Statistical Inference for Heavy and Super-Heavy-tailed distributions

M. Isabel Fraga Alves

DEIO, Faculty of Sciences, University of Lisbon, Portugal

Laurens de Haan

Econometric Institute, Erasmus University of Rotterdam, The Netherlands

Cláudia Neves

UIMA, Department of Mathematics, University of Aveiro, Portugal

Summary

- EVT: Modeling the tail of a distribution;

Summary

- EVT: Modeling the tail of a distribution;
- Characterization of Super-Heavy tails;

Summary

- EVT: Modeling the tail of a distribution;
- Characterization of Super-Heavy tails;
- Estimation of the index of tail heaviness $\alpha \geq 0$;

Summary

- EVT: Modeling the tail of a distribution;
- Characterization of Super-Heavy tails;
- Estimation of the index of tail heaviness $\alpha \geq 0$;
- Testing the presence of Super-Heavy tails;

Summary

- EVT: Modeling the tail of a distribution;
- Characterization of Super-Heavy tails;
- Estimation of the index of tail heaviness $\alpha \geq 0$;
- Testing the presence of Super-Heavy tails;
- Simulation results.

Heavy-tailed distributions

Let $\{X_i\}_{i=1}^n$ be i.i.d. random variables with common distribution function F.

Heavy-tailed distributions

Let $\{X_i\}_{i=1}^{n}$ be i.i.d. random variables with common distribution function F.

Heavy-tailed models suggested by EVT:

There exist constants $a_n > 0$, $n \in \mathbb{N}$ such that

$$\lim_{n\to\infty} P\Big\{\frac{\max(X_1,\ldots,X_n)}{a_n} \leq x\Big\} = \exp\Big(-\big(1+\frac{x}{\alpha}\big)^{-\alpha}\Big),$$

for all x for which $1+\alpha^{-1}x > 0$, $\alpha > 0$.

Heavy-tailed distributions

Let $\{X_i\}_{i=1}^n$ be i.i.d. random variables with common distribution function F.

Heavy-tailed models suggested by EVT:

There exist constants $a_n > 0$, $n \in \mathbb{N}$ such that

$$\lim_{n\to\infty} P\Big\{\frac{\max(X_1,\ldots,X_n)}{a_n} \le x\Big\} = \exp\Big(-\big(1+\frac{x}{\alpha}\big)^{-\alpha}\Big),$$

for all x for which $1+\alpha^{-1}x > 0$, $\alpha > 0$.

Models for the tail pertaining to $\alpha = 0$?

Heavy-tailed distributions

Let $\{X_i\}_{i=1}^n$ be i.i.d. random variables with common distribution function F.

Heavy-tailed models suggested by EVT:

There exist constants $a_n > 0$, $n \in \mathbb{N}$ such that

$$\lim_{n\to\infty} P\Big\{\frac{\max(X_1,\ldots,X_n)}{a_n} \le x\Big\} = \exp\Big(-(1+\frac{x}{\alpha})^{-\alpha}\Big),$$

for all x for which $1+\alpha^{-1}x > 0$, $\alpha > 0$.

Examples: log-Pareto, log-Cauchy, log-Weibull ...

Characterizing Heavy-tail behavior

F is of regular variation at infinity of index $\alpha > 0$ if

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha},$$

for all $x > 0$.

$$\boxed{1 - F \in RV_\alpha}$$

Characterizing Heavy-tail behavior

F is of regular variation at infinity of index $\alpha > 0$ if

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha},$$

for all $x > 0$.

$$\boxed{1 - F \in RV_\alpha}$$

F is a slowly varying function if

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = 1,$$

for all $x > 0$.

$$\boxed{1 - F \in RV_0}$$

Characterizing Heavy-tail behavior

F is of regular variation at infinity of index $\alpha > 0$ if

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha},$$

for all $x > 0$.

$$\boxed{1 - F \in RV_\alpha}$$

F is a slowly varying function if

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = 1,$$

for all $x > 0$.

$$\boxed{1 - F \in RV_0}$$

Heavy tails:

Suppose there exits a positive function a such that F satisfies

$$\lim_{t\to\infty} \frac{F(tx)-F(t)}{a(t)} = \frac{1-x^{-\alpha}}{\alpha},$$

for all $x > 0$, with $\alpha > 0$.

Characterizing Heavy-tail behavior

F is of regular variation at infinity of index $\alpha > 0$ if

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha},$$

for all $x > 0$.

$$1 - F \in RV_\alpha$$

F is a slowly varying function if

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = 1,$$

for all $x > 0$.

$$1 - F \in RV_0$$

Heavy tails:

Suppose there exits a positive function a such that F satisfies

$$\lim_{t\to\infty} \frac{F(tx)-F(t)}{a(t)} = \frac{1-x^{-\alpha}}{\alpha},$$

for all $x > 0$, with $\alpha > 0$.

Super-Heavy tails:

$$\lim_{t\to\infty} \frac{F(tx)-F(t)}{a(t)} = \log x, \quad x > 0.$$

$$F \in ERV_\alpha,\ \alpha \geq 0$$

de Haan (1984)

Π-variation and Γ-variation

Super-Heavy-tailed distributions

$$U(t) := \left(\frac{1}{1-F}\right)^{\leftarrow}(t) = \inf\left\{x : F(x) \geq 1 - \frac{1}{t}\right\}$$

The following are equivalent:

$$F \in \Pi(a) : \qquad \lim_{t\to\infty} \frac{F(tx) - F(t)}{a(t)} = \log x, \quad x > 0;$$

$$U \in \Gamma(q) : \qquad \lim_{t\to\infty} \frac{U(t + x\,q(t))}{U(t)} = e^x, \quad x \in \mathbb{R},$$

with $q(t) = t^2 a(U(t))$, $a \in RV_0$

A step towards estimation

$$\lim_{t\to\infty}\int_1^{\infty}\frac{F(tx)-F(t)}{a(t)}\,\frac{dx}{x^2}=\int_1^{\infty}\frac{1-x^{-\alpha}}{\alpha}\,\frac{dx}{x^2}=\frac{1}{1+\alpha}$$

