Dependence and Tail Modeling I

Evidence and Modeling of Heavy Tail Phenomena in Finance, Insurance and Telecommunications

[1]

Thomas Mikosch

University of Copenhagen

www.math.ku.dk/∼mikosch/Tehran10

[1] Tehran, May 19, 2010

Contents

1. Evidence of heavy tails in real-life data

1.1. Heavy tails in finance.

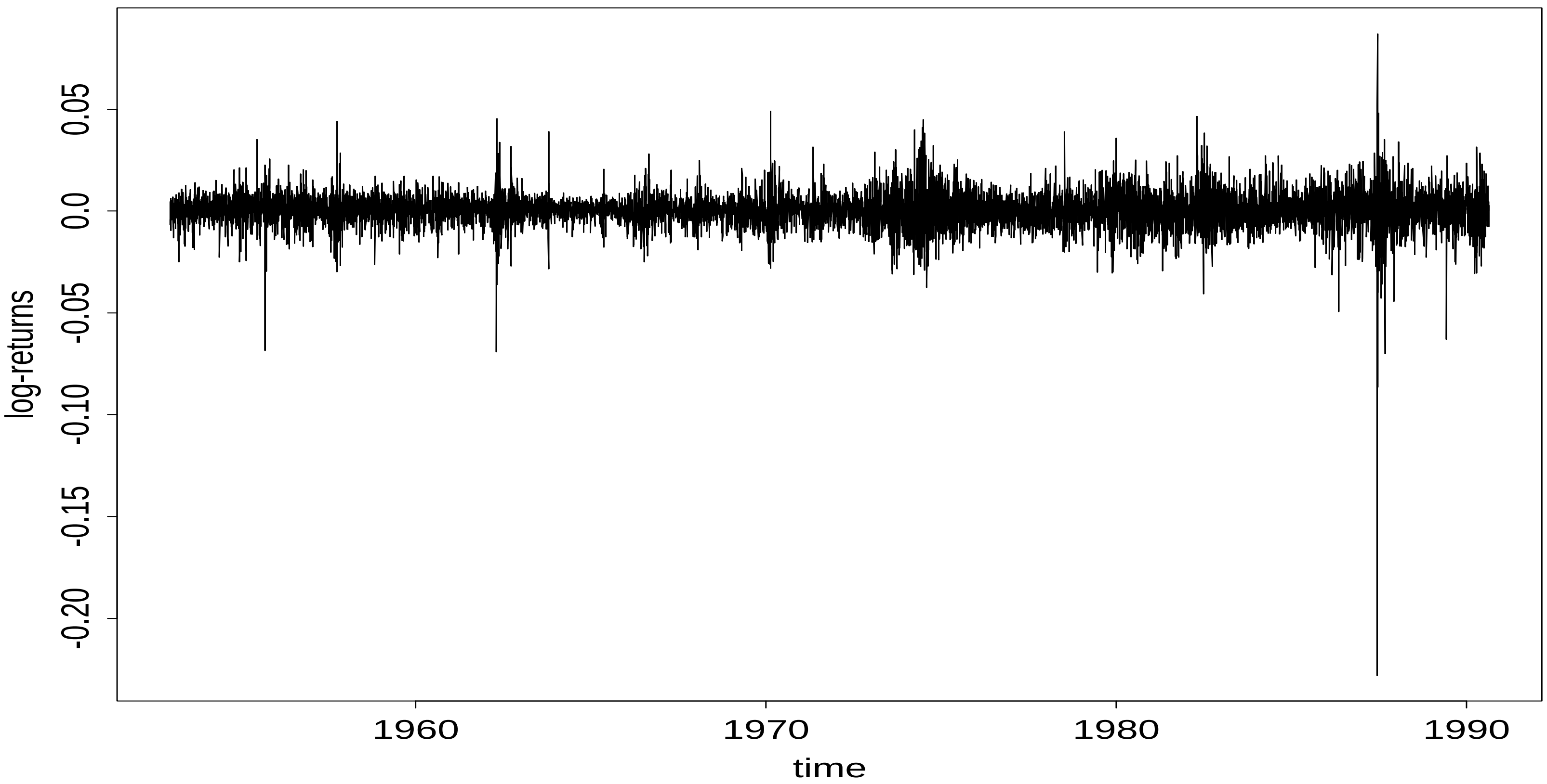

Figure 1. Plot of **9558** *S&P500* daily log-returns from January *2, 1953,* to December *31, 1990.* The year marks indicate the beginning of the calendar year.

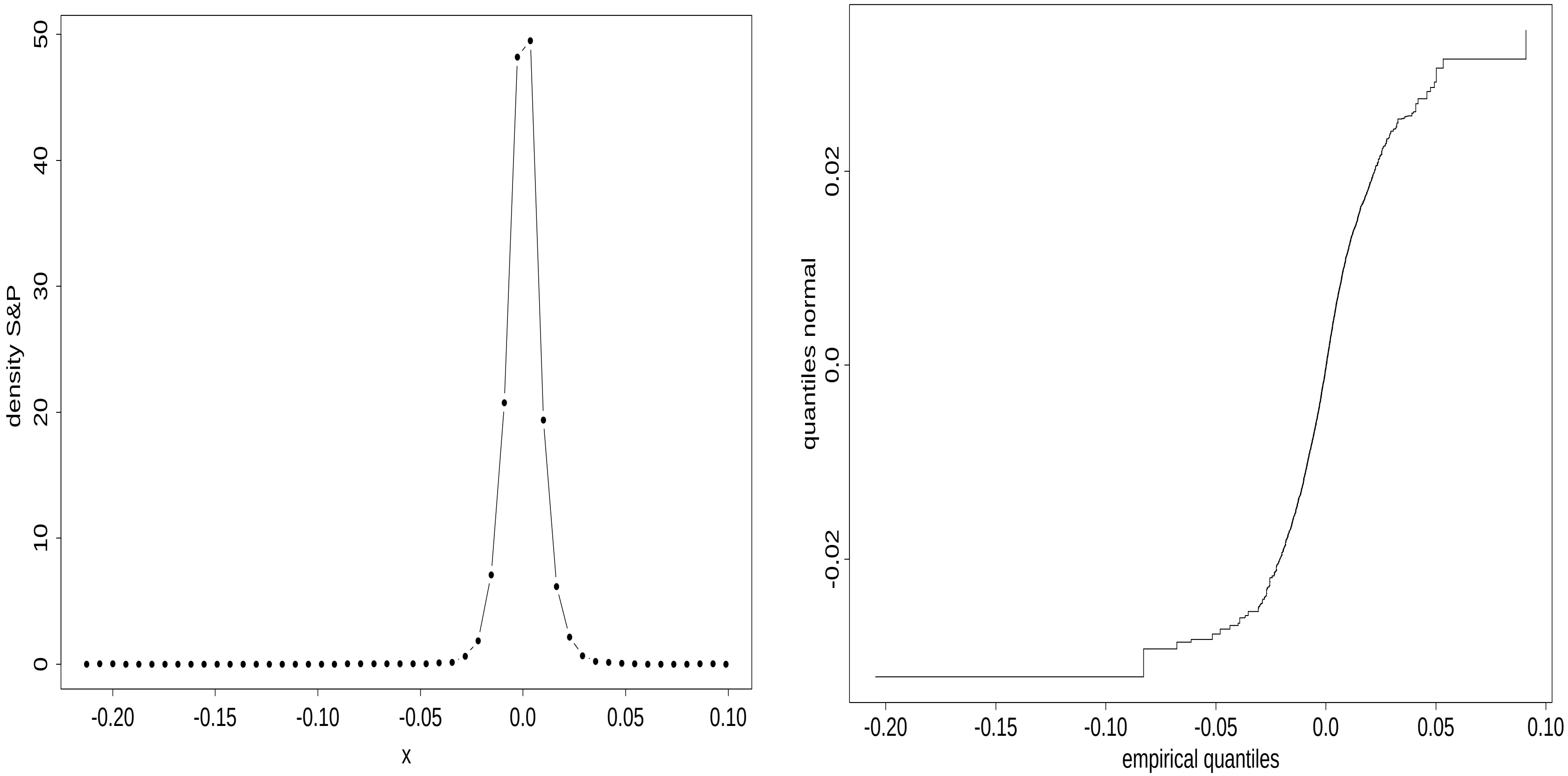

Figure 2. Left: Density plot of the *S&P500* data. The limits on the $\boldsymbol{x}$-axis indicate the range of the data. QQ-plot of the *S&P500* data against the normal distribution.

