigma}\right)\frac{1}{\sigma x}$$

$$= \frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right).$$

- Self similar

$$X = (X_t\,;\ t = 0, 1, 2\ \ldots),$$

$$r(k) \sim k^{-\beta} \quad as\ k \to \infty,$$

$$H = 1 - (\beta/2)$$

te series

$$X^{(m)} = \left(X_k^{(m)}\ :\ k = 1, 2, 3, \ldots\right),$$

$$0 < \beta < 1.$$

$$\frac{1}{2} < H < 1$$

original series X between non-overl

t has the same autocorrelation functi

$$r(k) = E\{(X_t)\ (X_{t+k})\}$$

Backup5

Self similar

one defines the m-aggregate series

$$X^{(m)} = \left(X_k^{(m)} : k = 1, 2, 3, \ldots\right),$$

obtained by summing the original series X between non-overlapping blocks of length m. Thus, if X is self similar, it has the same autocorrelation function

$$r(k) = E\left\{(X_t)\ (X_{t+k})\right\}$$

as the series $X^{(m)}$, for all values of m. This implies that the series is distributionally self similar: the distribution of the aggregate series is the same, except for changes in time scale, as the original one.

As a consequence, a self-similar process exhibits the so-called long-range dependence, i.e. it has an autocorrelation function of the form

$$r(k) \sim k^{-\beta} \quad as\ k \to \infty,$$

where

$$0 < \beta < 1.$$

Thus it decays hyperbolically, slow by comparison with the exponential trend shown by traditional traffic models.

An interesting feature of using self similar models for time series, is that the degree of LRD (long range dependence) or self similarity, is expressed using a single parameter, which expresses the decay rate of the autocorrelation function. Its name is Hurst's parameter, defined as

$$H = 1 - (\beta/2)$$

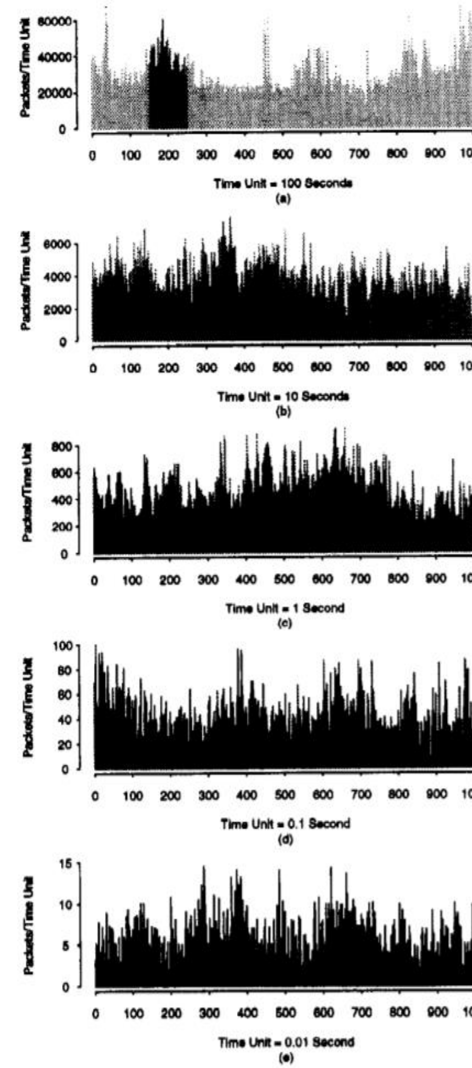
Thus, for self similar processes,

$$\frac{1}{2} < H < 1$$

as H approaches the value 1, the degree of self similarity increases, since as β decreases towards zero, the hyperbolic trend of the tails of the autocorrelation function tends to increase in level and become slower, leading to a LRD effect on increasingly larger leg times.