A step towards estimation

$$\lim_{t\to\infty}\int_1^\infty \frac{F(tx)-F(t)}{a(t)}\,\frac{dx}{x^3} = \int_1^\infty \frac{1-x^{-\alpha}}{\alpha}\,\frac{dx}{x^3} = \frac{1}{2(2+\alpha)}$$

A step towards estimation

For $j = 1, 2$

$$\lim_{t\to\infty}\int_1^\infty \frac{F(tx)-F(t)}{a(t)}\,\frac{dx}{x^{j+1}} = \int_1^\infty \frac{1-x^{-\alpha}}{\alpha}\,\frac{dx}{x^{j+1}} = \frac{1}{j\,(j+\alpha)}$$

A step towards estimation

For $j = 1, 2$

$$\lim_{t\to\infty} \int_1^\infty \frac{F(tx) - F(t)}{a(t)} \frac{dx}{x^{j+1}} = \int_1^\infty \frac{1 - x^{-\alpha}}{\alpha} \frac{dx}{x^{j+1}} = \frac{1}{j(j+\alpha)}$$

$$\int_1^\infty \big(F(tx) - F(t)\big) \frac{dx}{x^{j+1}} = \frac{t^j}{j} \int_t^\infty \frac{dF(u)}{u^j}$$

A step towards estimation

For $j = 1, 2$

$$\lim_{t\to\infty} \int_1^\infty \frac{F(tx) - F(t)}{a(t)} \frac{dx}{x^{j+1}} = \int_1^\infty \frac{1 - x^{-\alpha}}{\alpha} \frac{dx}{x^{j+1}} = \frac{1}{j\,(j+\alpha)}$$

$$\int_1^\infty \left(F(tx) - F(t)\right) \frac{dx}{x^{j+1}} = \frac{t^j}{j} \int_t^\infty \frac{dF(u)}{u^j}$$

$$\frac{2\int_1^\infty \left(F(tx) - F(t)\right) dx/x^3}{\int_1^\infty \left(F(tx) - F(t)\right) dx/x^2} = \frac{\int_t^\infty (t/u)^2\, dF(u)}{\int_t^\infty (t/u)\, dF(u)} \underset{t\to\infty}{\longrightarrow} \frac{1+\alpha}{2+\alpha} =: \psi(\alpha)$$

A step towards estimation

For $j = 1, 2$

$$\lim_{t\to\infty}\int_1^\infty \frac{F(tx)-F(t)}{a(t)}\,\frac{dx}{x^{j+1}} = \int_1^\infty \frac{1-x^{-\alpha}}{\alpha}\,\frac{dx}{x^{j+1}} = \frac{1}{j\,(j+\alpha)}$$

$$\int_1^\infty \left(F(tx)-F(t)\right)\frac{dx}{x^{j+1}} = \frac{t^j}{j}\int_t^\infty \frac{dF(u)}{u^j}$$

$$\frac{2\int_1^\infty \left(F(tx)-F(t)\right)dx/x^3}{\int_1^\infty \left(F(tx)-F(t)\right)dx/x^2} = \frac{\int_t^\infty (t/u)^2\,dF(u)}{\int_t^\infty (t/u)\,dF(u)} \underset{t\to\infty}{\longrightarrow} \frac{1+\alpha}{2+\alpha} =: \psi(\alpha)$$

Let $X_{1,n} \le X_{2,n} \le \ldots \le X_{n,n}$ be the order statistics corresponding to the random sample $(X_1, \ldots, X_n)$,

$$\hat{\psi}_n(k) := \frac{\sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)^2}{\sum_{i=0}^{k-1} X_{n-k,n}/X_{n-i,n}} \underset{n\to\infty}{\overset{P}{\longrightarrow}} \psi(\alpha), \quad 0 \le \alpha < \infty \quad ?$$

Estimation of the tail parameter

$$\hat{\alpha}_n(k) := \frac{2\sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)^2 - \sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)}{\sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right) - \sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)^2}$$

$k = k_n$ is a sequence of positive integers such that $k_n \to \infty$ and $k_n = o(n)$ as $n \to \infty$ ➽ k_n is an intermediate sequence

Estimation of the tail parameter

$$\hat{\alpha}_n(k) := \frac{2\sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)^2 - \sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)}{\sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right) - \sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)^2}$$

$k = k_n$ is a sequence of positive integers such that $k_n \to \infty$ and $k_n = o(n)$ as $n \to \infty$ ➻ k_n is an intermediate sequence

Consistency: If $(n/\sqrt{k})\, a\big(U(n/k)\big) \underset{n\to\infty}{\longrightarrow} \infty$, then $\hat{\alpha}_n(k) \overset{P}{\underset{n\to\infty}{\longrightarrow}} \alpha \geq 0$.

Estimation of the tail parameter

$$\hat{\alpha}_n(k) := \frac{2\sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)^2 - \sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)}{\sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right) - \sum_{i=0}^{k-1}\left(X_{n-k,n}/X_{n-i,n}\right)^2}$$

$k = k_n$ is a sequence of positive integers such that $k_n \to \infty$ and $k_n = o(n)$ as $n \to \infty$ ➽ k_n is an intermediate sequence

Consistency: If $(n/\sqrt{k})\, a(U(n/k)) \underset{n\to\infty}{\longrightarrow} \infty$, then $\hat{\alpha}_n(k) \overset{P}{\underset{n\to\infty}{\longrightarrow}} \alpha \geq 0$.

Since $(n/k)\, a(U(n/k)) \underset{n\to\infty}{\longrightarrow} \alpha$, we have that:

- for $\alpha > 0$, consistency holds for any intermediate sequence k_n;
- for $\alpha = 0$, we need to impose a lower bound to the sequence k_n.