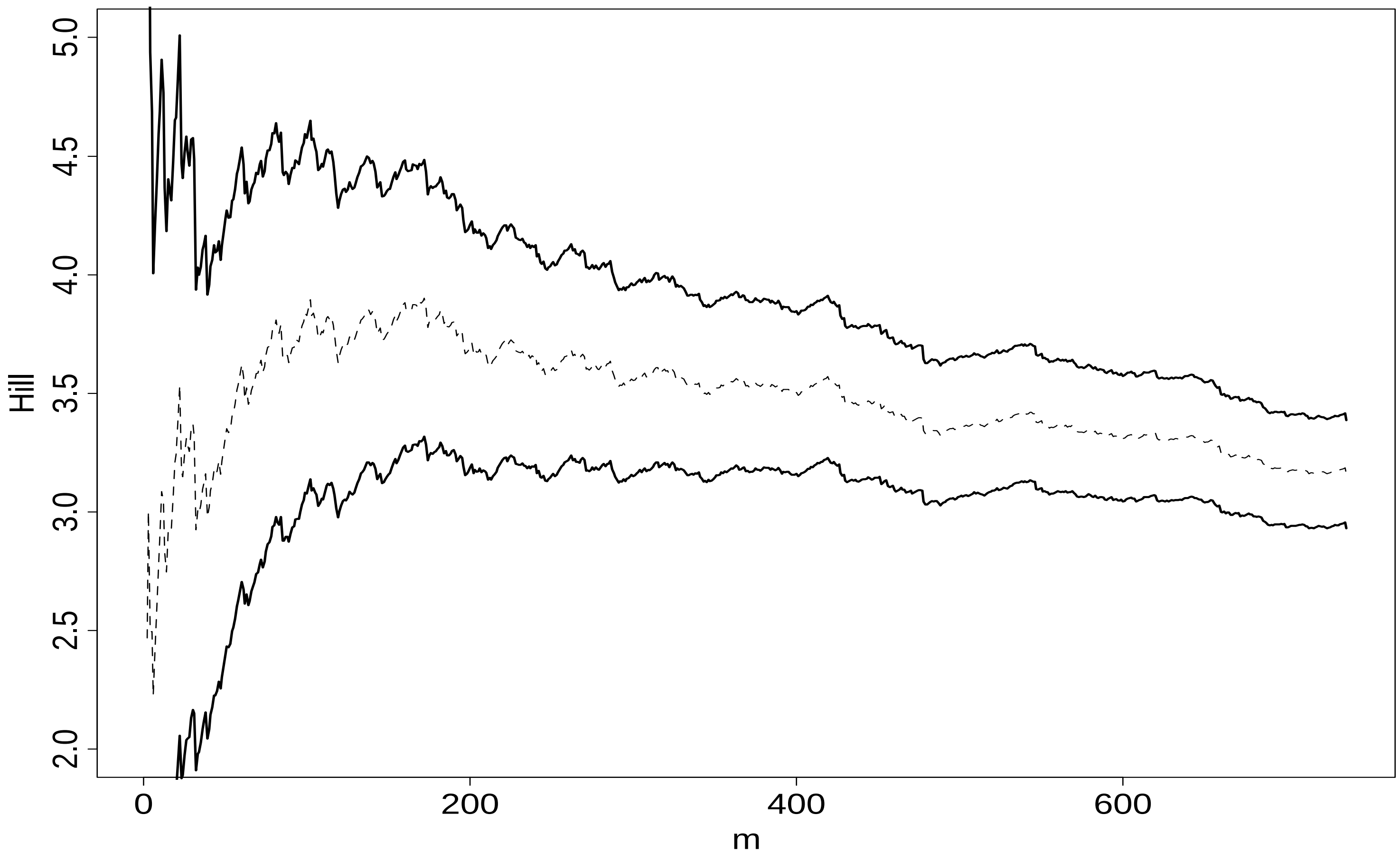

FIGURE 3. Hill plot (dotted line) for the *S&P500 data* with **95%** asymptotic confidence bounds. The Hill estimator approximates the tail index α in the model $P(X_1 > x) \sim c\,x^{-\alpha}$ as a function of the m upper order statistics in the return sample.

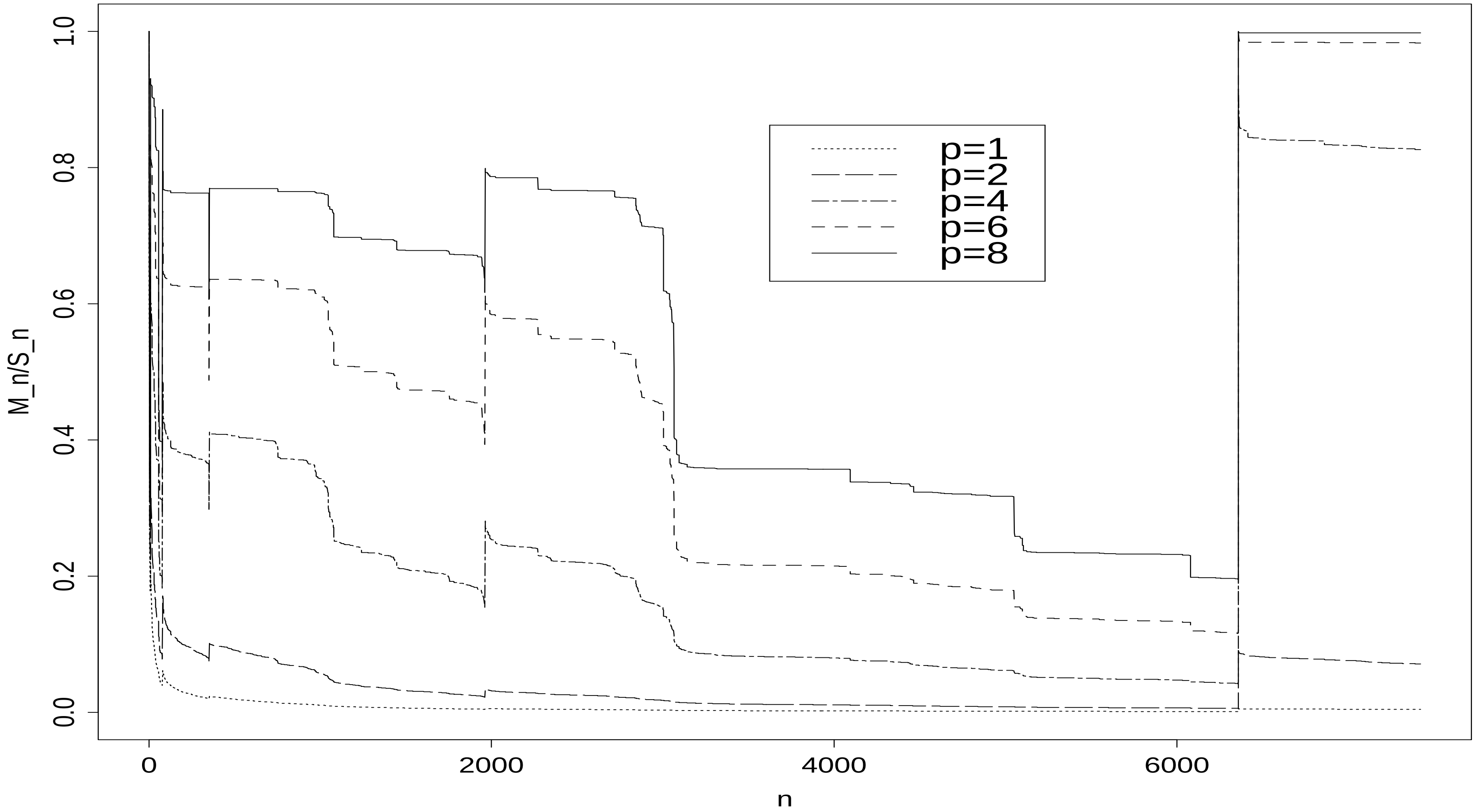

FIGURE 4. Plot of the ratio $T_n(p) = \max_{i=1,\ldots,n} |X_i|^p / (|X_1|^p + \cdots + |X_n|^p)$ for the *S&P500* data for various values of p. If $E|X_1|^p < \infty$ and the data came from a stationary ergodic model, the ratio should converge to zero a.s., by virtue of the strong law of large numbers.

