Applied Probability Modeling of Data Networks;

1. Random measures and point processes; weak convergence to PRM and Lévy processes.

Sidney Resnick
School of Operations Research and Industrial Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA
http://legacy.orie.cornell.edu/∼sid
sir1@cornell.edu

August 5, 2007

CORNELL

1. Outline of Lectures

1. Random measures and point processes; weak convergence to PRM and Lévy processes.
2. Multivariate regular variation; the Poisson transform; stable processes.
3. Introduction and survey of data network modeling. (Fairly non-technical; some statistics.)
4. Some stylized applied probability network models: a renewal input model.
5. Some (more) stylized applied probability network models: a large time scale input model.

5.5 Some (more more) stylized applied probability network models: Heavy traffic and heavy tails in GI/G/1.

Lectures primarily based on Resnick (2006).

2. Introduction: Point processes and random measures.

2.1. Setup.

(Resnick, 2006, Chapter 3)
Setup for study of random measures, point processes and associated topology yielding *vague convergence*.

- $\mathbb{E}$: a nice space (lccb). Typically $\mathbb{E}$ is a finite dimensional Euclidean space: Subset of compactified $\mathbb{R}^d$ or $\mathbb{R}^d_+$.
- $\mathcal{E}$: Borel σ-algebra.
- Space of all Radon measures on $\mathbb{E}$:

$$M_+(\mathbb{E}) = \{\mu : \mu \text{ is a non-negative measure on } \mathcal{E} \text{ and } \mu \text{ is Radon.}\} \tag{1}$$

Note a measure μ is called *Radon*, if

$$\mu(K) < \infty, \quad \forall K \in \mathcal{K}(\mathbb{E}) = \text{ compacta in } \mathbb{E}.$$

Thus, compact sets have finite μ-mass.

- Subset of $M_+(\mathbb{E})$ consisting of non-negative integer valued measures; ie, point measures: Call $m(\cdot) \in M_+(\mathbb{E})$ a point measure if

$$m(\cdot) = \sum_i \epsilon_{x_i}(\cdot), \quad x_i \in \mathbb{E}$$

  where for $A \in \mathcal{E}$

$$\epsilon_{x_i}(A) = \begin{cases} 1, & \text{if } x_i \in A, \\ 0, & \text{if } x_i \notin A. \end{cases}$$

  So $m(\cdot)$ is Radon: $m(K) < \infty$, for $K \in \mathcal{K}(\mathbb{E})$. Call $\{x_i\}$ the atoms or points and m is the function which counts how many atoms fall in a set. Then

$$M_p(\mathbb{E}) \subset M_+(\mathbb{E}) = \{ \text{ all Radon point measures.} \}$$

- Test functions: Those continuous functions which vanish on complements of compact sets:

$$C_K^+(\mathbb{E}) = \{f : \mathbb{E} \mapsto \mathbb{R}_+ : f \text{ is continuous with compact support.}\}$$

2.2. Convergence and topology in $M_+(\mathbb{E})$.

- Convergence concept: Vague convergence in $M_+(\mathbb{E})$ or $M_p(\mathbb{E})$: Sppse $\mu_n \in M_+(\mathbb{E})$, $n \geq 0$. Then $\mu_n \xrightarrow{v} \mu_0$, if for all $f \in C_K^+(\mathbb{E})$ we have

$$\mu_n(f) := \int_{\mathbb{E}} f(x)\mu_n(dx) \to \mu_0(f) := \int_{\mathbb{E}} f(x)\mu_0(dx), \quad (n \to \infty).$$

- Topology specified by basis sets of the form

$$\{\mu \in M_+(\mathbb{E}) : \mu(f_i) \in (a_i, b_i), i = 1, \ldots, d\}$$

where $f_i \in C_K^+(\mathbb{E})$ and $0 \leq a_i \leq b_i$.
Unions of basis sets form the class of open sets constituting the vague topology.

- Topology is metrizable as CSMS. There exists a metric

$$d(\cdot, \cdot) = \text{ vague metric}$$

specified as follows: there exists **some** sequence of functions

$$f_i \in C_K^+(\mathbb{E})$$

and for $\mu_1, \mu_2 \in M_+(\mathbb{E})$

$$d(\mu_1, \mu_2) = \sum_{i=1}^{\infty} \frac{|\mu_1(f_i) - \mu_2(f_i)| \wedge 1}{2^i}.$$

- Interpret $\mu \in M_+(\mathbb{E})$ as an object determined by components indexed by $C_K^+(\mathbb{E})$:

$$\mu = \{\mu(f), f \in C_K^+(\mathbb{E})\}$$

or

$$\mu = \{\mu(f_i), i \geq 1\}.$$

- Leads to compactness criterion: A subset $M \subset M_+(\mathbb{E})$ vaguely relatively compact iff

$$\sup_{\mu \epsilon M} \mu(f) < \infty, \quad \forall f \in C_K^+(\mathbb{E}).$$

2.3. Random elements.

- The open subsets of $M_+(\mathbb{E})$ generate the Borel σ-field:

$$\mathcal{M}_+(\mathbb{E}) = \text{ Borel } \sigma\text{-field of subsets in } M_+(\mathbb{E}).\}$$

- Random element: So

$$\big(M_+(\mathbb{E}), \mathcal{M}_+(\mathbb{E})\big)$$

is measurable space. A random measure is measurable map

$$(\Omega, \mathcal{B}) \mapsto \big(M_+(\mathbb{E}), \mathcal{M}_+(\mathbb{E})\big)$$

and a point process is measurable map

$$(\Omega, \mathcal{B}) \mapsto \big(M_p(\mathbb{E}), \mathcal{M}_p(\mathbb{E})\big).$$

3. Weak convergence of random measures.

The following are usual and convenient methods of showing weak convergence of random elements in $M_+(\mathbb{E})$.

- In $M_+(\mathbb{E})$, random measures $\{\eta_n(\cdot), n \geq 0\}$ converge weakly

  $$\eta_n \Rightarrow \eta_0$$

  iff for any family $\{h_j\}$, with $h_j \in C_K^+(\mathbb{E})$ we have

  $$(\eta_n(h_j), j \geq 1) \Rightarrow (\eta_0(h_j), j \geq 1) \quad \text{in } \mathbb{R}^\infty.$$

  Nice:

  - One assumes a sequence $\{h_j\}$ and prove $\mathbb{R}^\infty$ convergence;
  - This reduces to proving $\mathbb{R}^d$-convergence for any fixed d.
  - Often, in fact, this reduced to one dimensional convergence.

- Method of Laplace functionals: A convenient transform technique for manipulating distributions of point processes and random measures.

Definition. Let

$$\eta : (\Omega, \mathcal{A}, \mathbb{P}) \mapsto (M_+(\mathbb{E}), \mathcal{M}_+(\mathbb{E}))$$

be a random measure. The *Laplace functional* of the random measure η is the non-negative function $\Psi_\eta : C_K^+(\mathbb{E}) \mapsto \mathbb{R}_+$

$$\begin{aligned}\Psi_\eta(f) =& \mathbf{E}\exp\{-\eta(f)\} = \int_\Omega \exp\{-\eta(\omega, f)\} d\mathbb{P}(\omega) \\ =& \int_{M_+(\mathbb{E})} \exp\{-\mu(f)\} \mathbb{P} \circ \eta^{-1}(d\mu).\end{aligned}$$

Weak convergence of a sequence of random measures in $M_+(\mathbb{E})$ equivalent to the Laplace functionals of the random measures converging for each $f \in C_K^+(\mathbb{E})$. More detail: (Resnick, 1987, Section 3.5) or Kallenberg (1983), Neveu (1977).

Theorem 1 (Convergence criterion) *Let $\{\eta_n, n \geq 0\}$ be random elements of $M_+(\mathbb{E})$. Then*

$$\eta_n \Rightarrow \eta_0 \quad \text{in } M_p(\mathbb{E}),$$

iff

$$\Psi_{\eta_n}(f) = \mathbf{E}e^{-\eta_n(f)} \to \mathbf{E}e^{-\eta_0(f)} = \Psi_{\eta_0}(f), \quad \forall f \in C_K^+(\mathbb{E}). \tag{2}$$

So weak convergence is characterized by convergence of Laplace functionals on $C_K^+(\mathbb{E})$.

$\frac{1}{2}$ **Proof.** Suppose $\eta_n \Rightarrow \eta_0$ in $M_+(\mathbb{E})$. The map $M_+(\mathbb{E}) \mapsto [0, \infty)$ defined by

$$\mu \mapsto \mu(f)$$

is continuous. The continuous mapping theorem gives

$$\eta_n(f) \Rightarrow \eta_0(f) \quad \text{in } \mathbb{R}.$$

Thus

$$e^{-\eta_n(f)} \Rightarrow e^{-\eta_0(f)},$$

and by Lebesgue's dominated convergence theorem

$$\mathbf{E}e^{-\eta_n(f)} \to \mathbf{E}e^{-\eta_0(f)}.$$

□

4. Poisson random measure

Let $N : (\Omega, \mathcal{A}) \mapsto (M_p(\mathbb{E}), \mathcal{M}_p(\mathbb{E}))$ be a point process with state space $\mathbb{E}$, where $\mathcal{M}_p(\mathbb{E})$ is the Borel σ-algebra of subsets of $M_p(\mathbb{E})$ generated by open sets.