Auxiliary results I

Define, for $j = 1, 2$,

$$
\begin{aligned}
M_n^{(j)}(k) &:= \frac{n}{k}\, j \int_1^\infty \left(F_n(xX_{n-k,n}) - F_n(X_{n-k,n})\right) \frac{dx}{x^{j+1}} \\
&= \frac{1}{k}\sum_{i=0}^{k-1}\left(\frac{X_{n-k,n}}{X_{n-i,n}}\right)^j
\end{aligned}
$$

Auxiliary results I

Define, for $j = 1, 2$,

$$M_n^{(j)}(k) := \frac{n}{k}\, j \int_1^\infty \left(F_n(xX_{n-k,n}) - F_n(X_{n-k,n})\right) \frac{dx}{x^{j+1}} = \frac{1}{k}\sum_{i=0}^{k-1}\left(\frac{X_{n-k,n}}{X_{n-i,n}}\right)^j$$

Let $\{U_n\}_{n\geq 1}$ be a sequence of i.i.d. uniformly distributed r.v.'s. Define:

$$E_n(t) := \frac{1}{n}\sum_{i=1}^{n} I_{\{U_{i,n}\leq t\}},\ t \in [0,1]$$

and

$$\{e_n(t); 0 \leq t \leq 1\} := \{\sqrt{n}\,(E_n(t) - t);\ t \in [0,1]\}$$

n-th uniform empirical process

Auxiliary results I

Define, for $j = 1, 2$,

$$M_n^{(j)}(k) := \frac{n}{k}\, j \int_1^\infty \big(F_n(xX_{n-k,n}) - F_n(X_{n-k,n})\big)\, \frac{dx}{x^{j+1}} = \frac{1}{k}\sum_{i=0}^{k-1}\Big(\frac{X_{n-k,n}}{X_{n-i,n}}\Big)^j$$

Define $Q_n(k) := (k/n)/\big(1 - F(X_{n-k,n})\big)$ and $t^* = t_n^*(x) := \frac{1-F(xX_{n-k,n})}{1-F(X_{n-k,n})}$, $x \geq 1$.

For all $\nu \in (0, 1/2)$ as, $n \to \infty$,

$$F_n(xX_{n-k,n}) - F_n(X_{n-k,n}) = -\frac{\sqrt{k}}{n} B(t^*) + \big(F(xX_{n-k,n}) - F(X_{n-k,n})\big) Q_n(k) + O_p(1)\frac{k^\nu}{n}\big(t^*(1-t^*)\big)^\nu,$$

where $Q_n(k)$ is a r.v. independent of the Brownian bridge B and the O_p-term is uniform for $x \geq 1$.

[Proposition 4.3.1 of Csörgő and Horváth (1993)]

Auxiliary results I

Define, for $j = 1, 2$,

$$M_n^{(j)}(k) := \frac{n}{k}\, j \int_1^\infty \left(F_n(xX_{n-k,n}) - F_n(X_{n-k,n})\right) \frac{dx}{x^{j+1}} = \frac{1}{k}\sum_{i=0}^{k-1}\left(\frac{X_{n-k,n}}{X_{n-i,n}}\right)^j$$

$$\begin{aligned} &\frac{k}{n}\,\frac{M_n^{(j)}(k)}{a(X_{n-k,n})} - \frac{1}{j+\alpha} \\ = \; & -j\left(\frac{k}{n}\,\frac{1}{a(X_{n-k,n})}\right)\frac{1}{\sqrt{k}}\int_1^\infty B(t^*)\,\frac{dx}{x^{j+1}} + \frac{1}{j+\alpha}\left(Q_n(k)-1\right) \\ & + j\,Q_n(k)\int_1^\infty\left(\frac{F(xX_{n-k,n})-F(X_{n-k,n})}{a(X_{n-k,n})} - \frac{1-x^{-\alpha}}{\alpha}\right)\frac{dx}{x^{j+1}} \\ & + O_p(1)\,\frac{k^\nu}{n}\int_1^\infty \frac{\left(t^*(1-t^*)\right)^\nu}{a(X_{n-k,n})}\,\frac{dx}{x^{j+1}}. \end{aligned}$$

Auxiliary results II

- 2nd order Extended Regular Variation:

Suppose there exists a function $A(t) \to 0$, as $t \to \infty$, of constant sign near infinity and a second order parameter $\rho \leq 0$ such that

$$\lim_{t\to\infty} \frac{\frac{F(tx)-F(t)}{a(t)} - \frac{1-x^{-\alpha}}{\alpha}}{A(t)} = \frac{1}{\rho}\left(\frac{x^{-\alpha+\rho}-1}{-\alpha+\rho} + \frac{x^{-\alpha}-1}{\alpha}\right) =: H_{\alpha,\rho}(x),$$

According to de Haan and Stadtmüller (1996), $|A(t)| \in RV_\rho$.

$$F \in 2ERV(\alpha,\rho),\ \alpha \geq 0,\ \rho \leq 0$$

Auxiliary results II

- 2nd order Extended Regular Variation:

Suppose there exists a function $A(t) \to 0$, as $t \to \infty$, of constant sign near infinity and a second order parameter $\rho \leq 0$ such that

$$\lim_{t\to\infty} \frac{\frac{F(tx)-F(t)}{a(t)} - \frac{1-x^{-\alpha}}{\alpha}}{A(t)} = \frac{1}{\rho}\left(\frac{x^{-\alpha+\rho}-1}{-\alpha+\rho} + \frac{x^{-\alpha}-1}{\alpha}\right) =: H_{\alpha,\rho}(x),$$

According to de Haan and Stadtmüller (1996), $|A(t)| \in RV_\rho$.

$$F \in 2ERV(\alpha,\rho),\ \alpha \geq 0,\ \rho \leq 0$$

- Continuity-modulus property for Brownian bridge:

For any $\delta \in (0, 1/2)$, we have

$$\lim_{h\downarrow 0} \sup_{x>0} \frac{|B(x+h)-B(x)|}{h^{1/2-\delta}} = 0 \quad a.s.$$

Asymptotic normality

Given an intermediate sequence $k = k_n$ such that $(n/\sqrt{k})\, a\big(U(n/k)\big) \underset{n\to\infty}{\longrightarrow} \infty$,

if $F \in 2ERV(\alpha, \rho)$ with $\alpha \geq 0$ and $\Big(n\, a\big(U(n/k)\big)\Big)^{1/2} A\big(U(n/k)\big) \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}$,

Asymptotic normality

Given an intermediate sequence $k = k_n$ such that $(n/\sqrt{k})\, a\big(U(n/k)\big) \underset{n\to\infty}{\longrightarrow} \infty$, if $F \in 2ERV(\alpha,\rho)$ with $\alpha \geq 0$ and $\Big(n\, a\big(U(n/k)\big)\Big)^{1/2} A\big(U(n/k)\big) \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}$,

$$c_j = c_j(\alpha,\rho) := \int_1^\infty j\, H_{\alpha,\rho}(x)\, \frac{dx}{x^{j+1}} = \frac{1}{(j+\alpha)(j+\alpha-\rho)}, \quad j = 1, 2$$