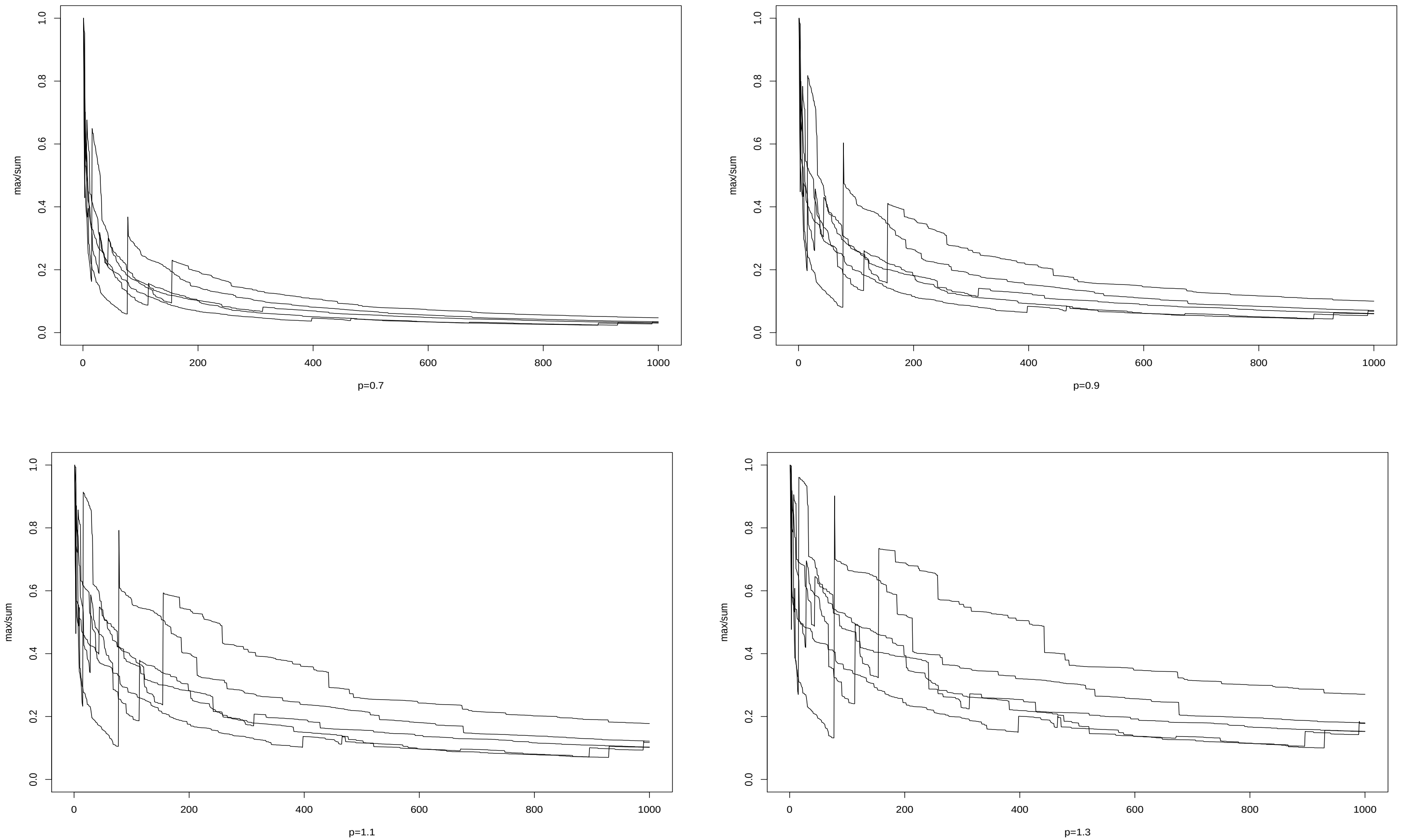

FIGURE 5. Plot of the ratio $T_n(p) = \max_{i=1,\dots,n} |X_i|^p/(|X_1|^p + \cdots + |X_n|^p)$ for iid simulated Cauchy variables with tail $P(|X| > x) \sim c\,x^{-1}$ and $p = 0.7, 0.9, 1.1, 1.3$

1.2. Heavy tails in insurance.

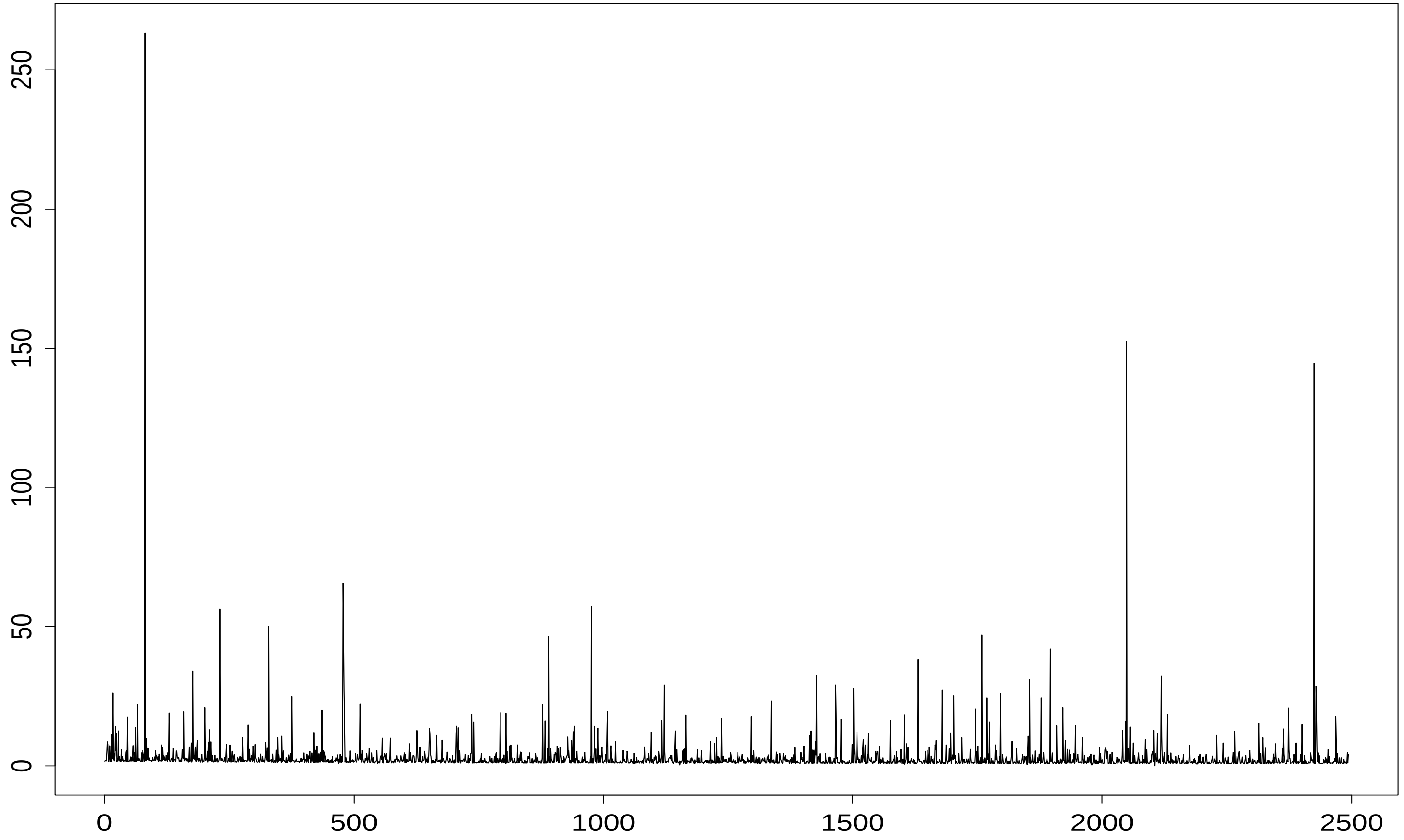

Figure 6. Danish fire insurance data.

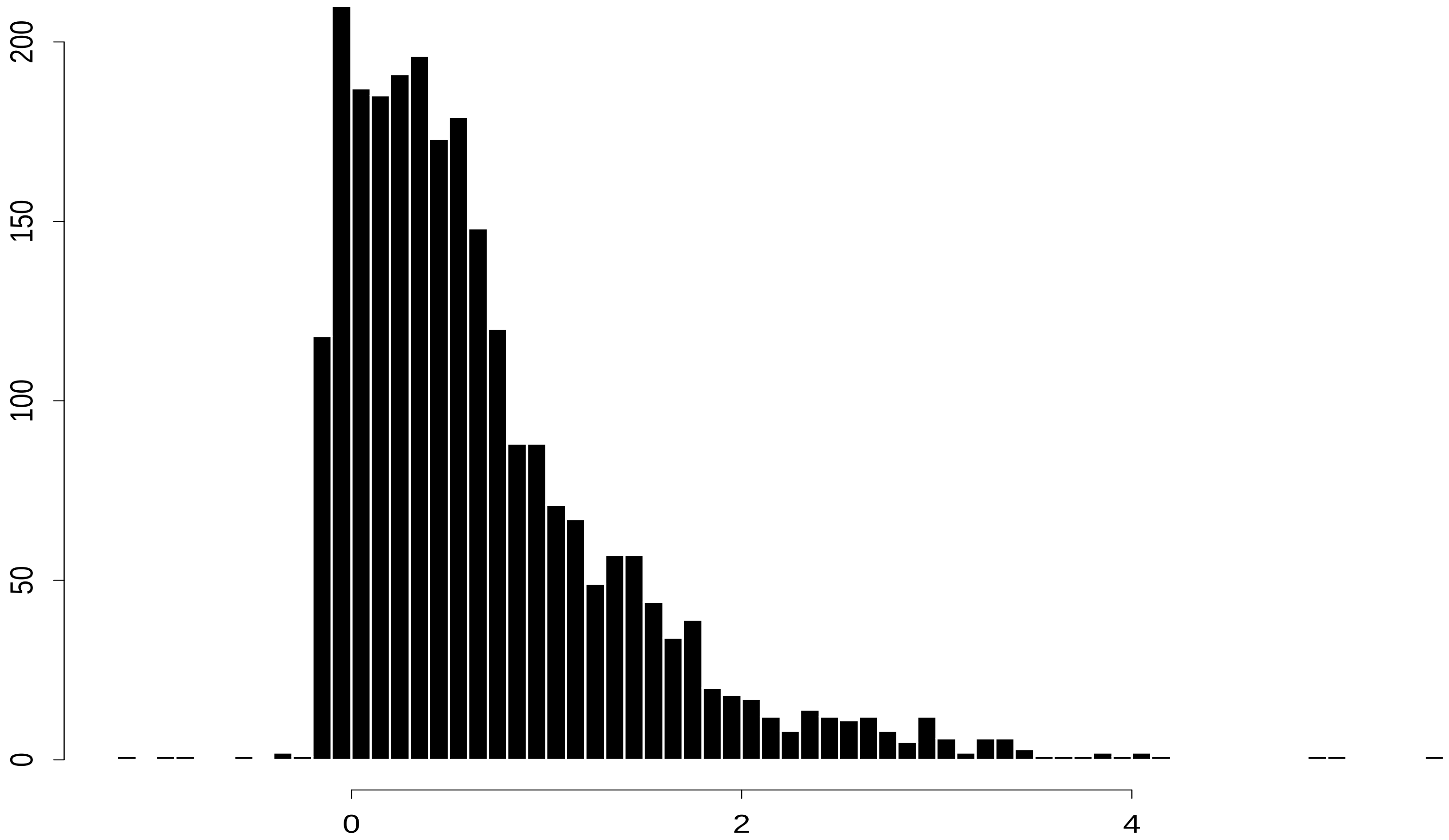

FIGURE 7. Histogram of the logarithmic Danish fire insurance data.