Heavy tailed distributions in TLC scenarios

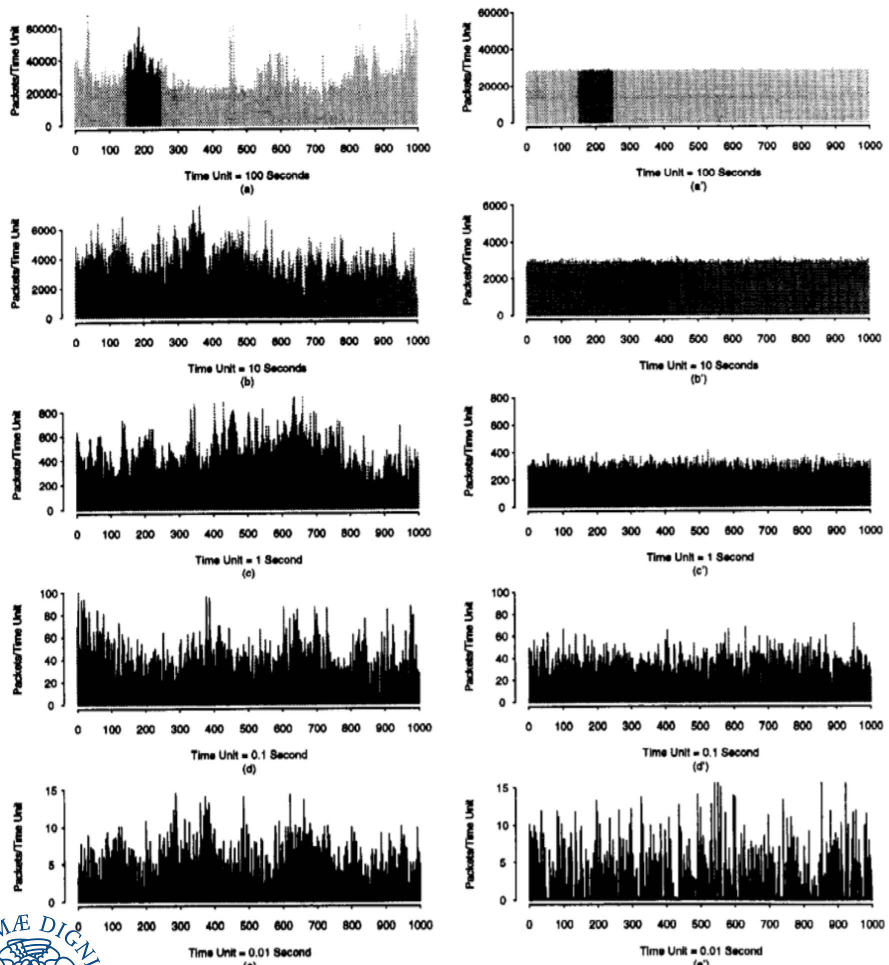
- Observed h-t behaviour in telecommunication systems, respect to quantities such as file sizes on a web server, uptime and silence times in remote communications, CPU times, peak rates, connection times.
- HT distributions can have "infinite" variance ➡ **Noah Effect** ➡ **"high variability"**
- Superimposing many ON/OFF source models, of which at least one with a HT distribution, with infinite variance for the length of the singles On and Off periods ➡ self-similarity traces in the collective traffic: Noah effect as a physical cause of the **Joseph Effect** or **self-similarity**.
- Measured traffic rates, in LAN environments ➡ scaling properties over a wide range of time scales