Definition 1 *N is a Poisson process with mean measure μ or synonomously a Poisson random measure* (PRM*(μ)) if*

1. *For $A \in \mathcal{E}$*

$$P[N(A) = k] = \begin{cases} \dfrac{e^{-\mu(A)}(\mu(A))^k}{k!}, & \text{if } \mu(A) < \infty \\ 0, & \text{if } \mu(A) = \infty. \end{cases}$$

2. *If $A_1, \ldots, A_k$ are disjoint subsets of $\mathbb{E}$ in $\mathcal{E}$, then $N(A_1), \ldots, N(A_k)$ are independent random variables.*

4.1. Properties.

- **Transformation preserves Poisson-i-ness.**

Proposition 1 *Suppose*

$$T : \mathbb{E} \mapsto \mathbb{E}'$$

is a measurable mapping of one nice space $\mathbb{E}$ *into another nice space* $\mathbb{E}'$ *such that*

$$\begin{aligned} &K' \in \mathcal{K}(\mathbb{E}') \text{ is compact in } \mathbb{E}' \\ &\Rightarrow \\ &T^{-1}K' := \{e \in E : Te \in K'\} \in \mathcal{K}(\mathbb{E}). \end{aligned}$$

If N *is* $\mathrm{PRM}(\mu)$ *on* $\mathbb{E}$ *then* $N' := N \circ T^{-1}$ *is* $\mathrm{PRM}(\mu')$ *on* $\mathbb{E}'$ *where* $\mu' := \mu \circ T^{-1}$.

Proof. Poisson marginals and independence preserved. $\square$

- **Augmentation preserves Poisson-i-ness.**

Proposition 2 *Suppose* $\{X_n\}$ *are random elements of a nice space* $\mathbb{E}_1$ *such that*

$$\sum_n \epsilon_{X_n}$$

is PRM(μ). *Suppose* $\{J_n\}$ *are iid random elements of a second nice space* $\mathbb{E}_2$ *with common probability distribution* F *and suppose the Poisson process and the sequence* $\{J_n\}$ *are defined on the same probability space and are independent. Then the point process on* $\mathbb{E}_1 \times \mathbb{E}_2$

$$\sum_n \epsilon_{(X_n, J_n)}$$

is PRM *with mean measure* $\mu \times F$.

- **Laplace functional of PRM(μ).** The Poisson process can be identified by the characteristic form of its Laplace functional.

Theorem 2 (Laplace functional of PRM) *The point process N is PRM(μ) iff its Laplace functional is of the form*

$$\Psi_N(f) = \exp\{-\int_{\mathbb{E}} (1 - e^{-f(x)})\mu(dx)\}, \quad f \in C_K^+(\mathbb{E}). \tag{3}$$

Proof. Verify (3) by *usual* measure theory argument:

- Let $f = \lambda 1_A$; for $\lambda > 0$. Then

$$\Psi_N(f) = E \exp\{-\lambda N(A)\}$$

  and (3) verified by direct calculation.
- Let

$$f = \sum_{i=1}^{k} \lambda_i 1_{A_i}$$

  for $\lambda_i > 0$ and $A_1, \ldots, A_k$ a partition. Then again verify (3) by direct calculation.
- Finish with a limiting argument, approximating general f by simple $f_n \uparrow f$. $\square$

- **Construction of PRM.** Suppose N is PRM(μ) on a nice space $\mathbb{E}$, such that
$$\mu(\mathbb{E}) < \infty.$$
Represent N as Poissonized sum of masses: Define the pm F
$$F(dx) = \mu(dx)/\mu(\mathbb{E})$$
on $\mathcal{E}$. Let
  - $\{\boldsymbol{X}_n, n \geq 1\}$ iid random elements of $\mathbb{E}$, common distribution F;
  - Let $\tau \perp\!\!\!\perp \{\boldsymbol{X}_n\}$, τ has Poisson distribution with parameter $\mu(\mathbb{E})$;
  - Define
$$N = \begin{cases} \sum_{i=1}^{\tau} \epsilon_{\boldsymbol{X}_i}, & \text{if } \tau \geq 1 \\ 0, & \text{if } \tau = 0. \end{cases}$$

Then N is PRM(μ).

Proof. Compute the Laplace functional. $\square$

5. Sample measures and PRM

Suppose that for each $n \geq 1$ we have

$$\{\boldsymbol{X}_{n,j}, j \geq 1\}$$

is a sequence of iid random elements of $(\mathbb{E}, \mathcal{E})$. We call

$$\sum_{j=1}^{n} \epsilon_{\boldsymbol{X}_{n,j}}$$

an *sample measure*.

Theorem 3 (Basic convergence) *We have*

$$\sum_{j=1}^{n} \epsilon_{\boldsymbol{X}_{n,j}} \Rightarrow \sum_{k} \epsilon_{\boldsymbol{j}_k} = \mathrm{PRM}(\mu) \tag{4}$$

in $M_p(\mathbb{E})$, *iff*

$$n\mathbb{P}[\boldsymbol{X}_{n,1} \in \cdot] = \mathbf{E}\left(\sum_{j=1}^{n} \epsilon_{\boldsymbol{X}_{n,j}}(\cdot)\right) \xrightarrow{v} \mu \tag{5}$$

in $M_+(\mathbb{E})$. *Furthermore* (4) *or* (5) *is equivalent to the version with a "time" component:*

$$\sum_{j=1}^{n} \epsilon_{\left(\frac{j}{n}, \boldsymbol{X}_{n,j}\right)} \Rightarrow \sum_{k} \epsilon_{\left(t_k, \boldsymbol{j}_k\right)} = \mathrm{PRM}(\mathbb{LEB} \times \mu) \tag{6}$$

in $M_p([0, \infty) \times \mathbb{E})$.

Proof. Compute Laplace functionals of the sample measures and decide when they converge: For $f \in C_K^+(\mathbb{E})$,

$$\begin{aligned}
\mathbf{E}e^{-\sum_{j=1}^n \epsilon_{\boldsymbol{X}_{n,j}}(f)} &= \mathbf{E}e^{-\sum_{j=1}^n f(\boldsymbol{X}_{n,j})} = \left(\mathbf{E}e^{-f(\boldsymbol{X}_{n,1})}\right)^n \\
&= \left(1 - \frac{\mathbf{E}(n(1-e^{-f(\boldsymbol{X}_{n,1})}))}{n}\right)^n \\
&= \left(1 - \frac{\int_{\mathbb{E}}(1-e^{-f(x)})nP[\boldsymbol{X}_{n,1} \in dx]}{n}\right)^n
\end{aligned}$$

and this converges to

$$\exp\{\int_{\mathbb{E}}(1-e^{-f(x)})\mu(dx)\},$$

the Laplace functional of PRM(μ), iff

$$\int_{\mathbb{E}}(1-e^{-f(x)})nP[\boldsymbol{X}_{n,1} \in dx] \to \int_{\mathbb{E}}(1-e^{-f(x)})\mu(dx).$$

and this last statement is equivalent to vague convergence in (5). $\square$

Corollary 1 (Variant of basic convergence) *Suppose additionally that* $0 < a_n \uparrow \infty$. *Then for a measure* $\mu \in M_+(\mathbb{E})$ *we have*

$$\frac{1}{a_n}\sum_{j=1}^{n} \epsilon_{X_{n,j}} \Rightarrow \mu \tag{7}$$

on $M_+(\mathbb{E})$ *iff*

$$\frac{n}{a_n} P[X_{n,1} \in \cdot\,] = \mathbf{E}\left(\frac{1}{a_n}\sum_{j=1}^{n} \epsilon_{X_{n,j}}(\cdot)\right) \xrightarrow{v} \mu \tag{8}$$

in $M_+(\mathbb{E})$.