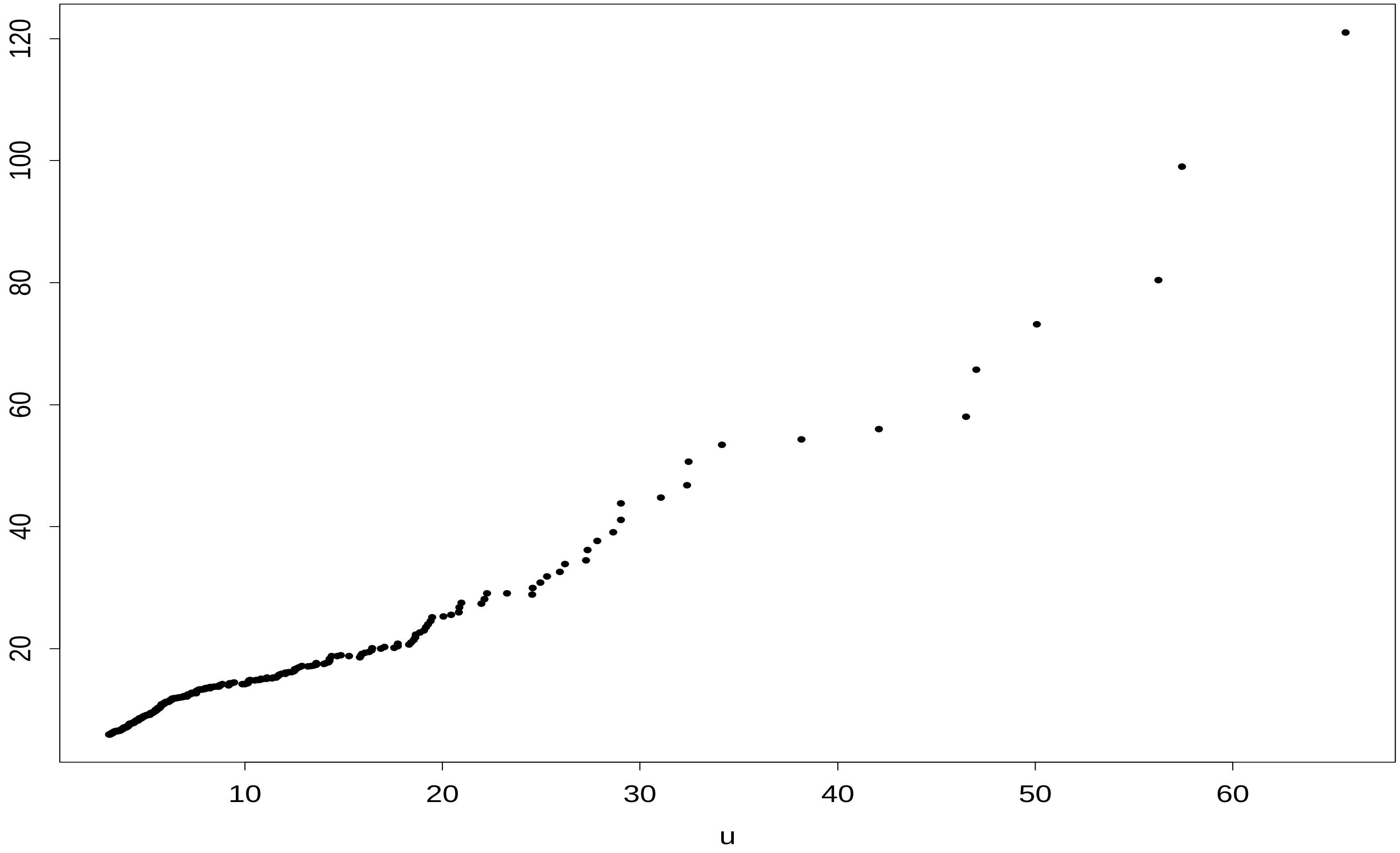

Figure 8. Empirical mean excess function of the Danish fire insurance data.

1.3. Heavy tails in teletraffic.

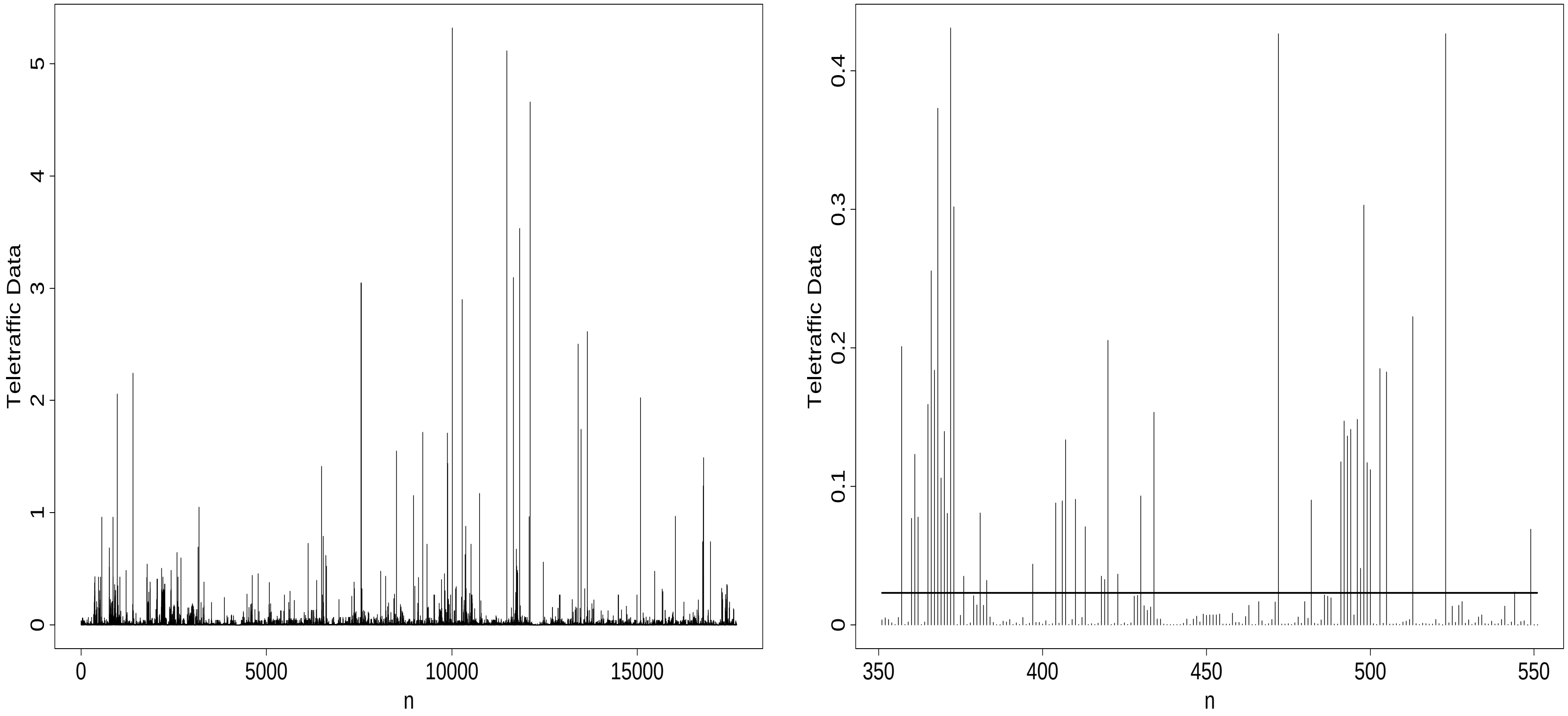

Figure 9. Time series of transmission durations (BU data).