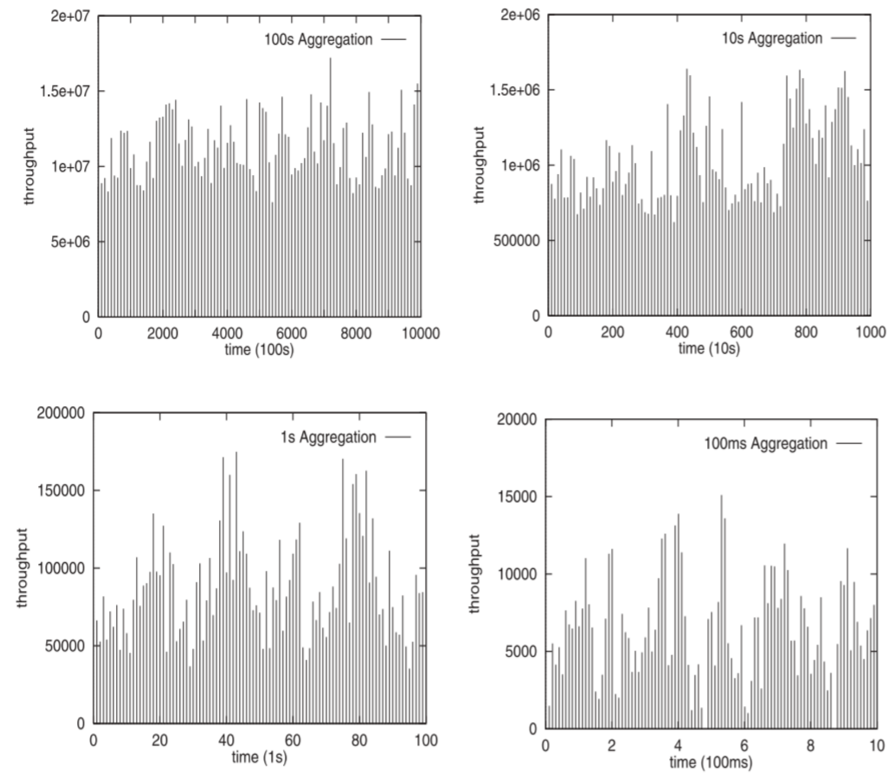


Answer: to model telecommunication traffic, among many other applications

Teletraffic field

- Statistically → traffic autocorrelation structure maintained for several time scales → **Long-Range Dependence.**





Stochastic self similarity is the «burstiness preservation sense»

Euclidean distributions:
a) The Euclidean Gaussian (1/2)

- Two working Matlab implementations that allow to generate pseudorandom numbers following Euclidean probability distributions, with “infinite” mean and variance (**i.e. two BANs**).

Gaussian

PDF: $f_X(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

```
N = 1e4;
Bd = 2;
setappdata(0, 'BAN_DEGREE', Bd);
true_mean = 3.4
true_sigma = 0.1;
true_var = true_sigma * true_sigma;
le=1;
x = BanArray(true_mean.*ones(N,1) + true_sigma.*randn(N,1),le);
```

Two BANs → $(true\_mean\ \ 0\ \ 0\ \ldots)\ \gamma^{le}$, $(true\_sigma\ \ 0\ \ 0\ \ldots)\ \gamma^{le}$

IN SUPREMÆ DIGNITATIS 1343

Euclidean distributions:
a) The Euclidean Gaussian (2/2)

```
est_mean = mean(x);
est_var = var(x);
```

Fitting Phase

Experimental results

With N=1e5 , le=1, we obtain:

```
True mean:
(3.4 0 0)G1
Estimated mean:
(3.3998 0 0)G1
```

$3.4\,\gamma$

$3.3998\,\gamma$

```
True variance:
(0.01 0 0)G2
Estimated variance:
(0.0100408 0 0)G2
```

$0.01\,\gamma^2$

$0.0100408\,\gamma^2$

```
x =

(3.46715 +0η +0η^{2})γ^{1}
(3.27925 +0η +0η^{2})γ^{1}
(3.47172 +0η +0η^{2})γ^{1}
(3.56302 +0η +0η^{2})γ^{1}
(3.44889 +0η +0η^{2})γ^{1}
(3.50347 +0η +0η^{2})γ^{1}
(3.47269 +0η +0η^{2})γ^{1}
(3.36966 +0η +0η^{2})γ^{1}
(3.42939 +0η +0η^{2})γ^{1}
(3.32127 +0η +0η^{2})γ^{1}
```

NA Log Normal

- A good statistical model to represent the amount of traffic per unit time.
- In Internet discussion fora → comment length distributions very regular and described by the log-normal form with a very high precision

Matlab Simulation:

```
x_as_bArr = BanArray(true_sigma_as_ban.*randn(N,1), le);
x_as_bArr = x_as_bArr + mu_as_ban;
x_as_bArr = true_theta_as_ban * exp(x_as_bArr);
```

But how to compute e^{BAN} ??

e^{γ} = ?