Proof. Similar to previous proof. Compute Laplace functional for the quantity on the left side of (7):

$$\mathbf{E}e^{-\frac{1}{a_n}\sum_{i=1}^{n}\epsilon_{X_{n,1}}(f)} = \left(\mathbf{E}e^{-\frac{1}{a_n}f(X_{n,1})}\right)^n$$
$$= \left(1 - \frac{\int_{\mathbb{E}}\big(1-e^{-\frac{1}{a_n}f(x)}\big)n\mathbb{P}[X_{n,1}\in dx]}{n}\right)^n$$

and this converges to $e^{-\mu(f)}$, the Laplace functional of μ, iff

$$\int_{\mathbb{E}}(1-e^{-\frac{1}{a_n}f(x)})nP[X_{n,1}\in dx] \to \mu(f). \tag{9}$$

Since $a_n \to \infty$,

$$\int_{\mathbb{E}}(1-e^{-f(x)/a_n})n\mathbb{P}[X_{n,1}\in dx] \approx \int_{\mathbb{E}} f(x)\frac{n}{a_n}P[X_{n,1}\in dx] \to \mu(f)$$

iff

$$\frac{n}{a_n}P[X_{n,1}\in\cdot\,] \xrightarrow{v} \mu.$$

$\square$

6. Weak convergence of partial sums to Lévy processes

Result for d dimensions. Set $\mathbb{E} = [-\infty, \infty] \setminus \{\mathbf{0}\}$. Denote random vectors by $\boldsymbol{X} = (X^{(1)}, \dots, X^{(d)})$.

Theorem 4 *Suppose for each $n \geq 1$, that $\{\boldsymbol{X}_{n,j}, j \geq 1\}$ are iid random vectors such that*

$$n\mathbb{P}[\boldsymbol{X}_{n,1} \in \cdot] \xrightarrow{v} \nu(\cdot) \tag{10}$$

in $M_+(\mathbb{E})$, where ν is a Lévy measure and for each $j = 1, \dots, d$,

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} n\mathbf{E}\Big(\big(X_{n,1}^{(j)}\big)^2 1_{[|X_{n,1}^{(j)}| \leq \varepsilon]}\Big) = 0. \tag{11}$$

Define the partial sum process based on the nth array row:

$$\boldsymbol{X}_n(t) := \sum_{k=1}^{[nt]} \Big(\boldsymbol{X}_{nk} - \mathbf{E}\big(\boldsymbol{X}_{n,k} 1_{[\|\boldsymbol{X}_{n,k}\| \leq 1]}\big)\Big), \quad t \geq 0.$$

Then (10) and (11) imply

$$\boldsymbol{X}_n \Rightarrow \boldsymbol{X}_0,$$

in $D([0,\infty), \mathbb{R}^d)$, where $\boldsymbol{X}_0(\cdot)$ is a Lévy jump process with Lévy measure ν.

Outline
Pt processes, RM's
Weak convergence RM's
PRM
Sample & PRM
Sums to Lévy
Title Page
Page 21 of 29
Go Back
Full Screen
Close
Quit

Proof from Resnick and Greenwood (1979), Resnick (2006). Proof in several steps:

STEP 1: Basic convergence, Theorem 3, and (10) imply

$$\sum_{k=1}^{\infty} \epsilon_{(\frac{k}{n}, \boldsymbol{X}_{n,k})} \Rightarrow \sum_{k} \epsilon_{(t_k, \boldsymbol{j}_k)} = \mathrm{PRM}(\mathbb{LEB} \times \nu) \tag{12}$$

in $M_p([0,\infty) \times \mathbb{E})$.

STEP 2: Two continuity assertions:

(i) With respect to the distribution of PRM $(\mathbb{LEB} \times \nu)$, the restriction map $M_p([0,\infty) \times \mathbb{E}) \mapsto M_p([0,\infty) \times \{\mathbf{x} : \|\mathbf{x}\| > \varepsilon\})$ defined by

$$m \mapsto m|_{[0,\infty)\times\{\mathbf{x}:\|\mathbf{x}\|>\varepsilon\}}$$

is almost surely continuous.

(ii) On $M_p([0,\infty] \times \{\mathbf{x} : \|\mathbf{x}\| > \varepsilon\})$ the summation functional from $M_p([0,\infty) \times \{\mathbf{x} : \|\mathbf{x}\| > \varepsilon\}) \mapsto D([0,T], \mathbb{R}^d)$ defined by

$$\sum_{k} \epsilon_{(\tau_k, \boldsymbol{J}_k)} \to \sum_{\tau_k \le (\cdot)} \boldsymbol{J}_k$$

is almost surely continuous with respect to the distribution of $\mathrm{PRM}(\mathbb{LEB} \times \nu)$.

STEP 3: From first continuity assertion in Step 2, the convergence statement in Step 1, and continuous mapping Theorem we get the restricted convergence

$$\sum_k 1_{[\,\|\boldsymbol{X}_{n,k}\|>\varepsilon]}\epsilon_{(\frac{k}{n},\boldsymbol{X}_{n,k})} \Rightarrow \sum_k 1_{[\,\|\boldsymbol{j}_k\|>\varepsilon]}\epsilon_{(t_k,\boldsymbol{j}_k)} \tag{13}$$

in $M_p([0,\infty)\times\{\mathbf{x}:\|\mathbf{x}\|>\varepsilon\})$. From the second continuity assertion in Step 2, we get from (13)

$$\sum_{k=1}^{[n\cdot]} \boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|>\varepsilon]} \Rightarrow \sum_{t_k\leq(\cdot)} \boldsymbol{j}_k 1_{[\,\|\boldsymbol{j}_k\|>\varepsilon]} \tag{14}$$

in $D([0,T],\mathbb{R}^d)$. Similarly we get

$$\sum_{k=1}^{[n\cdot]} \boldsymbol{X}_{n,k}1_{[\varepsilon<\|\boldsymbol{X}_{n,k}\|\leq 1]} \Rightarrow \sum_{t_k\leq(\cdot)} \boldsymbol{j}_k 1_{[\varepsilon<\|\boldsymbol{j}_k\|\leq 1]}. \tag{15}$$

STEP 4. In (15), take expectations and apply (10) to get

$$[n\cdot]\mathbf{E}\big(\boldsymbol{X}_{n,1}1_{[\varepsilon<\|\boldsymbol{X}_{n,1}\|\leq 1]}\big) \to (\cdot)\int_{\{\mathbf{x}:\varepsilon<\|\mathbf{x}\|\leq 1\}} \mathbf{x}\nu(d\mathbf{x}) \tag{16}$$

in $D([0,T],\mathbb{R}^d)$. To justify this, observe first for any $t>0$ that

$$[nt]\mathbf{E}\big(\boldsymbol{X}_{n,1}1_{[\varepsilon<\|\boldsymbol{X}_{n,1}\|\le 1]}\big) = \frac{[nt]}{n}\int_{\{\mathbf{x}:\|\mathbf{x}\|\in(\varepsilon,1]\}} \mathbf{x} n\mathbb{P}[\boldsymbol{X}_{n,1}\in d\mathbf{x}]$$
$$\to t\int_{\{\mathbf{x}:\|\mathbf{x}\|\in(\varepsilon,1]\}} \mathbf{x}\nu(d\mathbf{x})$$

since $n\mathbb{P}[\boldsymbol{X}_{n,1}\in\cdot]\xrightarrow{v}\nu(\cdot)$. Convergence is locally uniform in t and hence convergence takes place in $D([0,T],\mathbb{R}^d)$.

STEP 5. Difference (14)–(16). The result is

$$\boldsymbol{X}_n^{(\varepsilon)}(\cdot) = \sum_{k=1}^{[n\cdot]} \boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|>\varepsilon]} - [n\cdot]\mathbf{E}\big(\boldsymbol{X}_{n,1}1_{[\varepsilon<\|\boldsymbol{X}_{n,1}\|\le 1]}\big)$$
$$\Rightarrow \boldsymbol{X}_0^{(\varepsilon)}(\cdot) := \sum_{t_k\le\cdot} \boldsymbol{j}_k 1_{[\,\|\boldsymbol{j}_k\|>\varepsilon]} - (\cdot)\int_{\left\{\mathbf{x}:\,\|\mathbf{x}\|\in(\varepsilon,1]\right\}} \mathbf{x}\nu(d\mathbf{x}). \tag{17}$$

From the Itô representation of a Lévy process, for almost all ω, as $\varepsilon\downarrow 0$,

$$\boldsymbol{X}_0^{(\varepsilon)}(\cdot)\to\boldsymbol{X}_0(\cdot),$$

locally uniformly in t. Let $d(\cdot,\cdot)$ be the Skorohod metric on $D[0,\infty)$. Local uniform convergence $\Rightarrow$ implies Skorohod convergence, so

$$d\left(\boldsymbol{X}_0^{(\varepsilon)}(\cdot),\boldsymbol{X}_0(\cdot)\right)\to 0$$

almost surely as $\varepsilon \downarrow 0$. Almost sure convergence $\Rightarrow$ weak convergence, so

$$\boldsymbol{X}_0^{(\varepsilon)}(\cdot) \Rightarrow \boldsymbol{X}_0(\cdot).$$

in $D([0,\infty),\mathbb{R}^d)$.