- $\alpha > 0$,

$$\Big(n\, a(X_{n-k,n})\Big)^{1/2} \left(\frac{k}{n}\, \frac{M_n^{(j)}(k)}{a(X_{n-k,n})} - \frac{1}{j+\alpha}\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} -\frac{j}{\sqrt{\alpha}} \int_1^\infty B(x^{-\alpha})\, \frac{dx}{x^{j+1}} + \frac{\sqrt{\alpha}}{j+\alpha}\, N + \lambda\, c_j,$$

where $\{B(t);\, 0 \leq t \leq 1\}$ is a Brownian bridge and N denotes a normal r.v. independent of B;

Asymptotic normality

Given an intermediate sequence $k = k_n$ such that $(n/\sqrt{k})\, a(U(n/k)) \underset{n\to\infty}{\longrightarrow} \infty$,

if $F \in 2ERV(\alpha, \rho)$ with $\alpha \geq 0$ and $\Big(n\, a(U(n/k))\Big)^{1/2} A(U(n/k)) \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}$,

$$c_j = c_j(\alpha, \rho) := \int_1^\infty j\, H_{\alpha,\rho}(x)\, \frac{dx}{x^{j+1}} = \frac{1}{(j+\alpha)(j+\alpha-\rho)}, \quad j = 1, 2$$

- $\alpha > 0$,

$$\Big(n\, a(X_{n-k,n})\Big)^{1/2} \Big(\frac{k}{n}\, \frac{M_n^{(j)}(k)}{a(X_{n-k,n})} - \frac{1}{j+\alpha}\Big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} -\frac{j}{\sqrt{\alpha}} \int_1^\infty B(x^{-\alpha})\, \frac{dx}{x^{j+1}} + \frac{\sqrt{\alpha}}{j+\alpha}\, N + \lambda\, c_j,$$

where $\{B(t); 0 \leq t \leq 1\}$ is a Brownian bridge and N denotes a normal r.v. independent of B;

- $\alpha = 0$,

$$\Big(n\, a(X_{n-k,n})\Big)^{1/2} \Big(\frac{k}{n}\, \frac{M_n^{(j)}(k)}{a(X_{n-k,n})} - \frac{1}{j+\alpha}\Big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} -j \int_1^\infty W(\log x)\, \frac{dx}{x^{j+1}} + \lambda\, c_j,$$

where $\{W(t); 0 \leq t < \infty\}$ is a standard Brownian motion.

Main result

Given an intermediate sequence $k = k_n$ such that $(n/\sqrt{k})\, a\big(U(n/k)\big) \underset{n\to\infty}{\longrightarrow} \infty$,

if $F \in 2ERV(\alpha,\rho)$ with $\alpha \geq 0$ and $\Big(n\, a\big(U(n/k)\big)\Big)^{1/2} A\big(U(n/k)\big) \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}$,

Main result

Given an intermediate sequence $k = k_n$ such that $(n/\sqrt{k})\, a(U(n/k)) \underset{n\to\infty}{\longrightarrow} \infty$,

if $F \in 2ERV(\alpha, \rho)$ with $\alpha \geq 0$ and $\Big(n\, a(U(n/k))\Big)^{1/2} A(U(n/k)) \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}$,

$$\Big(\sum_{i=0}^{k-1} \frac{X_{n-k,n}}{X_{n-i,n}}\Big)^{1/2} \Big(\hat{\alpha}_n(k) - \alpha\Big) \overset{d}{\underset{n\to\infty}{\longrightarrow}} N(b_1, \sigma_1^2),$$

with

$$b_1 = b_1(\alpha, \rho) := \frac{-\lambda\sqrt{1+\alpha}\,(2+\alpha)}{(1+\alpha-\rho)(2+\alpha-\rho)},$$

$$\sigma_1^2 = \sigma_1^2(\alpha) := \frac{(1+\alpha)(2+\alpha)}{(3+\alpha)(4+\alpha)}\,(4+3\alpha+\alpha^2).$$

Main result

Given an intermediate sequence $k = k_n$ such that $(n/\sqrt{k})\, a(U(n/k)) \underset{n\to\infty}{\longrightarrow} \infty$,

if $F \in 2ERV(\alpha, \rho)$ with $\alpha \geq 0$ and $\Big(n\, a(U(n/k))\Big)^{1/2} A(U(n/k)) \underset{n\to\infty}{\longrightarrow} \lambda \in \mathbb{R}$,

$$\Big(\sum_{i=0}^{k-1} \frac{X_{n-k,n}}{X_{n-i,n}}\Big)^{1/2} \Big(\hat{\alpha}_n(k) - \alpha\Big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N(b_1, \sigma_1^2),$$

with

$$b_1 = b_1(\alpha, \rho) := \frac{-\lambda\sqrt{1+\alpha}\,(2+\alpha)}{(1+\alpha-\rho)(2+\alpha-\rho)},$$

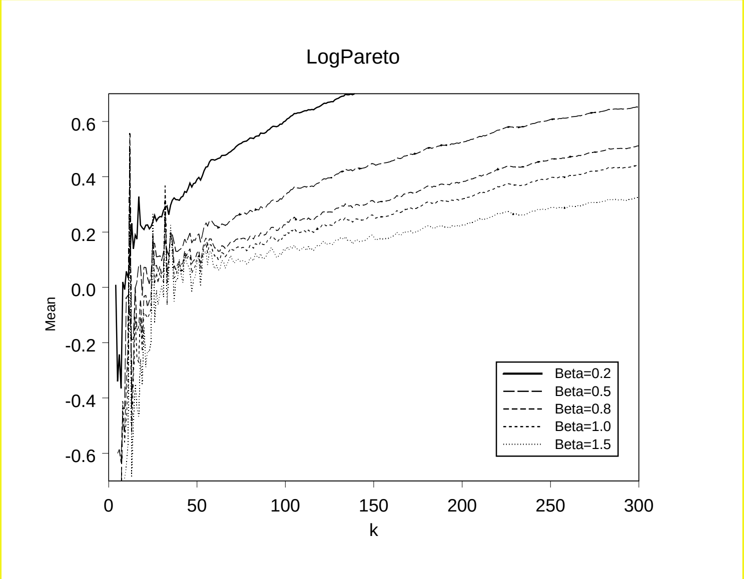
$$\sigma_1^2 = \sigma_1^2(\alpha) := \frac{(1+\alpha)(2+\alpha)}{(3+\alpha)(4+\alpha)}\,(4+3\alpha+\alpha^2).$$