$\sim \gamma^{100}$ VAN Class

Non-Archimedean Analysis and Gamma Theory

- **Axiom (of Archimedes)**. Let U be any totally ordered field. Then, $\forall$ x, y $\in$ U, $0 < x < y$, $\exists$ n $\in$ N : $y < nx$.
- Gamma Theory, a non-standard model originally proposed by V. Benci in 1995
- **Axiom 1**. Exists an ordered field E $\supset$ R whose numbers are called Euclidean numbers
- **Axiom 2**. Exists a function num, num : U $\rightarrow$ E which satisfies
  - $\gamma = \text{num}(N)$
  - $\text{num}(A \cup B) = \text{num}(A) + \text{num}(B) - \text{num}(A \cap B)$
  - $\text{num}(A \times B) = \text{num}(A) \cdot \text{num}(B)$
- **Axiom 3**. Given a real function φ, $\exists!$ φ^* defined over E such that:
  - $\varphi(x) = \varphi^*(x) \quad \forall\, x \in R$
  - $Id^*(R) = Id(E)$, where Id(A) is the identity function on A

Gamma Theory and Algorithmic Numbers

- Any couple of real functions φ, ψ satisfies:
  - a) $(\varphi + \psi)^* = \varphi^* + \psi^*$
  - b) $(\varphi \cdot \psi)^* = \varphi^* \cdot \psi^*$
  - c) $(\varphi \circ \psi)^* = \varphi^* \circ \psi^*$

- $\mathbb{R}$ is “too rich” to be entirely managed by a finite machine… we must use **Algorithmic Field**

```
>> m1 = single(-2^127);
>> m2 = single(2^127);
>> m3 = single(1);
>> m1+(m2+m3)
ans = 0
>> (m1+m2)+m3
ans = 1
```

- The importance of fixed-length representations of numbers in symbolic computations

➢ **Algorithmic Numbers (ANs)**, introduced by V. Benci and M. Cococcioni … **Definition:**

$$\xi = \sum_{k=0}^{l} r_k \gamma^{s_k} \;\; ; \quad r_k \in \mathrm{R},\; s_k \in \mathrm{Q}; \quad s_k > s_{k+1} \, .$$

Algorithmic Numbers & BANs

- "Normal form": $\xi = \gamma^p P\left(\eta^{\frac{1}{m}}\right)$

  where p ∈ Q, m ∈ N and P(x) is a polynomial with real coefficients such that P(0) = $r_0 \neq 0$.

BANs: Bounded Algorithmic Numbers

- **Definition:** $\gamma^p P(\eta)$, where P(x) is a polynomial with real coefficients of degree n such that P(0) ≠0 and p ∈ Z.
- **Operations between two BANs:**

Sum: (assuming p ≥ q)

$$\xi + \zeta = \gamma^p P(\eta) + \gamma^p (Q(\eta)\eta^{p-q}) = \gamma^p (P(\eta) + \mathrm{tr}_n [Q(\eta)\eta^{p-q}])$$

Product:

$$\xi\zeta = \gamma^{p+q}\, \mathrm{tr}_n [P(\eta) \cdot Q(\eta)]$$

- We implemented the class of BANs (Ban.m) and two-dimensional arrays of BANs (BanArray.m) in Matlab