STEP 6. By Converging Together Theorem suffices to show

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\mathbb{P}[d(\boldsymbol{X}_n^{(\varepsilon)},\boldsymbol{X}_n)>\delta]=0.$$

Convergence in $D([0,\infty),\mathbb{R}^d)$, if Skorohod convergence in $D([0,T],\mathbb{R}^d)$ for any T. Skorohod metric on $D([0,T],\mathbb{R}^d)$ bounded above by the uniform metric on $D([0,T],\mathbb{R}^d)$. Suffices to show

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\mathbb{P}[\sup_{0\le t\le T}\|\boldsymbol{X}_n^{(\varepsilon)}(t)-\boldsymbol{X}_n(t)\|>\delta]=0,$$

for any $\delta>0$. Recalling definitions of $\boldsymbol{X}_n^{(\epsilon)}$ and $\boldsymbol{X}_n$:

$$\begin{aligned}\|\boldsymbol{X}_n^{(\varepsilon)}(t)-\boldsymbol{X}_n(t)\| &=\Big\|\sum_{k=1}^{[nt]}\boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}-[nt]\mathbf{E}\big(\boldsymbol{X}_{n,1}1_{[\,\|\boldsymbol{X}_{n,1}\|\le\varepsilon]}\big)\Big\|\\ &=\Big\|\sum_{k=1}^{[nt]}\Big(\boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}-\mathbf{E}\big(\boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}\big)\Big)\Big\|\end{aligned}$$

and so

$$\mathbb{P}[\sup_{0\le t\le T}\|\boldsymbol{X}_n^{(\varepsilon)}(t)-\boldsymbol{X}_n(t)\|>\delta]$$

$$\le\mathbb{P}\Big[\sup_{0\le t\le T}\Big\|\sum_{k=1}^{[nt]}\Big(\boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}-\mathbf{E}\big(\boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}\big)\Big)\Big\|>\delta\Big]$$

$$=\mathbb{P}\Big[\sup_{0\le j\le nT}\Big\|\sum_{k=1}^{j}\Big(\boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}-\mathbf{E}\big(\boldsymbol{X}_{n,k}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}\big)\Big)\Big\|>\delta\Big].$$

Now use the fact that $\|\mathbf{x}\|\le d\bigvee_{i=1}^{d}|x^{(i)}|$ and we get the bound

$$\le\sum_{i=1}^{d}\mathbb{P}\Big[\sup_{0\le j\le nT}\Big|\sum_{k=1}^{j}\Big(X_{n,k}^{(i)}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}-\mathbf{E}\big(X_{n,k}^{(i)}1_{[\,\|\boldsymbol{X}_{n,k}\|\le\varepsilon]}\big)\Big)\Big|>\frac{\delta}{d}\Big]$$

and by Kolmogorov's inequality this has upper bound

$$
\begin{aligned}
&\leq (\delta/d)^{-2} \sum_{i=1}^{d} \operatorname{Var}\Big(\sum_{k=1}^{[nT]} X_{n,k}^{(i)} 1_{[\,\|\boldsymbol{X}_{n,k}\| \leq \varepsilon]}\Big) \\
&= (\delta/d)^{-2} \sum_{i=1}^{d} [nT] \operatorname{Var}\Big(X_{n,1}^{(i)} 1_{[\,\|\boldsymbol{X}_{n,1}\| \leq \varepsilon]}\Big) \\
&\leq (\delta/d)^{-2} \sum_{i=1}^{d} [nT] \mathbf{E}\Big(\big(X_{n,1}^{(i)}\big)^2 1_{[\,|X_{n,1}^{(i)}| \leq \varepsilon]}\Big).
\end{aligned}
$$

Taking $\lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty}$, we easily get 0 by (11). $\square$

Contents

References

O. Kallenberg. *Random Measures.* Akademie-Verlag, Berlin, third edition, 1983. ISBN 0-12-394960-2.

J. Neveu. Processus ponctuels. In *École d'Été de Probabilités de Saint-Flour, VI—1976*, pages 249–445. Lecture Notes in Math., Vol. 598, Berlin, 1977. Springer-Verlag.

S.I. Resnick and P. Greenwood. A bivariate stable characterization and domains of attraction. *J. Multivariate Anal.*, 9(2):206–221, 1979. ISSN 0047-259X.

S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2006. ISBN: 0-387-24272-4.

S.I. Resnick. *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, New York, 1987.

Applied Probability Modeling of Data Networks;

2. Multivariate regular variation; the Poisson transform; stable processes.

Sidney Resnick
School of Operations Research and Industrial Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA
http://legacy.orie.cornell.edu/~sid
sir1@cornell.edu

August 5, 2007

1. Multivariate regular variation of functions

Basic when studying models built from iid objects.

1.1. Regular variation of functions.

- $\mathbb{C} \subset \mathbb{R}^d$ is a *cone*; ie,

$$\mathbf{x} \in \mathbb{C} \Rightarrow t\mathbf{x} \in \mathbb{C}, \quad \forall\, t > 0.$$

  Examples:

$$\begin{aligned} \mathbb{C} =& (0,\infty], \quad (d=1), \quad [0,\infty]^d \setminus \{\mathbf{0}\}, \qquad (0,\infty]^d \\ =& [-\infty,\infty]^d \setminus \{\mathbf{0}\} \quad [-\infty,\infty]^d \setminus \{-\boldsymbol{\infty}\} \quad [0,\infty] \times (0,\infty] \quad (d=2). \end{aligned}$$

- A function $h : \mathbb{C} \mapsto (0,\infty)$ is *monotone* if it is either
  - non-decreasing
  - non-increasing

  in each component.

  h non-decreasing, $\Leftrightarrow \mathbf{x}, \boldsymbol{y} \in \mathbb{C}$ and $\mathbf{x} \le \boldsymbol{y} \Rightarrow h(\mathbf{x}) \le h(\boldsymbol{y})$.

Definition 1 *Sppse $h \geq 0$ is measurable function defined on $\mathbb{C}$. Sppse $\mathbf{1} = (1, \dots, 1) \in \mathbb{C}$. Call h multivariate regularly varying with limit function λ provided:*

- $\lambda(\mathbf{x}) > 0$ *for* $\mathbf{x} \in \mathbb{C}$*; and*
- *for all* $\mathbf{x} \in \mathbb{C}$ *we have*

$$\lim_{t \to \infty} \frac{h(t\mathbf{x})}{h(t\mathbf{1})} = \lambda(\mathbf{x}). \tag{1}$$

Properties

Scaling arguments give the following properties:

1. For $d = 1$, $\mathbb{C} = (0, \infty)$ this says

$$\lim_{t \to \infty} \frac{h(tx)}{h(t)} = \lambda(x), \quad x > 0,$$

and a scaling argument shows

$$\lambda(x) = x^{\rho}, \quad \text{for some } \rho \in \mathbb{R},$$

since

$$\lambda(xy) = \lambda(x)\lambda(y), \quad x > 0, y > 0.$$

2. For $d > 1$, fix $\mathbf{x} \in \mathbb{C}$ and

$$U(t) = h(t\mathbf{x})$$

is one dimensional regularly varying so For any $s > 0$

$$\lim_{t\to\infty} \frac{U(ts)}{U(t)} = \lim_{t\to\infty} \frac{h(ts\mathbf{x})}{h(t\mathbf{x})} = \lim_{t\to\infty} \frac{h(ts\mathbf{x})}{h(t\mathbf{1})} \Big/ \frac{h(t\mathbf{x})}{h(t\mathbf{1})} = \frac{\lambda(s\mathbf{x})}{\lambda(\mathbf{x})},$$

and conclude for some $\rho(\mathbf{x}) \in \mathbb{R}$ that $U \in RV_{\rho(\mathbf{x})}$ and

$$\frac{\lambda(s\mathbf{x})}{\lambda(\mathbf{x})} = s^{\rho(\mathbf{x})}.$$

3. Can check that $\rho(\mathbf{x})$ does not depend on $\mathbf{x}$ so limit function satisfies homogeneity

$$\lambda(s\mathbf{x}) = s^{\rho}\lambda(\mathbf{x}).$$

2. Regular variation and measures

2.1. $d = 1$, $\mathbb{C} = (0, \infty]$.

Suppose $Z \geq 0$ is a random variable with distribution function F. Regular variation of $1 - F$ has the following equivalences:

(i) $\bar{F} \in RV_{-\alpha}$, $\alpha > 0$.

(ii) There exists a sequence $\{b_n\}$, with $b_n \to \infty$ such that

$$\lim_{n\to\infty} n\bar{F}(b_n x) = x^{-\alpha}, \quad x > 0.$$

(iii) There exists a sequence $\{b_n\}$ with $b_n \to \infty$ such that

$$\mu_n(\cdot) := n\mathbb{P}\left[\frac{Z}{b_n} \in \cdot\right] \xrightarrow{v} \nu_\alpha(\cdot) \tag{2}$$

in $M_+(0, \infty]$, where $\nu_\alpha(x, \infty] = x^{-\alpha}$.