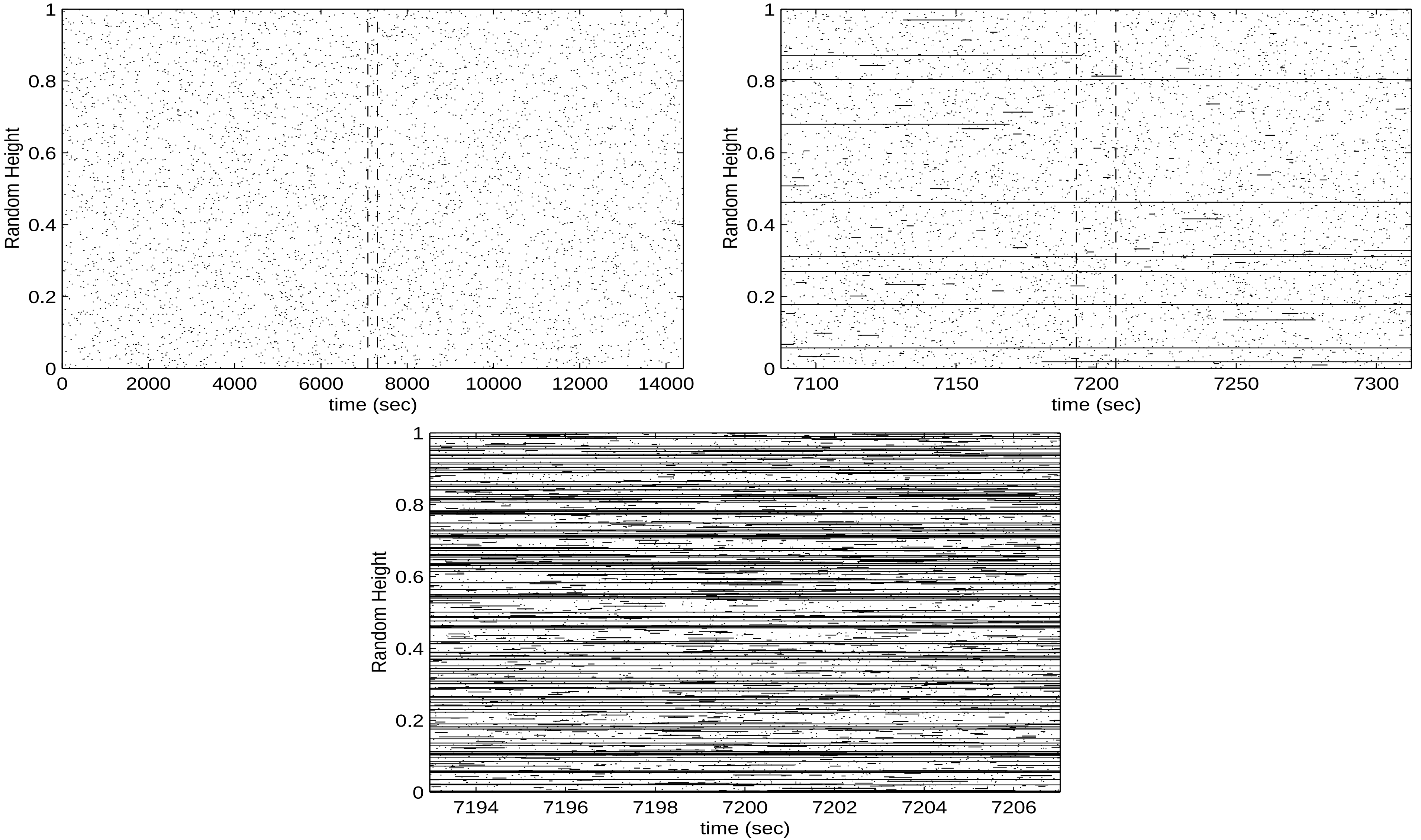

Figure 10. Mice and elephants plots (S. Marron).

2. Concepts of heavy-tailed distributions

2.1. Long-tailed distributions. **A positive random variable and its distribution[2] F with tail $\overline{F} = 1 - F$ are long-tailed (see Embrechts et al. (1997)) if**

$$\frac{\overline{F}(x-y)}{\overline{F}(x)} \to 1\,, \quad x \to \infty\,, \quad y \in \mathbb{R}\,. \tag{2.1}$$

- **Notice that**

$$P(X > x + y \mid X > x) \to 1\,, \quad x \to \infty\,, \quad y > 0\,.$$

- **The notion is inconvenient since the class of long-tailed distributions is too large.**

---

[2] All distributions are supposed to have infinite support.

- (2.1) is equivalent to slow variation of $L(x) = \overline{F}(\log x)$, i.e.

$$\frac{L(cx)}{L(x)} \to 1\,, \quad x \to \infty\,, \quad c > 0\,.$$

This means that $\overline{F}(x) = L(\mathrm{e}^x)$.

- A slowly varying function has Karamata representation, see Bingham et al. (1987)

$$L(x) = c(x)\, \exp\left\{-\int_z^x \frac{\varepsilon(t)}{t}\,dt\right\}\,, \quad x \geq z\,,$$

for some functions $\varepsilon(t) \to 0$ and $c(t) \to c > 0$ as $t \to \infty$.

- It satisfies for any $\delta > 0$ and sufficiently large x

$$x^{-\delta} \leq L(x) \leq x^{\delta}\,.$$

- **It is reasonable to define a heavy-tailed/light-tailed distribution in relation to some probabilistic structure and to study phase transitions of different behaviors of this structure when crossing borderlines.**

2.2. Subexponential distributions.

- A positive random variable X and its distribution F are subexponential if for iid copies X_i of X and any (some) $n \geq 2$, with $S_n = X_1 + \cdots + X_n$, $M_n = \max(X_1, \ldots, X_n)$,

$$\frac{P(S_n > x)}{P(M_n > x)} \sim \frac{P(S_n > x)}{n\,\overline{F}(x)} \to 1\,, \quad x \to \infty\,.$$

- Subexponential distributions are long-tailed, see EKM, p. 41
- Subexponential distributions do not have finite moment generating function. (Exponential moments do not exist.) See EKM, p. 42

- **Examples of subexponential distributions:** regularly varying distributions (power law tails), log-normal, heavy-tailed Weibull $\overline{F}(x) = \mathrm{e}^{-x^\tau}$, $\tau \in (0,1)$, $\overline{F}(x) = \mathrm{e}^{-x/\log^\gamma(x)}$, $\gamma > 0$.

- **Examples of non-subexponential distributions:** exponential, gamma, (truncated) normal, any distribution with finite upper endpoint.

- Subexponential distributions are considered as natural heavy-tailed distributions in the context of insurance mathematics, queuing, storage, dam, renewal theory.

- **There one is interested in the tail behavior of a random walk $(S_{N(t)})$ with iid step sizes X_i, independent of the counting process N. For a fixed $t > 0$, with $p_n = P(N(t) = n)$,**

$$P(S_{N(t)} > x) = \sum_{n=1}^{\infty} p_n \, P(S_n > x) \, .$$

- **If X_i is subexponential, then $P(S_n > x)/\overline{F}(x) \le K(\varepsilon)\,(1+\varepsilon)^n$.**

  See EKM p. 41.

- **Hence if $Ee^{hN(t)} < \infty$ for some $h > 0$,**

$$\frac{P(S_{N(t)} > x)}{\overline{F}(x)} = \sum_{n=1}^{\infty} p_n \frac{P(S_n > x)}{\overline{F}(x)} \sim \sum_{n=1}^{\infty} p_n \, n = EN(t) \, .$$

- **The total claim amount $S_{N(t)}$ of an insurance portfolio at a high threshold: $P(S_{N(t)} > x) \sim EN(t)\,\overline{F}(x)$.**

- If N is homogeneous Poisson with intensity $\lambda > 0$, the ruin probability of the portfolio is given by

$$\psi(x) = P\left(\inf_{t\geq 0}(x + c\,t - S_{N(t)}) < 0\right)$$
$$= \rho(1+\rho)^{-1}\sum_{n=1}^{\infty}(1+\rho)^{-n}\,P(S_n^* > x)\,,$$

where $\rho = c/(\lambda EX) - 1$ is assumed positive and (S_n^*) is a random walk with iid positive step sizes with distribution $F_*(x) = (EX)^{-1}\int_0^x \overline{F}(t)\,dt$.

- **TFAE** EKM, p. 581 **(1)** F_* is subexponential, **(2)** $1-\psi$ is subexponential, **(3)** the following relation holds

$$\frac{\psi(x)}{\overline{F}_*(x)} \sim \rho(1+\rho)^{-1}\sum_{n=1}^{\infty}(1+\rho)^{-n}\,n = \rho^{-1}\,.$$