Backup BANs

Division: After having rewritten ζ as

$$\zeta = \gamma^q \left(q_0 + \sum_{k=1}^{n} q_k \, \eta^k \right) = q_0 \gamma^q (1 - \varepsilon)$$

where $\varepsilon = -\sum_{k=1}^{n} \frac{q_k}{q_0} \eta^k$, the division definition becomes:

$$\frac{\xi}{\zeta} = \gamma^{p-q} \mathrm{tr}_n \left[\frac{P(\eta)}{q_0} (1 + \varepsilon + \varepsilon^2 + \cdots + \varepsilon^n) \right]$$
$$= \gamma^{p-q} \left(\frac{P(\eta)}{q_0} + \mathrm{tr}_n \left[\varepsilon \frac{P(\eta)}{q_0} \right] + \cdots + \mathrm{tr}_n \left[\varepsilon^n \frac{P_n(\eta)}{q_0} \right] \right)$$

- We implemented the class of BANs (Ban.m) and two-dimensional arrays of BANs (BanArray.m) in Matlab

Backup VAN 1

The idea is to represent such transcendental infinite numbers as sum of monosemia, similar to Equation (1) except that the exponents are taken not in $\mathbb{Q}$ but in a vector space over $\mathbb{Q}$. To do that, one only requires the definition of a proper transcendental basis V. For the case $e^{2\gamma+3}$ it is enough to set V as

$$V = \{\beta_0, \beta_1\} = \left\{1, \frac{\gamma}{\ln \gamma}\right\} \tag{2}$$

and to define $\mathbb{V}$ as the set of all the Euclidean numbers having the powers of γ in the space spanned by V. Indeed, it holds true

$$\begin{aligned} e^{2\gamma+3} &= e^3 e^{2\gamma} = e^3 \gamma^{\log_\gamma e^{2\gamma}} = e^3 \gamma^{2\gamma \log_\gamma e} = \\ &= e^3 \gamma^{2\gamma \frac{\ln e}{\ln \gamma}} = e^3 \gamma^{2\frac{\gamma}{\ln \gamma}} = e^3 \gamma^{2\beta_1} \end{aligned}$$

and more generally, for instance,

$$e^{2\gamma+3} + 5\gamma = e^3 e^{2\gamma} + 5\gamma = e^3 \gamma^{2\beta_1} + 5\gamma^{\beta_0}$$

$$2^\gamma = e^{\ln 2^\gamma} = e^{\ln(2)\gamma} = \gamma^{\ln(2)\beta_1}$$

The case of $\ln \eta$ is quite similar but not straightforward, since it requires to pass through $\ln \gamma$. To numerically embed it, ones can use the basis

$$V = \{\beta_{-1}, \beta_0\} = \left\{\frac{\ln \ln \gamma}{\ln \gamma}, 1\right\} \tag{3}$$

obtaining the following identity

$$\ln \gamma = \gamma^{\log_\gamma \ln \gamma} = \gamma^{\frac{\ln \ln \gamma}{\ln \gamma}} = \gamma^{\beta_{-1}}$$

Backup VAN 2

Since $\ln \eta$ is the opposite of $\ln \gamma$, one evinces that

$$\ln \eta = \ln \frac{1}{\gamma} = \ln \gamma^{-1} = -\ln \gamma = -\gamma^{\beta_{-1}}$$

Finally, an example which merges both the approaches is η^η. The first step requires to apply the Taylor approximation, even if by means of a transcendental function rather than a polynomial one.

$$\eta^\eta = e^{\ln \eta^\eta} = e^{\eta \ln \eta} \cong 1 + \eta \ln \eta + \frac{1}{2}\eta^2 \ln^2 \eta$$

Then, one can leverage the Algorithmic field $\mathbb{V}$ induced by Equation (3) to rewrite $\ln \eta$ in a polynomial-like form

$$\eta^\eta \cong 1 + \eta \ln \eta + \frac{1}{2}\eta^2 \ln^2 \eta = 1 - \gamma^{\beta_0+\beta_{-1}} + \frac{1}{2}\gamma^{2(\beta_0+\beta_{-1})}$$

Alternative approaches may have been rewriting η^η as $\gamma^{-\frac{1}{\gamma}}$ and use the transcendental basis

$$V = \{\beta_0, \beta_1\} = \left\{1, \frac{1}{\gamma}\right\}$$

obtaining the exact representation $\eta^\eta = \gamma^{-\beta_1}$; or resorting on V in Equation (2) and, after some computations, rewriting η^η as $e^{-\frac{1}{\beta_1}}$. Of course, the latter idea does not bring to a polynomial-like representation.