To generalize (i) or (ii) to higher dimensions can be done but dealing with distribution functions in higher dimensions is awkward. However, (iii) generalizes nicely.

2.2. $d > 1$, $\mathbb{C} = [\mathbf{0}, \boldsymbol{\infty}] \setminus \{\mathbf{0}\}$.

Equivalences for regular variation of probability measures on $\mathbb{C} = [\mathbf{0}, \boldsymbol{\infty}] \setminus \{\mathbf{0}\}$. In each, we understand the phrase *Radon measure* to mean a Radon measure which is not identically zero and which is not degenerate at a point.

Theorem 1 *Repeated use of the symbols ν, $b(\cdot)$, $\{b_n\}$ from statement to statement, does not require these objects to be exactly the same in different statements.*

1. *There exists a Radon measure ν on $\mathbb{C}$ such that*

$$\lim_{t\to\infty} \frac{1 - F(t\mathbf{x})}{1 - F(t\mathbf{1})} = \lim_{t\to\infty} \frac{\mathbb{P}\left[\frac{\boldsymbol{Z}}{t} \in [\mathbf{0}, \mathbf{x}]^c\right]}{\mathbb{P}\left[\frac{\boldsymbol{Z}}{t} \in [\mathbf{0}, \mathbf{1}]^c\right]} = \nu\big([\mathbf{0}, \mathbf{x}]^c\big), \qquad (3)$$

   for all points $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$ which are continuity points of the function $\nu([\mathbf{0}, \cdot]^c)$.

2. *There exists a function $b(t) \to \infty$ and a Radon measure ν on $\mathbb{C}$ called the limit measure, such that in $M_+(\mathbb{C})$*

$$t\mathbb{P}\left[\frac{\boldsymbol{Z}}{b(t)} \in \cdot\right] \xrightarrow{v} \nu, \quad t \to \infty. \qquad (4)$$

3. *There exists a sequence* $b_n \to \infty$ *and a Radon measure* ν *on* $\mathbb{C}$ *such that in* $M_+(\mathbb{C})$

$$n\mathbb{P}[\frac{\boldsymbol{Z}}{b_n} \in \cdot] \xrightarrow{v} \nu, \quad n \to \infty. \tag{5}$$

4. *Polar coordinate version: Define*

$$(R, \boldsymbol{\Theta}) = (\|\boldsymbol{Z}\|, \frac{\boldsymbol{Z}}{\|\boldsymbol{Z}\|}$$

and

$$\aleph_+ = \{\mathbf{x} \in \mathbb{C} : \|\mathbf{x}\| = 1\}.$$

There exists a probability measure $S(\cdot)$ *on* $\aleph_+$ *called the angular measure, and a function* $b(t) \to \infty$ *such that for* $(R, \boldsymbol{\Theta}) = \left(\|\boldsymbol{Z}\|, \frac{\boldsymbol{Z}}{\|\boldsymbol{Z}\|}\right)$ *we have*

$$t\mathbb{P}[(\frac{R}{b(t)}, \boldsymbol{\Theta}) \in \cdot] \xrightarrow{v} c\nu_\alpha \times S \tag{6}$$

in $M_+\big((0,\infty] \times \aleph_+\big)$ *, for some* $c > 0$.

5. *There exists a probability measure $S(\cdot)$ on $\aleph_+$ and a sequence $b_n \to \infty$ such that for $(R, \boldsymbol{\Theta}) = \left(\|\boldsymbol{Z}\|, \frac{\boldsymbol{Z}}{\|\boldsymbol{Z}\|}\right)$ we have*

$$n\mathbb{P}[(\frac{R}{b_n}, \boldsymbol{\Theta}) \in \cdot] \xrightarrow{v} c\nu_\alpha \times S \tag{7}$$

in $M_+\big((0,\infty] \times \aleph_+\big)$, for some $c > 0$.

Remark 1

- *Regular variation for functions: limit function scales; i.e., it is homogeneous.*
- *Regular variation for measures: the homogeneity property in Cartesian coordinates translates into a product measure in polar coordinates.*

3. Regular variation and the Poisson random measure

Multivariate regular variation additionally equivalent to induced sample measures weakly converging to Poisson random measure limits.

Theorem 2 *Suppose* $\{\boldsymbol{Z}, \boldsymbol{Z}_1, \boldsymbol{Z}_2, \dots\}$ *are iid; after transformation with polar coordinates the sequence is* $\{(R, \boldsymbol{\Theta}), (R_1, \boldsymbol{\Theta}_1), (R_2, \boldsymbol{\Theta}_2), \dots\}$. *Any of the equivalences in Theorem 1, are also equivalent to*

6. *There exists* $b_n \to \infty$ *such that*

$$\sum_{i=1}^{n} \epsilon_{\boldsymbol{Z}_i/b_n} \Rightarrow \mathrm{PRM}(\nu), \tag{8}$$

in $M_p(\mathbb{E})$.

7. *There exists a sequence* $b_n \to \infty$ *such that*

$$\sum_{i=1}^{n} \epsilon_{(R_i/b_n, \Theta_i)} \Rightarrow \mathrm{PRM}(c\nu_\alpha \times S) \tag{9}$$

in $M_p((0,\infty] \times \aleph_+)$.

These conditions imply that for any sequence $k = k(n) \to \infty$ *such that* $n/k \to \infty$ *we have*

8. *In* $M_+(\mathbb{E})$

$$\frac{1}{k}\sum_{i=1}^{n}\epsilon_{\boldsymbol{Z}_i/b(\frac{n}{k})} \Rightarrow \nu \tag{10}$$

and

9. *In* $M_+\big((0,\infty]\times\aleph_+\big)$

$$\frac{1}{k}\sum_{i=1}^{n}\epsilon_{(R_i/b(\frac{n}{k}),\boldsymbol{\Theta}_i)} \Rightarrow c\nu_\alpha \times S \tag{11}$$

and 8 or 9 is equivalent to any of 1–7, provided $k(\cdot)$ *satisfies* $k(n) \sim k(n+1)$.

3.1. Variant with time coordinate

The result with a time coordinate needed for proving weak convergence of partial sum processes or maximal processes in the space $D[0,\infty)$.

Theorem 3 *Suppose* $\{\boldsymbol{Z}, \boldsymbol{Z}_1, \boldsymbol{Z}_2, \dots\}$ *are iid random elements of* $[\mathbf{0}, \boldsymbol{\infty})$. *Then multivariate regular variation of the distribution of* $\boldsymbol{Z}$ *in* $\mathbb{C} = [\mathbf{0}, \boldsymbol{\infty}] \setminus \{\mathbf{0}\}$

$$n\mathbb{P}[\frac{\boldsymbol{Z}}{b_n} \in \cdot] \xrightarrow{v} \nu$$

is also equivalent to

$$\sum_j \epsilon_{(\frac{j}{n}, \boldsymbol{Z}_j/b_n)} \Rightarrow \mathrm{PRM}(\mathbb{LEB} \times \nu) \tag{12}$$

in $M_+([0,\infty) \times \mathbb{C})$.

4. Weak convergence of partial sums to stable processes

We can specialize the result for convergence of sums to Lévy processes. Assume, for simplicity, $d = 1$ and

$$\mathbb{C} = [-\infty, \infty] \setminus \{0\}.$$

Regular variation of the tail probabilities on the cone $\mathbb{R} \setminus \{0\}$,

$$t\mathbb{P}[\frac{Z_1}{b(t)} \in \cdot\,] \xrightarrow{v} \nu(\cdot),$$

is equivalent to

$$\lim_{x\to\infty} \frac{\mathbb{P}[Z_1 > x]}{\mathbb{P}[|Z_1| > x]} = px^{-\alpha}, \quad \lim_{x\to\infty} \frac{\mathbb{P}[Z_1 \le -x]}{\mathbb{P}[|Z_1| > x]} = qx^{-\alpha},$$

and

$$\mathbb{P}[|Z_1| > x] \in RV_{-\alpha}.$$

Corollary 1 *Consider the special case where* $\{Z_n, n \geq 1\}$ *are iid random variables on* $\mathbb{R}$ *and set* $X_{n,j} = Z_j/b_n$ *for some* $b_n \to \infty$. *Define* ν *for* $x > 0$ *and* $0 < \alpha < 2$ *by*

$$\nu((x,\infty]) = px^{-\alpha}, \quad \nu((-\infty,-x]) = qx^{-\alpha}, \tag{13}$$

where $0 \leq p \leq 1$ *and* $q = 1 - p$. *Then*

$$\sum_{j=1}^{[n\cdot]} \frac{Z_j}{b(n)} - [n\cdot]\mathbf{E}\Big(\frac{Z_1}{b(n)} 1_{[|\frac{Z_1}{b(n)}|\leq 1]}\Big) \Rightarrow X_\alpha(\cdot), \tag{14}$$

in $D[0,\infty)$, *where the limit is* α*-stable Lévy motion with Lévy measure* ν, *iff*

$$n\mathbb{P}[\frac{Z_1}{b_n} \in \cdot] \xrightarrow{v} \nu, \tag{15}$$

in $M_+(\mathbb{C})$.