An alternative formulation is

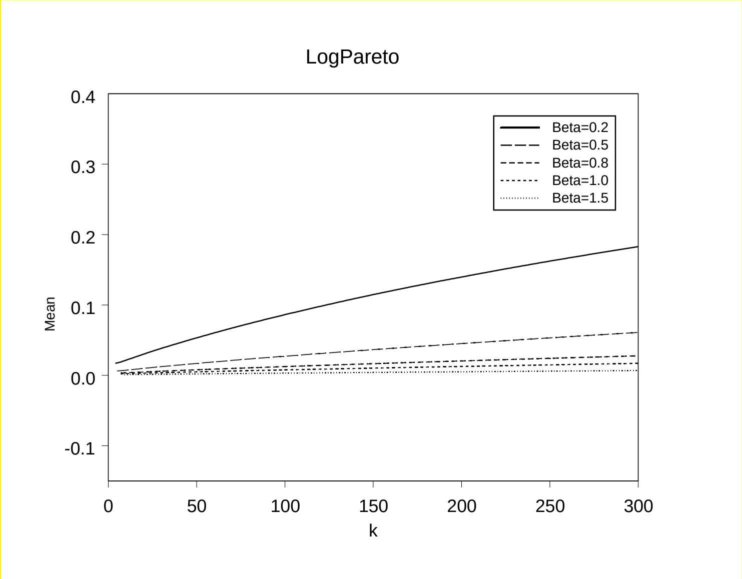
$$\Big(n\, a(U(n/k))\Big)^{1/2} (\hat{\alpha}_n(k) - \alpha) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N(b_1^*, \sigma_1^{*2}),$$

$$b_1^* = b_1^*(\alpha,\rho) = (1+\alpha)^{1/2} b_1(\alpha,\rho) \quad \text{and} \quad \sigma_1^{*2} = \sigma_1^{*2}(\alpha) = (1+\alpha)\,\sigma_1^2(\alpha)$$

Finite sample behavior I



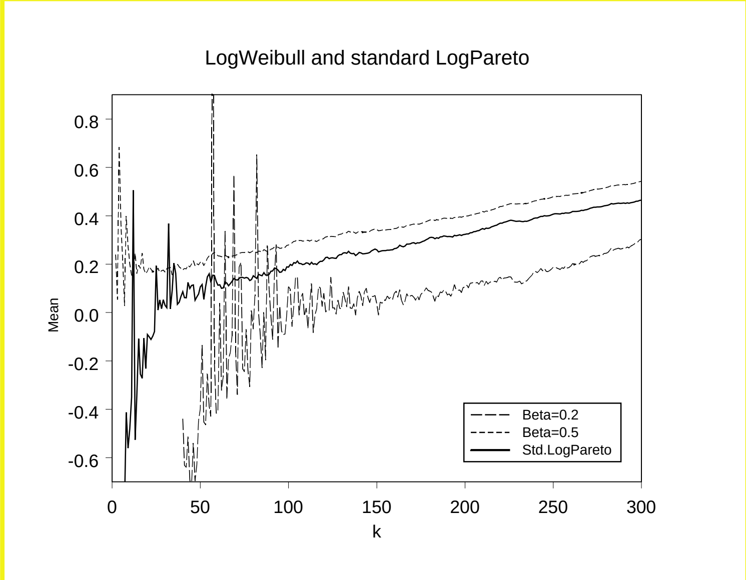
Patterns of the estimated mean of $\hat{\alpha}_n(k)$



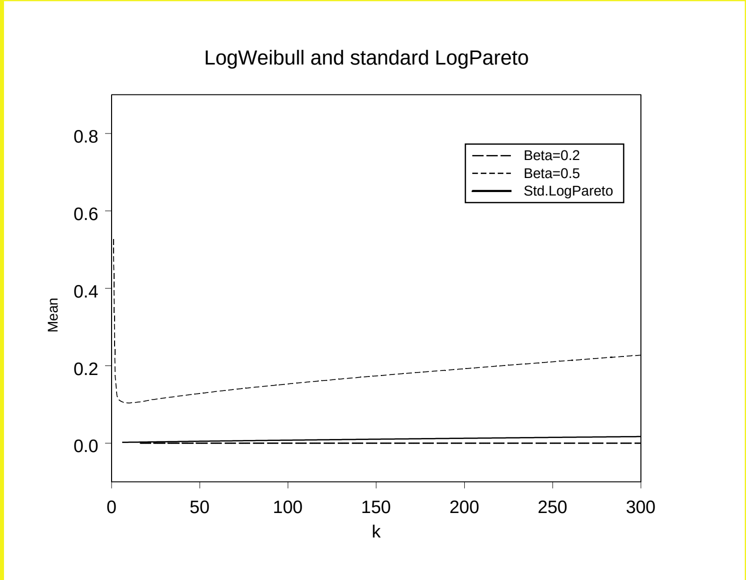
Patterns of the estimated mean of the Hill estimator

Let Y be a random variable with standard Pareto distribution.
A random variable W follows a log-Pareto distribution with parameter $\beta > 0$ if and only if $W = (e^{\beta Y} - 1)/\beta$.