Backup VAN 3

Definition 15. We say that the set of euclidean numbers

$$\{\beta_{-m}, .., \beta_{-1}, \beta_0, \beta_1, ..., \beta_n\}$$

is a transcendent basis if $\beta_0 = 1$, and

$$h < k \Rightarrow \forall r \in \mathbb{R},\ \alpha^{r\beta_h} \ll \alpha^{\beta_k}$$

The real vector space V generated by such a basis will be called the order group. An Euclidean number of the form

$$r\alpha^v,\ r \in \mathbb{R},\ v \in V$$

will be called V-monosemium. An Euclidean number of the form

$$\xi = \sum_{k=0}^{\ell} r_k \alpha^{v_k};\quad v_k > v_{k+1}$$

is called V-algorithmic number [2].

The V-algorithmic numbers generalize the notion of algorithmic numbers. It is necessary to say that they are not very suitable for numeric computations since they do not have a normal form defined by a polynomial such as (5). We introduced them just for completeness since they might be used in some particular problem. In any case they form a ring and they have an approximate inverse. It is possible to prove that there exists an infinite-dimensional real vector pace $V_{\mathbb{E}}$ such that any Euclidean number can be approximated by a transcendental AN with an order group $V \subset V_{\mathbb{E}}$.

Example. Take

$$V = \{\beta_0, \beta_1\} = \left\{1, \frac{\alpha}{\log \alpha}\right\}$$

The number $e^{2\alpha+3} + 5\alpha^2$ can be represented as follows:

$$e^{2\alpha+3} + 5\alpha^2 = e^3 e^{2\alpha} + 5\alpha^2 = e^3 \alpha^{\frac{2\alpha}{\log \alpha}} + 5\alpha^2 = e^3 \alpha^{2\beta_1} + 5\alpha^{2\beta_0}.$$

Backup2

- TheBanclassB.1.1 How to set the degree of the BanUse the command setappdata(0, ’BANDEGREE’, 3) to set the degree to 3, as an example.
- B.1.2 How to set the format for displaying a Ban: Use the command setappdata(0, ’BANFORMAT’, 0); to display the Ban in ASCII format, like: (1 2 3 4)G-3
- Use the command setappdata(0, ’BANFORMAT’, 2); to display the Ban in LATEX format, like: (1 + 2η+ 3η2+ 4η3)γ−3
- Use the command setappdata(0, ’BANFORMAT’, 1); to display the Ban in an INTERMEDIATE format, like: (1 +2η+3η{2}+4η{3})γ{-3}
- The constant 1 (i.e., the Euclidean number one) as aBan3: γ0(1 + 0η+ 0η2+ 0η3) = γ0[1 0 0 0] =(0)[1 0 0 0] = [1 0 0 0] =“1Ban3”
- The constant 0 (i.e., the Euclidean zero) as aBan3: γ0(1 + 0η+ 0η2+ 0η3) =γ0[0 0 0 0] =(0)[0 0 0 0] = [0 0 0 0] =“0Ban3”
- The real value 7.6 as aBan3: γ0(7.6+0η+0η2+0η3) =γ0[7.6 0 0 0] =(0)[7.6 0 0 0] = [7.6 0 0 0] =“7.6Ban3” (the value 7.6 will be stored as a double precision floating point number)
- The constant π as aBan3: γ0(3.14 + 0η+ 0η2+ 0η3) =γ0[3.14 0 0 0] =(0)[3.14 0 0 0] = [3.14 0 0 0] =“3.14Ban3”or even “PIBan3”(of course we had to approximateπusing finite a finite decimal ap-proximation)

Testo da inserire