Proof sufficiency. Given regular variation, plug into the result about convergence of sums to a Lévy process. We only need to check the *truncated 2nd moment condition*

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty} n\mathbf{E}\Big(\big(X_{n,1}\big)^2 1_{[\,|X_{n,1}|\le\varepsilon]}\Big) = 0,$$

where

$$X_{n,1} = \frac{Z_1}{b(n)}.$$

We have by Karamata's theorem,

$$n\mathbf{E}\left(\Big(\frac{Z_1}{b(n)}\Big)^2 1_{\left[\,|\frac{Z_1}{b(n)}|\le\varepsilon\right]}\right) \to \int_{[\,|x|\le\varepsilon\,]} x^2\nu(dx), \quad (n\to\infty)$$
$$= \frac{p\alpha\varepsilon^{2-\alpha}}{2-\alpha} + \frac{q\alpha\varepsilon^{2-\alpha}}{2-\alpha} = (\text{ const })\varepsilon^{2-\alpha}$$

and as $\varepsilon\to 0$, we have $\varepsilon^{2-\alpha}\to 0$ as required for the partial sum process to converge to the Lévy process.

Contents

References

O. Kallenberg. *Random Measures.* Akademie-Verlag, Berlin, third edition, 1983. ISBN 0-12-394960-2.

J. Neveu. Processus ponctuels. In *École d'Été de Probabilités de Saint-Flour, VI—1976*, pages 249–445. Lecture Notes in Math., Vol. 598, Berlin, 1977. Springer-Verlag.

S.I. Resnick and P. Greenwood. A bivariate stable characterization and domains of attraction. *J. Multivariate Anal.*, 9(2):206–221, 1979. ISSN 0047-259X.

S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2006. ISBN: 0-387-24272-4.

S.I. Resnick. *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, New York, 1987.

Applied Probability Modeling of Data Networks;

3. Intro & survey of data network modeling.

Sidney Resnick
School of Operations Research and Industrial Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA
http://legacy.orie.cornell.edu/∼sid
sir1@cornell.edu

August 5, 2007

1. Introduction

Does data network traffic behave statistically
like telephone network traffic?

Action:

- Stop assuming the two types of networks behave the same.
- Start checking.

Initially at Bellcore (now Telcordia) and later at AT&T Labs-Research, high resolution measurements (data) were collected. Usually the data consisted of counts of bits, bytes, packets etc per unit time (eg millisecond). This could then be aggregated to coarser time scales. For example

- ...
- 10 seconds
- 10 milliseconds

- ...
- 1 second
- 1 millisecond

Significant examples:

- LAN's and WAN's
  - Willinger et al. (1997)
  - Duffy et al. (1993)
  - Leland et al. (1993)
  - Willinger et al. (1995)
- WWW traffic
  - Crovella and Bestavros (1996),
  - Crovella and Bestavros (1997).

Measurements on data networks exhibit features surprising by the standards of classical queueing and telephone network models. These are called

- *invariants*

which is to networks what

- *stylized facts*

are to finance.

2. Stylized facts

1. Heavy tails abound for such things as

   - file sizes,
   - transmission rates,
   - durations (file transfers, connection lengths).

   (See Arlitt and Williamson (1996), Leland et al. (1994), Maulik et al. (2002), Resnick and Rootzén (2000), Resnick (2003), Willinger and Paxson (1998), Willinger et al. (1998), Willinger (1998).)

   Note: a random variable X has a heavy tail if

   $$P[X > x] \approx x^{-\alpha}, \quad \alpha > 0, \ x \text{ large}.$$

   - Tail exhibits power law decay.
   - Limited moments:

   $$E(|X|^{\alpha+\delta}) = \infty, \quad \delta > 0.$$

   - If $1 < \alpha < 2$, no variance!

2. The number of bits or packets per slot exhibits *long range dependence* across time slots (eg, Leland et al. (1993), Willinger et al. (1995). There is also a perception of *self-similarity* as the width of the time slot varies across a range of time scales exceeding a typical round trip time.

   Note: A stationary process $\{X_n\}$ possesses long range dependence if dependence between variables decays slowly as the gap between the variables increases:

$$|\mathrm{Corr}(X_0, X_h)| \leq (const)h^{-\beta}, \quad 0 < \beta < 1.$$

3. Network traffic is *bursty* with rare but influential periods of very high transmission rates punctuating typical periods of modest activity.
   - Bursty is a somewhat ill defined concept associated with heavy tailed transmissions rates.
   - Introduces peak loads to the network.
   - Associated with large files transmitted over fast links.
   - Not associated with the truism: *Traffic at a heavily loaded link is normally distributed.*

3. Broad Issues (BI's)

BI-1 Role of statistics and applied probability:

- Statistics: Empirically identify phenomena and properties of the data so as to better understand what network data *in the wild* should look like. NOT prediction typically.
  - Examples: Identify presence of heavy tails, long range dependence, self-similarity;
  - Understand different statistical properties of various applications and protocols (ftp, http, mail, streaming audio).
- Applied probability: Build models which explain relations and explain empirically observed phenomena.
  - Example: Sizes of files stored on a server follows Pareto power law tail which causes long range dependence.

    Pardigm: Heavy tails cause long range dependence.

  - Build models of end user behavior which allow construction of simulation tools to study effect of tweaking protocols.

BI-2 Problem of time scales: Can Applied Math, Applied Prob & Statistics make contributions to data network analysis and planning in *Internet time*.

- Developers have short attention spans and little patience with outsider's toys.
- Two year time horizon to write a PhD thesis and really understand something is ridiculously long time horizon.
- Pessimists view: the best the mathy community can hope for is to cause paradigm shifts with explanations which may lag behind developments.
- Take the money and run mentality. ("*Anyone who gets a PhD does not understand economics*.") Long development time for a project means other people are earning.

4. An approach to modeling: Infinite source Poisson model–A fluid model

Suppose *sessions* characterized by

- Session initiation times are $\{\Gamma_k\}$ where

$$\{\Gamma_k\} \sim \text{ homogeneous Poisson on } (-\infty, \infty), \text{ rate } \lambda.$$

- Sequence of iid marks independent of $\{\Gamma_k\}$: Each Poisson point Γ_k receives a *mark* which characterizes input characteristics:

$$(F_k, L_k, R_k) = (\text{file, duration, rate }),$$

where

$$F_k = L_k R_k.$$

All three quantities are seen empirically to be marginally heavy tailed:

$$\begin{aligned} P[F > x] &\sim x^{-\alpha_F} \\ P[L > x] &\sim x^{-\alpha_L} \\ P[R > x] &\sim x^{-\alpha_R}, \end{aligned}$$

with (usually) $1 < \alpha_F, \alpha_R, \alpha_L < 2$.

Examples of mark (F_k, L_k, R_k) structures:

– Early simple models assumed constant input rates,

$$R_k = 1$$

so that total input rate at time t is

$$M(t) = \ \#\text{ active sources at time } t.$$

– Random but constant rates R_k during the transmission interval. (Can even make rate time varying.)

– What is the distribution of the triple (F_k, L_k, R_k)?

 * Can depend on application protocol.
 * Can depend on how data is segmented and how sessions defined from connection level data of packet headers.
 * Different distributional assumptions lead to radically different model predictions.

- Fluid input models for cumulative input in time $(a, b]$:

$$A(a, b] = \text{ work inputted in time interval}(a, b].$$

  – Process approximation to cumulative input.

    * Large time scale approximation

$$d - \lim_{T\to\infty} \frac{A(0, Tt] - b(T)}{a(T)} = X(t),$$

      where possible limits include (Mikosch et al., 2002)

      · fractional Brownian motion (Gauss marginals, lrd)

      · stable Lévy motion (heavy tailed marginals, sii, ss)

      · BUT: not easy to find agreement with these approximate models and data (Guerin et al., 2003)

    * Small time scale approximations: Block time into small time slots $(k\delta, (k+1)\delta]$ and consider as $\delta \to 0$

$$\{A(k\delta, (k+1)\delta], k \in \mathbb{N}\}.$$

      depending on the interaction of input rates and tails. Will need $\lambda = \lambda(\delta) \uparrow \infty$ (a la heavy traffic limit theorems).

- Small time scale approximation results dependent on distributional assumptions on (F, L, R). Either

  D-limit is approximately highly correlated
  normal random variables or
  D-limit is approximately highly dependent
  stable infinite variance random variables.

- Compute dependence measure across different slots. Models predict lrd for $\{A(k\delta, (k+1)\delta], k \in \mathbb{N}\}$ (D'Auria and Resnick, 2006a,b).