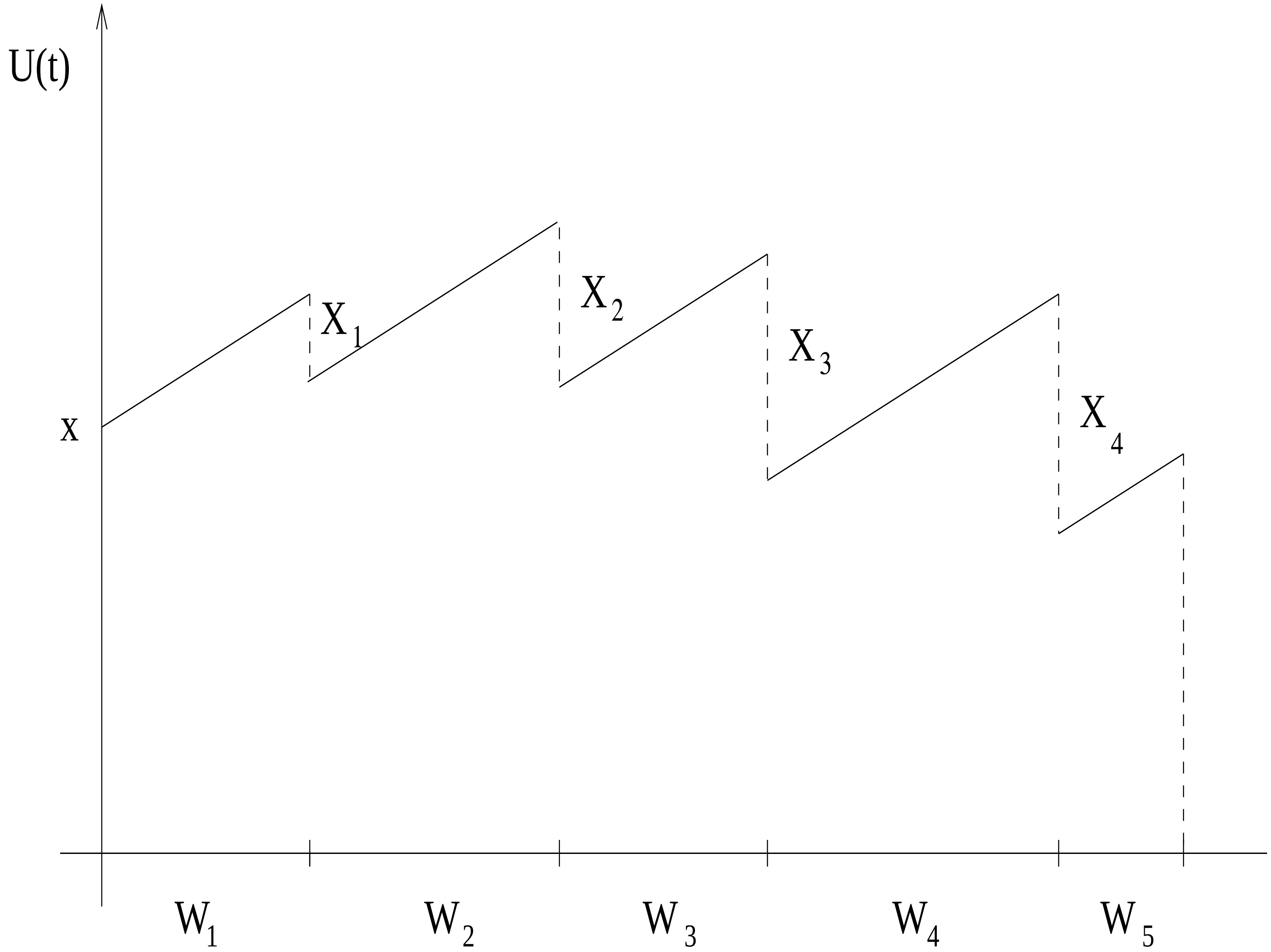

Figure 11. Ruin in an insurance portfolio.

- **In contrast, if F is light-tailed in the sense that $E\mathrm{e}^{hX} < \infty$ in some neighborhood of the origin, then $\psi(x)$ decays exponentially fast as $x \to \infty$.** Cramér (1930)
- **These results describe a phase transition of different behaviors of the ruin probability when passing from light-tailed to heavy-tailed distributions.**

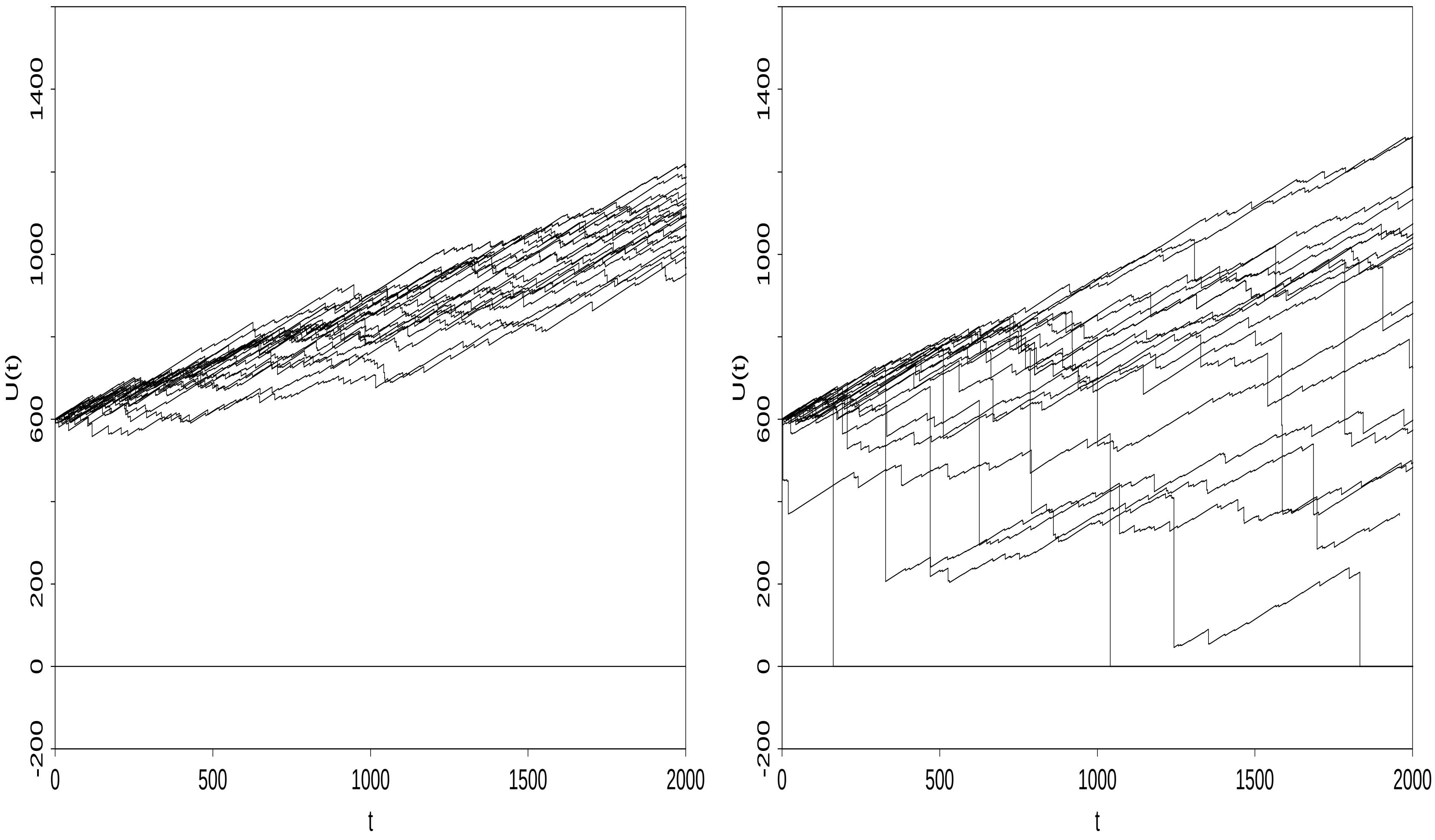

FIGURE 12. Ruin in a insurance portfolio. Light tails (left) and heavy tails (right).

2.3. Regularly varying distributions.

- A positive measurable function on $(0,\infty)$ is **regularly varying with index** $\rho \in \mathbb{R}$ if $f(x) = L(x)\, x^{\rho}$ for some slowly varying function L. Bingham et al. (1987)
- A positive function f on $(0,\infty)$ is regularly varying **if and only if** as $x \to \infty$, $\dfrac{f(cx)}{f(x)} \to c^{\rho}\,, c > 0$.
- A positive random variable X and its distribution F are **regularly varying with index** $\alpha > 0$ if for some slowly varying function L

$$\overline{F}(x) = P(X > x) = \frac{L(x)}{x^{\alpha}}, \quad x > 0\,.$$

- **Regularly varying distributions are subexponential** see e.g. Feller (1971), EKM p. 37, **hence long-tailed.**
- **Examples:**

Pareto

log-gamma

infinite variance stable

Cauchy

student

Fréchet.

- **Regular variation is a natural condition in the context of extreme value theory and limit theory for partial sums of iid random variables: For an iid non-negative sequence (X_i) with distribution $\overline{F}(x) = P(X > x) = \dfrac{L(x)}{x^\alpha}$, $x > 0$,**

$$\textbf{Fréchet limit:}\quad n^{-1/\alpha}\,\ell(n)\,M_n \xrightarrow{d} Y_M \sim \Phi_\alpha\,,\quad \alpha > 0\,,$$

$$\textbf{Stable limit:}\quad n^{-1/\alpha}\,\ell(n)\,(S_n - b_n) \xrightarrow{d} Y_S \sim P_\alpha\,,\quad \alpha \in (0,2)\,.$$

The Fréchet distribution $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$, $x > 0$, is one of the max-stable or extreme value distributions. The distribution P_α is called α-stable.

See Feller (1971), Ibragimov, Linnik (1971), Petrov (1975,1995) for sums; Galambos (1978), Leadbetter et al. (1983), Resnick (1987,2006), EKM for maxima

- **There is also joint convergence for $\alpha \in (0,2)$** see Resnick (1986)**:**

$$n^{-1/\alpha}\ell(n)(M_n, S_n - b_n) \xrightarrow{d} (Y_M, Y_S)\,,$$

and Y_S, Y_M are dependent.