Finite sample behavior II



Patterns of the estimated mean of $\hat{\alpha}_n(k)$



Patterns of the estimated mean of the Hill estimator

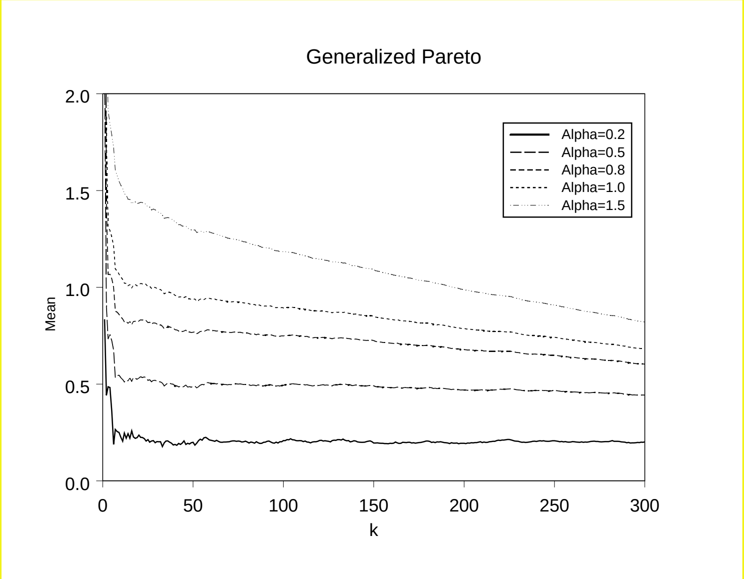
log-Weibull distribution:

$$F(x) = 1 - \exp\{-(\log x)^{\beta}\}, \; x \geq 1, \quad 0 < \beta < 1$$

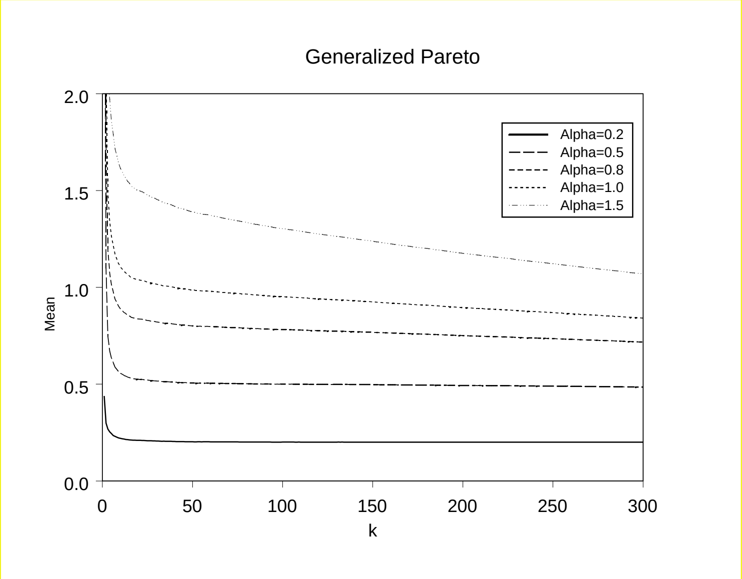
Standard log-Pareto distribution:

$$F_X(x) = 1 - (\log x)^{-1}, \; x \geq e$$

Finite sample behavior III



Patterns of the estimated mean of $\hat{\alpha}_n(k)$



Patterns of the estimated mean of the Hill estimator

Generalized Pareto distribution:

$$F(x) = 1 - \left(1 + \frac{x}{\alpha}\right)^{-\alpha}, x > 0, \quad \alpha \geq 0$$

Drawback...

... of the Hill estimator...

Consider a random sample taken from a standard log-Pareto population $X = e^Y$, with Y denoting a standard Pareto random variable. Then, the Hill estimator acquires the simple form

$$H_n(k) = \frac{1}{k}\sum_{i=0}^{k-1} Y_{n-i,n} - Y_{n-k,n}$$

Hence,

$$H_n(k) \stackrel{d}{=} Y_{n-k,n}\left(\frac{1}{k}\sum_{i=1}^{k} Y_i^* - 1\right) = \frac{n}{k}\left(S_k + b_k\right)\left(1 + o_p(1)\right),$$

with $\{Y_i^*\}_{i=1}^k$ denoting i.i.d. standard Pareto random variables independent of the random threshold $Y_{n-k,n}$, $b_k = O(\log k)$ and where S_k denotes a random variable with limiting sum-stable distribution.

Drawback...

... of the Hill estimator...

Undertaking the log-Weibull distribution, the Hill estimator may be written as

$$H_n(k) = \frac{1}{k}\sum_{i=0}^{k-1}\left(\log Y_{n-i,n}\right)^{1/\beta} - \left(\log Y_{n-k,n}\right)^{1/\beta}$$

Hence,

$$H_n(k) \overset{d}{=} \left(\log Y_{n-k,n}\right)^{1/\beta-1}\frac{1}{\beta}\left(\frac{1}{k}\sum_{i=1}^{k}\log Y_i^* + o_p(1)\right)$$

$$= \left(\log \frac{n}{k}\right)^{1/\beta-1}\frac{1}{\beta}\left(S_k^* + \sqrt{k}\right)\left(1+o_p(1)\right),$$

where S_k^* converges to a normal random variable.

Test of hypothesis

$$H_0 : \alpha = 0 \qquad \textit{versus} \qquad H_1 : \alpha > 0$$

Test of hypothesis

$$H_0 : \alpha = 0 \qquad \textit{versus} \qquad H_1 : \alpha > 0$$

<u>Test Statistic:</u>

$$T_n(k) := \sqrt{24}\left(\sum_{i=0}^{k-1} \frac{X_{n-k,n}}{X_{n-i,n}}\right)^{1/2}\left(\hat{\psi}_n(k) - \frac{1}{2}\right)$$

Reject H_0 in favor of the unilateral alternative if $T_n(k) > z_{1-\bar{\alpha}}$, where z_ε denotes the ε-quantile of the standard normal distribution.