5. Summary: Stylized facts and small time scale approximation

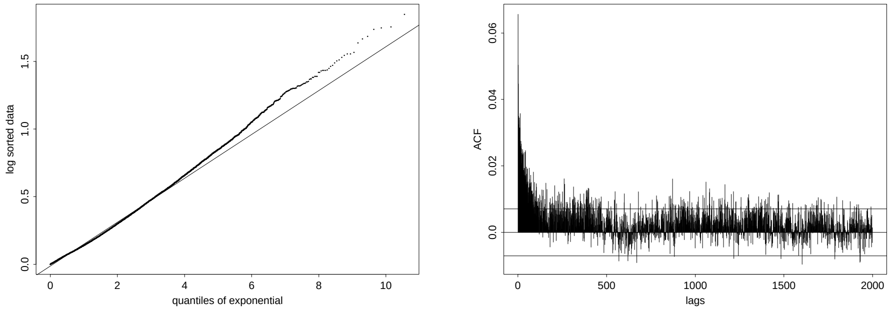
Stylized Facts	$F \perp\!\!\!\perp R$ Model	$L \perp\!\!\!\perp R$ Model
1. Heavy tails	Built in	Built in
2. $P[A(0,\delta] > x] \sim$	$x^{-(\alpha_R+\alpha_F)}$, fixed δ; $x \to \infty$	$x^{-\alpha_R}$, fixed δ; $x \to \infty$
3. LRD across slots	$\mathrm{Cov}(k) \sim c\bar{G}^{(0)}(k)$; fixed δ; $k \to \infty$	$\mathrm{EDM}(k) \sim c\bar{F}_L^{(0)}(k)$; fixed δ; $k \to \infty$
4. Cum traffic/slot is $N(0,1)$?	$\dfrac{(A(0,\delta] - (\mathrm{ctering}(\delta))}{a(\delta)}$ $\overset{d}{\approx} N(0,1)$	$\dfrac{(A(0,\delta] - (\mathrm{ctering}(\delta))}{b(\delta)}$ $\overset{d}{\approx} X_{\alpha_R}(\cdot)$

6. Some Technical Points

Tech Pt 1: Identify in the connection level data **Poisson time points** from packet headers and validate the choice.

- Quick & dirty (Q&D) solution: Check if interpoint distances are iid (sample acf almost 0) and exponentially distributed (qq-plots).
- Q&D Rules of Thumb:
  - Behavior of lots of humans acting independently is often well modelled by a Poisson process.
  - Starting times of machine triggered downloads cannot be modelled as Poisson process.

Example: UCB: Inter-arrival times of requests in http sessions via modem.



UCB data; http sessions via modem. *Left*: qqplot against exponential distribution. *Right*: autocorrelation function of interarrival times.

Tech Pt 2: **Heavy tails**: A rv X has a heavy (right) tail if

$$P[X > x] \sim x^{-\alpha}, \quad x \to \infty.$$

Notes

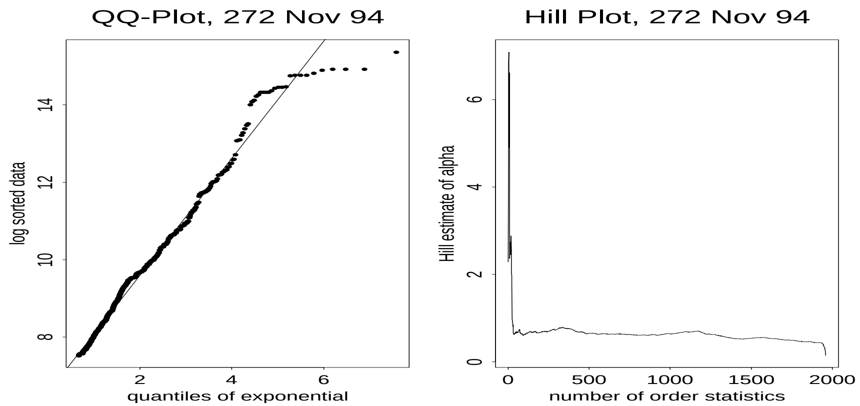
- $0 < \alpha < 1$: Very heavy: mean & variance infinite.
- $1 < \alpha < 2$: Heavy: Frequent case where mean finite but variance is infinite.
- $\alpha > 2$: Heavy with finite variance: Typical of financial data.
- For large x,

  $$P[\log X > x] \sim e^{-\alpha x}, \quad x \to \infty.$$

  So inference can be based on exponential density and thresholding techniques to account for the distribution following this law only for large x.
- For many purposes, do not need to know the whole distribution but just the tail.

Example: BU data: Influential study from mid '90's: File sizes downloaded in a web session.

Figure: Sizes of www downloads; BU experiment: QQ-plot and Hill plot.



Tech Pt 3. Checking for **independence**.

- Q&D method 1: Standard time series method checks if sample correlation function

$$\hat{\rho}(h) = \frac{\sum_{i=1}^{n-h}(X_i - \bar{X})(X_{i+h} - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}, \quad h = 1, 2, \dots,$$

  is **close** to identically 0.

  How to put meaning to phrase **close** to 0? If

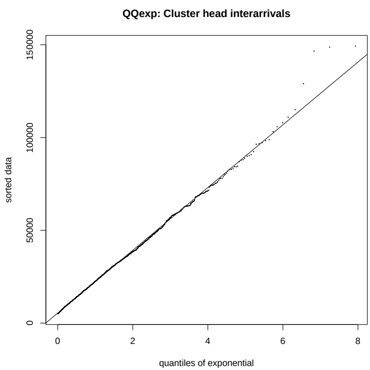
  – If finite variances, Bartlett's formula provides asymptotic normal theory.

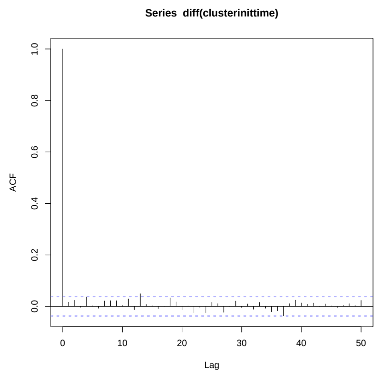
  – If heavy tailed, Davis and Resnick formula provides asymptotic distributions for $\hat{\rho}(h)$.

- Q&D method 2: If data heavy tailed, take a function of the data (say the log) to get lighter tail and test. (But this may obscure the importance of large values.)
- Q&D method 3: Subset method. Split data into (say) 2 subsets. Plot acf of each half separately. If iid, pics should look same.

Example: UNC connection data. Data contains 173604 connection vectors ordered by the connection initiation times. Clusters are obtained by considering only the connection start times ordered as they appeared in time on the UNC link and arbitrarily using a time threshold of 5 milliseconds to separate clusters. This yields 16417 clusters.

Do the cluster heads look Poisson?





CORNELL
Introduction
Stylized facts
Broad Issues (BI's)
Approach to Modeling
Model predictions
Tech points
Heavy tails ⇒ LRD
References
Title Page
Page 18 of 30
Go Back
Full Screen
Close
Quit

Tech Pt 4. Is the data **stationary**? Usually not and there are, for example, diurnal cycles. (There is only so much Red Bull (Jolt, Coke, ...) that a human can consume.)

- Q&D Coping Method: Take a slab of the copius data which looks stationary.
- Usually there is too much data. (!!??)
- Rule of thumb: don't take more than 4 hours. Depending on the data, it can be a couple of minutes; eg connection data.
- Should we try to model the cycles?

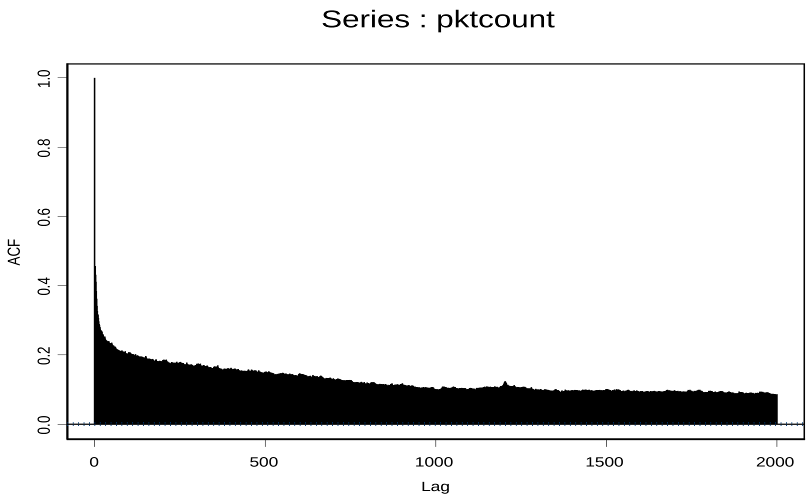
Tech Pt 5. **Long range dependence**. A stationary L_2 sequence $\{\xi_n, n \geq 1\}$ has long-range dependence if

$$\mathrm{Cov}(\xi_n, \xi_{n+h}) \sim h^{-\beta}, \quad h \to \infty$$

for $0 < \beta < 1$.