- **By the continuous mapping theorem,**

$$(S_n - b_n)/M_n \xrightarrow{d} Y_S/Y_M\,.$$

- **In particular, for $\alpha \in (0,1)$,**

$$S_n/M_n \xrightarrow{d} Y_S/Y_M$$

and for $\alpha \in (1,2)$

$$(S_n - n\,EX)/M_n \xrightarrow{d} Y_S/Y_M\,.$$

- If $\alpha \geq 2$, there exist $a_n, b_n \to \infty$ such that

$$(a_n^{-1}(S_n - n\,EX)\,, b_n^{-1}\,M_n) \stackrel{d}{\to} (Y_S, Y_M)\,,$$

Y_S, Y_M are independent, $Y_S \sim N(0,1)$, $Y_M \sim \Phi_\alpha$ and $b_n/a_n \to 0$.

- In particular,

$$M_n/(S_n - n\,EX) \stackrel{P}{\to} 0\,.$$

- For general iid non-negative X_i, $M_n/S_n \stackrel{P}{\to} 0$ if and only if $EX < \infty$ or $P(X > x) = L(x)x^{-1}$ for some slowly varying L O'Brien (1980), and $M_n/S_n \stackrel{P}{\to} 1$ if and only if $\overline{F}(x) = L(x)$ Arov and Bobrov (1960), Maller and Resnick (1984).

2.4. Alternative definitions of regular variation.

- $X > 0$ is regularly varying with index $\alpha > 0$ if and only if

$$\frac{P(X > tx)}{P(X > x)} \to t^{-\alpha}, \quad x \to \infty, \quad t > 0.$$

- Replacing x by a_n with $P(X > a_n) \sim n^{-1}$, one can show equivalence with

$$n\,P(a_n^{-1}X > t) \to t^{-\alpha}, \quad n \to \infty, \quad t > 0. \tag{2.2}$$

- For iid copies X_i of X, (2.2) has the interpretation

$$E\Big(\sum_{i=1}^{n} I_{(t,\infty)}(X_i/a_n)\Big) \to t^{-\alpha}, \quad t > 0.$$

Recalling Poisson's limit theorem, (2.2) is equivalent to

$$\sum_{i=1}^{n} I_{(t,\infty)}(X_i/a_n) \xrightarrow{d} \text{Poisson}(t^{-\alpha}), \quad t > 0,$$

- (2.2) is equivalent to the point process convergence

$$N_n = \sum_{i=1}^{n} \varepsilon_{X_i/a_n} \xrightarrow{d} N \sim \text{PRM}(\mu),$$

where $\mu(t,\infty] = t^{-\alpha}$, $t > 0$, defines the mean measure of the limiting Poisson random measure N with state space $(0,\infty]$.

- (2.2) can be shown to be equivalent to vague convergence of the measures

$$n\,P(a_n^{-1} X \in \cdot) \xrightarrow{v} \mu(\cdot) \quad \text{on } (0,\infty].$$

3. Some objectives in modeling extremal and heavy tail phenomena

- **The classical approach to EVT and extreme value statistics.**
  Assume (X_i) iid. Find appropriate limit distributions H, constants $c_n > 0$, $d_n \in \mathbb{R}$ such that $c_n^{-1}(M_n - d_n) \stackrel{d}{\rightarrow} Y \sim H$. (Fisher-Tippett theorem, 1928), **see EKM, p. 121****.**

- **Use this limit relation for estimating tails $\overline{F}(x)$ for large x and high quantiles $F^{\leftarrow}(p)$, *possibly outside the range of the data.***

de Haan and Fereira (2006)

- **More recent problems. EVT and statistics for spatio-temporal structures: stationary or non-stationary time series and random fields. Study the extremal properties of these structures.**
- **Multivariate structures: dependence measures for extremes beyond covariances. The interplay between tails and extremal dependence**
- **Infinite-dimensional structures: EVT for stochastic processes and random fields**
- **Use notions from EVT to build models for description of heavy tail phenomena, e.g. in telecommunications (file sizes, transmission durations, transmission rates).**

4. Some point process theory

The theory of point processes plays a central role in extreme value theory. Applications include:

- **Derivation of joint limiting distribution of order statistics, i.e., k^{th} largest order statistic, limiting distribution of maximum and minimum, etc.**
- **Calculation of limit distribution of exceedances of a high level.**
- **Extensions to stationary processes.**
- **Provides a useful tool in heavy-tailed case for deriving limiting behavior of various statistics, e.g., sample mean, sample autocovariances, etc, which are often determined by the behavior of the extreme order statistics.**

4.1. Definition and basic results.

- Suppose (X_t) is an iid sequence with common distribution F.
- Assume that there exist sequences of constants $a_n > 0$ and b_n such that

$$P(a_n^{-1}(M_n - b_n) \le x) = F^n(a_n x + b_n) \to G(x) \tag{4.1}$$

for all x, where $M_n = \max(X_1, \ldots, X_n)$ and G is a nondegenerate distribution function.

- By extremal types theorem, G has to be an extreme value distribution of which there are only three types, see Leadbetter et al. (1983) or EKM.

- Taking logarithms and using a Taylor series expansion, (4.1) holds if and only if for any $x \in \mathbb{R}$,

$$n\,P(a_n^{-1}(X_1 - b_n) > x) \to -\log G(x)\,. \tag{4.2}$$

(If $G(x) = 0$ we interpret $-\log G(x)$ as ∞.)

- Now (4.2) can be strengthened to the statement,

$$n\,P(a_n^{-1}(X_1 - b_n) \in B) \to \nu(B) \tag{4.3}$$

for all suitably chosen Borel sets B, where the measure ν is defined by its value on intervals of the form $(a, b]$ as

$$\nu(a, b] = \log G(b) - \log G(a)\,.$$

The convergence in (4.3) can be connected with the convergence in distribution of a sequence of point processes.

For a bounded Borel set B in the product space $(0,1] \times \mathbb{R}$, define the sequence of point processes (N_n) by

$$N_n(B) = \#\{(t/n, a_n^{-1}(X_t - b_n)) \in B\,, t = 1, \ldots, n\}$$
$$= \sum_{t=1}^{n} \varepsilon_{(t/n, a_n^{-1}(X_t - b_n))}(B)\,,$$

where ε_y is the Dirac measure at the point y.

Properties:

- **If B is the rectangle $(a,b] \times (c,d]$ with $0 \le a < b \le 1$ and $-\infty < c < d < \infty$, then since the X_j are iid, $N_n(B) \sim \mathrm{Bin}([nb] - [na], p_n)$ ($[s]$ = integer part of s), and**

$$p_n = P(a_n^{-1}(X_1 - b_n) \in (c,d])\,.$$

- Provided $\nu(c,d] < \infty$, it follows from

$$n\, P(a_n^{-1}(X_1 - b_n) \in B) \to \nu(B)$$

that $N_n(B)$ converges in distribution to a Poisson random variable $N(B)$ with mean $\mu(B) = (b-a)\,\nu(c,d]$.

- In fact, we have the stronger point process convergence,

$$N_n \stackrel{d}{\to} N\,, \tag{4.4}$$

where N is a Poisson process on $(0,1] \times \mathbb{R}$ with mean measure $\mu(dt, dx) = dt \times \nu(dx)$ and $\stackrel{d}{\to}$ denotes convergence in distribution of point processes.

4.2. **Convergence for point processes.** For our purposes, $\stackrel{d}{\to}$ for point processes means that for any collection of *bounded* Borel sets $B_1, \ldots, B_k$ for which $P(N(\partial B_j) > 0) = 0$, $j = 1, \ldots, k$, we have

$$(N_n(B_1), \ldots, N_n(B_k)) \stackrel{d}{\to} (N(B_1), \ldots, N(B_k))$$

on $\mathbb{R}^k$; see EKM, Leadbetter et al. (1983), Resnick (1987).

Technical remarks:

- In the heavy-tailed case, the state space of the point process is often defined to be $(0, 1] \times ([-\infty, \infty] \setminus \{0\})$.
- For the space, $[-\infty, \infty] \setminus \{0\}$, the roles of zero and infinity have been interchanged so that bounded sets are now those sets which are bounded away from 0.