Test of hypothesis

$$H_0 : \alpha = 0 \qquad \textit{versus} \qquad H_1 : \alpha > 0$$

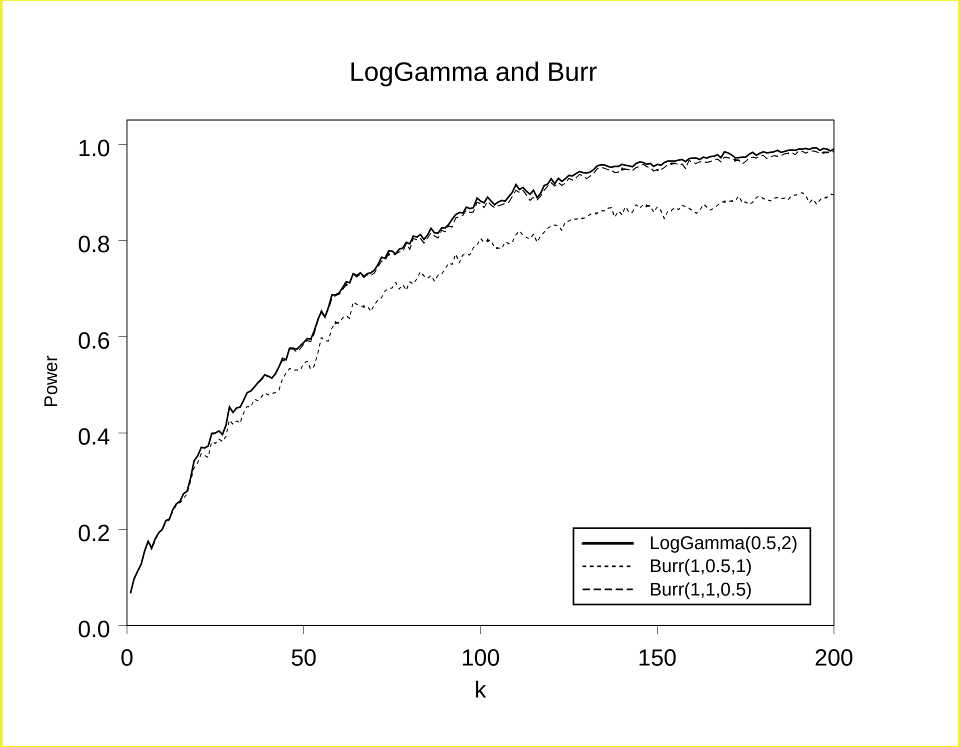
Test Statistic:

$$T_n(k) := \sqrt{24}\left(\sum_{i=0}^{k-1} \frac{X_{n-k,n}}{X_{n-i,n}}\right)^{1/2}\left(\hat{\psi}_n(k) - \frac{1}{2}\right)$$

Reject H_0 in favor of the unilateral alternative if $T_n(k) > z_{1-\bar{\alpha}}$, where z_ε denotes the ε-quantile of the standard normal distribution.

- The presented test is asymptotically of size $\bar{\alpha}$;
- the test discriminates between distributions lying in the class of ERV_α for which the second order condition holds with any intermediate sequence k_n such that $\left(n\, a\big(U(n/k_n)\big)\right)^{1/2} A\big(U(n/k_n)\big) \underset{n\to\infty}{\longrightarrow} 0$.