- **A bounded set on the product space is contained in the rectangle $[0, c] \times ([-\infty, -d] \cup [d, \infty])$ for some positive and finite constants c and d. Under this topology, the intensity measure of the limit Poisson process is ensured to be finite on all bounded Borel sets.**

Application:

Define $M_{n,2}$ to be the second largest among $X_1, \ldots, X_n$. The event $\{a_n^{-1}(M_{n,2} - b_n) \le y\}$ is the same as $\{N_n((0,1] \times (y, \infty)) \le 1\}$, we conclude from (4.4) that

$$\begin{aligned} P(a_n^{-1}(M_{n,2} - b_n) \le y) &= P(N_n((0,1] \times (y,\infty)) \le 1) \\ &\to P(N((0,1] \times (y,\infty)) \le 1) \\ &= G(y)\,(1 - \log G(y))\,. \end{aligned}$$

Similarly, the joint limiting distribution of $(M_n, M_{n,2})$ can be calculated by noting that for $y \le x$,

$$\{a_n^{-1}(M_n - b_n) \le x, a_n^{-1}(M_{n,2} - b_n) \le y\}$$

$$= \{N_n((0,1] \times (x,\infty)) = 0, N_n((0,1] \times (y,x]) \le 1\}.$$

Hence,

$$P(a_n^{-1}(M_n - b_n) \le x, a_n^{-1}(M_{n,2} - b_n) \le y)$$

$$= P(N_n((0,1] \times (x,\infty)) = 0, N_n((0,1] \times (y,x]) \le 1)$$

$$\to P(N((0,1] \times (x,\infty)) = 0, N((0,1] \times (y,x]) \le 1))$$

$$= G(y)(1 + \log G(x) - \log G(y)) \,.$$

4.3. **More on convergence of point processes:** If E denotes the state-space for our point measures, (e.g., in the heavy-tailed case $E = (0,1] \times ([-\infty,\infty] \setminus \{0\})$ or $E = (0,1] \times (0,\infty])$, define

$$M_p(E) = \text{Radon point measures}$$

$$C_K(E) = \text{continuous functions on } E \text{ with compact support.}$$

Here $M_p(E)$ is endowed with the topology induced by vague convergence so that

$$m_n \to m \text{ iff } m_n(f) \to m(f) \text{ for all } f \in C_K(E).$$

It follows that for a sequence of point processes (N_n),

$$N_n \xrightarrow{d} N \text{ iff } N_n(f) \xrightarrow{d} N(f) \text{ for all } f \in C_K(E).$$

4.4. **Laplace functional convergence.** Sometimes it is more convenient to work with Laplace functionals in which case $N_n \stackrel{d}{\rightarrow} N$ is equivalent to

$$E\left(\exp\{-N_n(f)\}\right) \rightarrow E\left(\exp\{-N(f)\}\right) \quad f \in C_K^+(E). \tag{4.5}$$

Remarks:

- The test functions f in (4.5) can be discontinuous as long as the f-discontinuity set D_f satisfies

$$P(N(D_f) = 0) = 1$$

- If N is a Poisson random measure with intensity measure λ, then

$$E\left(\exp\{-N(f)\}\right) = \exp\left\{-\int_E (1 - \mathrm{e}^{-f(x)})\, \lambda(dx)\right\}.$$

- **Try taking $f = 1_{(y,\infty)}(x)$ to show that $N(f) = N((y,\infty))$ has a Poisson distribution with mean $\lambda(f) = \lambda(y,\infty)$.**
- **Suppose (X_t), $X_t \geq 0$ a.s. is an iid regularly varying sequence with index $\alpha > 0$. Then**

$$n\,P(X_1 > a_n x) \to x^{-\alpha}\,, \quad x > 0\,,$$

where $P(X_1 > a_n) \sim n^{-1}$. With $E = (0,\infty]$, then one has

$$\mu_n(\cdot) = n\,P(X_1/a_n \in \cdot) \to \mu(\cdot)$$

on $M_p(E)$, where $\mu(x,\infty] = x^{-\alpha}$, $x > 0$. In particular,

$$\mu_n(g) \to \mu(g) \text{ for all } g \in C_K(E).$$

Since $f \in C_K^+(E)$ implies $g = (1 - \mathrm{e}^{-f}) \in C_K^+(E)$,

$$\mu_n(g) = n\,E(1 - \mathrm{e}^{-f(X_1/a_n)}) \to \mu(g) = \int_E (1 - \mathrm{e}^{-f(x)})\mu(dx)$$

- Now consider the sequence of point processes

  $N_n = \sum_{t=1}^{n} \varepsilon_{a_n^{-1} X_t}$. with Laplace functional

$$
\begin{aligned}
E(\exp\{-N_n(f)\}) &= E(\exp\{-\sum_{t=1}^{n} f(X_t/a_n)\}) \\
&= (E(\exp\{-f(X_1/a_n)\}))^n \\
&= (1 - n\, E\left(1 - \exp\{-f(X_1/a_n)\}\right)/n)^n \\
&\to \exp\{-\mu(g)\} \\
&= E\exp\{-N(f)\}\,.
\end{aligned}
$$

- Thus, $N_n \xrightarrow{d} N$.
- The points of the limit Poisson process with mean measure

  $\mu(x,\infty) = x^{-\alpha}$, $x > 0$, can be displayed in an explicit fashion.

- Set $\Gamma_k = E_1 + \cdots + E_k$, where $E_1, E_2, \ldots$ are iid unit exponentials. These are the points of a unit rate homogeneous Poisson process on $(0, \infty)$.

- Then

$$N_n = \sum_{t=1}^{n} \varepsilon_{a_n^{-1} X_t} \xrightarrow{d} N \stackrel{d}{=} \sum_{k=1}^{\infty} \varepsilon_{\Gamma_k^{-1/\alpha}} .$$

- If we order the data, then we can read off the weak convergence for the k^{th}-largest $M_{n,k}$, i.e.,

$$a_n^{-1} M_{n,k} \xrightarrow{d} \Gamma_k^{-1/\alpha} .$$

(These are joint in k as well.)

4.5. Continuous mapping theorem: partial sum convergence.

- **We illustrate the power of point process convergence in a simple application of the continuous mapping theorem.**
- **Recall the continuous mapping theorem. If $N_n \stackrel{d}{\rightarrow} N$ and $T : M_p(E) \rightarrow \mathbb{R}$ is an a.s. continuous mapping (relative to N), then $T(N_n) \stackrel{d}{\rightarrow} T(N)$.**
- **Application to partial sums: Suppose (X_t) is an iid sequence of positive random variables which are regularly varying with index $\alpha < 1$. Then**

$$N_n = \sum_{t=1}^{n} \varepsilon_{a_n^{-1} X_t} \stackrel{d}{\rightarrow} N = \sum_{k=1}^{\infty} \varepsilon_{\Gamma_k^{-1/\alpha}}$$

- For $m = \sum_j \varepsilon_{y_j}$ define the mapping $T_\epsilon(m) = \sum_j y_j 1_{\{y_j > \epsilon\}}$ which is a.s. continuous relative to the limit point process N.

- By CMT,

$$T_\epsilon(N_n) = \sum_{t=1}^{n} a_n^{-1} X_t 1_{\{X_t > \epsilon a_n\}} \xrightarrow{d} T_\epsilon(N) = \sum_{k=1}^{\infty} \Gamma_k^{-1/\alpha} I_{\{\Gamma_k^{-1/\alpha} > \epsilon\}}$$

Now, as $\epsilon \downarrow 0$,

$$T_\epsilon(N) \to S = \sum_{k=1}^{\infty} \Gamma_k^{-1/\alpha} \text{ a.s.}$$

and (with $S_n = \sum_{t=1}^n X_t$)

$$\begin{aligned} E|a_n^{-1} S_n - T_\epsilon(N_n)| &\leq n a_n^{-1} E X_1 1_{(X_1 \leq \epsilon a_n)} \\ &\sim \frac{\alpha}{1-\alpha} n a_n^{-1} \epsilon a_n P(X_1 > a_n \epsilon) \text{ (by Karamata's theorem)} \\ &\to \frac{\alpha}{1-\alpha} \epsilon^{1-\alpha} \text{ as } n \to \infty \\ &\to 0 \text{ as } \epsilon \to 0\,. \end{aligned}$$

- **It follows from a standard weak convergence result that**

$$a_n^{-1} S_n \xrightarrow{d} S = \sum_{k=1}^{\infty} \Gamma_k^{-1/\alpha} .$$

The limit is the series representation of an α-stable random variable.

References

[1] Arov, D.Z. and Bobrov, A.A. (1960) The extreme members of a sample and their role in the sum of independent variables. (Russian) *Teor. Verojatnost. i Primenen.* **5**, 415–435.

[2] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[3] Cramér, H. On the Mathematical Theory of Risk. Skandia Jubilee Volume. Reprinted in: Martin-Löf, A. (Ed.) 1994 Cramér, H. *Collected Works.* Springer, Berlin.

[4] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[5] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications.* Vol. II. Second edition. Wiley, New York.

[6] Galambos, J. (1978) *The Asymptotic Theory of Order Statistics.* Wiley, New York.

[7] Haan, L. de, Fereira, A. (2006) *Extreme Value Theory: An Introduction.* Springer, Berlin, New York.

[8] Ibragimov, I.A. and Linnik, Yu.V. (1971) *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.

[9] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

[10] Maller, R.A. and Resnick, S.I. (1984) Limiting behaviour of sums and the term of maximum modulus. *Proc. London Math. Soc.* **49** , 385–422.

[11] O'Brien, G. L. (1980) A limit theorem for sample maxima and heavy branches in Galton-Watson trees. *J. Appl. Probab.* **17**, 539–545.

[12] Petrov, V.V. (1975) *Sums of Independent Random Variables.* Springer, Berlin.

[13] Petrov, V.V. (1995) *Limit Theorems of Probability Theory.* Clarendon Press, Oxford.

[14] Resnick, S.I. (1986) Point processes, regular variation and weak convergence. *Adv. Appl. Prob.* **18**, 66–138.

[15] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[16] Resnick, S.I. (2006) *Heavy Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York.