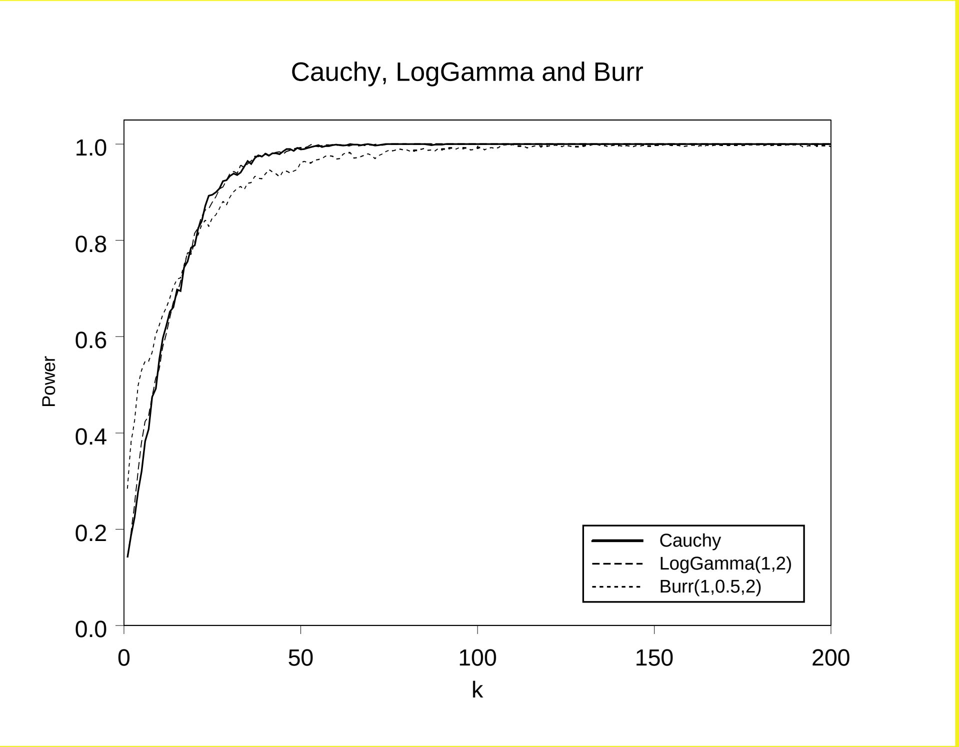
Empirical power



Empirical power at a nominal level $\bar{\alpha} = 0.05$

Underlying distribution function $F \in ERV_{0.5}$

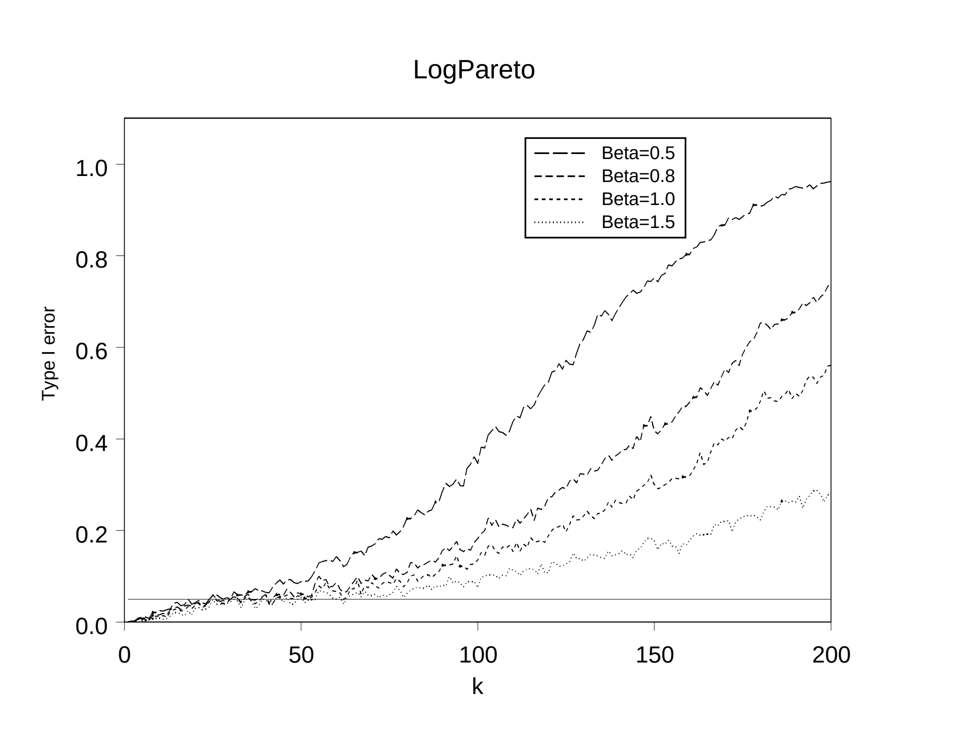
Empirical power



Empirical power at a nominal level $\bar{\alpha} = 0.05$

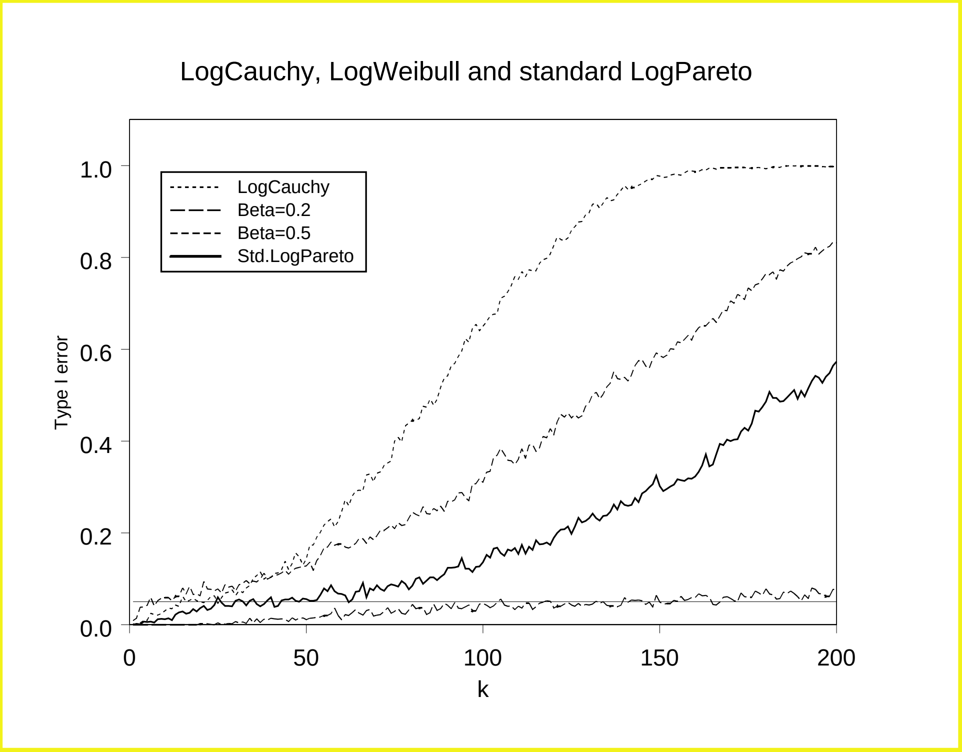
Underlying distribution function $F \in ERV_1$

Estimated type I error



Estimated type I error probability at a nominal level $\bar{\alpha} = 0.05$

Estimated type I error



Estimated type I error probability at a nominal level $\bar{\alpha} = 0.05$

References

- Csörgő and Horváth (1993). *Weighted Approximations in Probability and Statistics*. Wiley.
- de Haan, L. (1984). Slow variation and characterization of domains of attraction. *Statistical Extremes and Applications*.Reidel Publishing, Dordrecht.
- de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Austral. Math. Soc.*.

References

- Csörgő and Horváth (1993). *Weighted Approximations in Probability and Statistics*. Wiley.
- de Haan, L. (1984). Slow variation and characterization of domains of attraction. *Statistical Extremes and Applications*.Reidel Publishing, Dordrecht.
- de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Austral. Math. Soc.*.

Heavy Tails

- Adler, R. and Feldman, R. and Taqqu, M. S. (1998). *Practical Guide to Heavy Tails*. Birkhäuser Verlag.
- Heffernan, J. and Resnick, S. (2004). Limit laws for random vectores with an extreme component. `www.orie.cornell.edu/~sid/`
- Diebolt, J. and El-Aroui, M. and Garrido, M. and Girard, S. (2003). Quasi-conjugate bayes estimates for GPD parameters and application to heavy tails modeling. INRIA. Toulouse, France.

References

- Csörgő and Horváth (1993). *Weighted Approximations in Probability and Statistics*. Wiley.
- de Haan, L. (1984). Slow variation and characterization of domains of attraction. *Statistical Extremes and Applications*.Reidel Publishing, Dordrecht.
- de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Austral. Math. Soc.*.

Heavy Tails

- Adler, R. and Feldman, R. and Taqqu, M. S. (1998). *Practical Guide to Heavy Tails*. Birkhäuser Verlag.
- Heffernan, J. and Resnick, S. (2004). Limit laws for random vectores with an extreme component. `www.orie.cornell.edu/~sid/`
- Diebolt, J. and El-Aroui, M. and Garrido, M. and Girard, S. (2003). Quasi-conjugate bayes estimates for GPD parameters and application to heavy tails modeling. INRIA. Toulouse, France.

Super-Heavy tails

- Desgagné, A. and Angers, J.-F. (2004). Importance sampling with the generalized exponential power density. *Stat. Comp. (accepted)*.
- Mikosch, T. and Resnick, S. (2004). Activity rates with very heavy tails. `www.orie.cornell.edu/~sid/`
- Zeevi, A. and Glynn, P.W. (2004). Recurrence properties of autoregressive processes with super-heavy tailed innovations. *J. Appl. Probab. (to appear)*.