How to test? Q&D method: The sample acf should not $\to 0$ quickly as the lag increases.

EXAMPLE: CompanyX–packet counts per unit time on CompanyX's WAN (including trans-Atlantic traffic).


Series : pktcount
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000

7. How do heavy tails cause long range dependence?

Assume infinite source Poisson model with

$$R_k = 1 \qquad \Rightarrow \qquad F_k = L_k.$$

$$1 < \alpha_L < 2.$$

Recall

$$\begin{aligned} M(t) &= \#\text{ active sources at time } t,\\ &= \text{total input rate at time } t,\\ &= \text{analogue of packet count per unit time.} \end{aligned}$$

Background and warmup:

For each fixed t, $M(t)$ is a Poisson random variable.
Why? When $1 < \alpha_L < 2$, $M(\cdot)$ has a stationary version. Assume

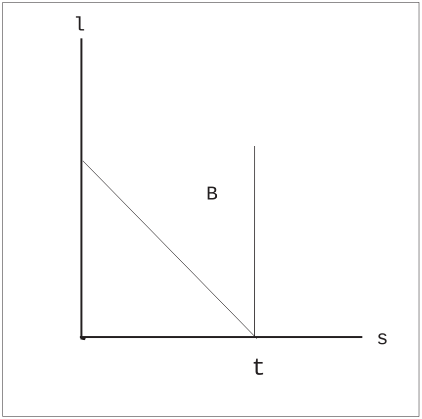
$$\sum_k \epsilon_{\Gamma_k} = \mathrm{PRM}(\lambda dt)$$

on $\mathbb{R}$. Then

$$\xi := \sum_k \epsilon_{(\Gamma_k, L_k)} = \mathrm{PRM}(\lambda dt \times F_L)$$

on $\mathbb{R} \times [0, \infty)$ and

$$M(t) = \sum_k 1_{[\Gamma_k \le t < \Gamma_k + L_k]}$$
$$= \xi(\{(s, l) : s \le t < s + l\} = \xi(B)$$



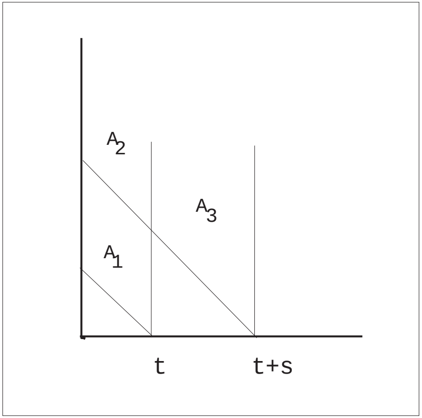
is Poisson with mean

$$E\Big(\xi(\{(s, l) : s \le t < s + l\}\Big)$$
$$= \int_{s=-\infty}^{t} \bar{F}_L(t - s)\lambda ds = \lambda \mu_L$$

The process $\{M(t), t \in \mathbb{R}\}$ is stationary with covariance function

$$\begin{aligned}&\mathrm{Cov}(M(t), M(t+s))\\ &\quad =\mathrm{Cov}\big(\xi(A_1)+\xi(A_2), \xi(A_2)+\xi(A_3)\big)\end{aligned}$$

and because $\xi(A_1) \perp\!\!\!\perp \xi(A_3)$, this is



$$\begin{aligned}&=\mathrm{Cov}\big(\xi(A_2), \xi(A_2)\big)\\ &=\mathrm{Var}\big(\xi(A_2)\big)\\ &=E\big(\xi(A_2)\big)\\ &=\int_{u=-\infty}^{t} \lambda du\, \bar{F}_L(t+s-u)\\ &=\lambda \int_s^\infty \bar{F}_L(v)dv\\ &=\lambda s \bar{F}_L(s) \cdot c = c' s^{-(\alpha-1)} \ell(s).\end{aligned}$$

The slow decay of the covariance as a function of the lag s characterizes *long range dependence*. □

8. References

Contents

References

M. Arlitt and C.L. Williamson. Web server workload characterization: The search for invariants (extended version). In *Proceedings of the ACM Sigmetrics Conference, Philadelphia, Pa*, 1996. Available from {mfa126,carey}@cs.usask.ca.

M. Crovella and A. Bestavros. Self-similarity in world wide web traffic: evidence and possible causes. In *Proceedings of the ACM SIGMETRICS '96 International Conference on Measurement and Modeling of Computer Systems*, volume 24, pages 160–169, 1996.

M. Crovella and A. Bestavros. Self-similarity in world wide web traffic: evidence and possible causes. *IEEE/ACM Transactions on Networking*, 5(6):835–846, 1997.

B. D'Auria and S.I. Resnick. Data network models of burstiness. *Adv. in Appl. Probab.*, 38(2):373–404, June 2006a.

B. D'Auria and S.I. Resnick. The influence of dependence on data network models of burstiness. Technical report, Cornell University, 2006b. Report #1449, Available at legacy.orie.cornell.edu/~sid.

D.E. Duffy, A.A. McIntosh, M. Rosenstein, and W. Willinger. Analyzing telecommunications traffic data from working common channel signaling subnetworks. In M.E. Tarter and M.D. Lock, editors, *Computing Science and Statistics Interface, Proceedings of the 25th Symposium on the Interface*, volume 25, pages 156– 165, San Diego, California, 1993.

C.A. Guerin, H. Nyberg, O. Perrin, S.I. Resnick, H. Rootzén, and C. Stărică. Empirical testing of the infinite source poisson data traffic model. *Stochastic Models*, 19(2):151–200, 2003.

W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. Statistical analysis of high time-resolution ethernet Lan traffic measurements. In *Proceedings of the 25th Symposium on the Interface between Statistics and Computer Science*, pages 146– 155, 1993.

Will E. Leland, Murad S. Taqqu, Walter Willinger, and Daniel V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Trans. Netw.*, 2(1):1–15, 1994. ISSN 1063-6692. doi: http://dx.doi.org/10.1109/90.282603.

K. Maulik, S.I. Resnick, and H. Rootzén. Asymptotic independence and a network traffic model. *J. Appl. Probab.*, 39(4):671–699, 2002. ISSN 0021-9002.

T. Mikosch, S.I. Resnick, H. Rootzén, and A.W. Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.*, 12(1):23–68, 2002.

CORNELL

S.I. Resnick and H. Rootzén. Self-similar communication models and very heavy tails. *Ann. Appl. Probab.*, 10:753–778, 2000.

S.I. Resnick. Modeling data networks. In B. Finkenstadt and H. Rootzén, editors, *SemStat: Seminaire Europeen de Statistique, Extreme Values in Finance, Telecommunications, and the Environment*, pages 287–372. Chapman-Hall, London, 2003.

W. Willinger and V. Paxson. Where mathematics meets the Internet. *Notices of the American Mathematical Society*, 45(8):961–970, 1998.

W. Willinger, M.S. Taqqu, M. Leland, and D. Wilson. Self–similarity in high–speed packet traffic: analysis and modelling of ethernet traffic measurements. *Statist. Sci.*, 10:67–85, 1995.

W. Willinger, M.S. Taqqu, M. Leland, and D. Wilson. Self–similarity through high variability: statistical analysis of ethernet lan traffic at the source level (extended version). *IEEE/ACM Transactions on Networking*, 5(1):71–96, 1997.

W. Willinger, V. Paxson, and M.S. Taqqu. Self-similarity and heavy tails: Structural modeling of network traffic. In R.J. Adler, R.E. Feldman, and M.S. Taqqu, editors, *A Practical Guide to Heavy Tails. Statistical Techniques and Applications*, pages 27–53. Birkhäuser Boston Inc., Boston, MA, 1998.

W. Willinger. Data network traffic: heavy tails are here to stay. Presentation at Extremes–Risk and Safety, Nordic School of Public Health, Gothenberg Sweden, August 1998.

Title Page

Page 30 of 30

Go Back

Full Screen

Close

Quit

Applied Probability Modeling of Data Networks;

4. Some Stylized Applied Probability Network Models: A Renewal Input Model.

Sidney Resnick
School of Operations Research and Industrial Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA
http://legacy.orie.cornell.edu/∼sid
sir1@cornell.edu

August 1, 2007

1. Introduction: Very Heavy Tails.

Attempts to explain network self-similarity focus on heavy tailed transmission times of sources sending data to one or more servers:

$$P[\text{On-period duration } > x] \approx x^{-\beta}.$$

Reasons for assuming $1 < \beta < 2$:

- (Applied) Willinger et al (Bellcore) analyzed 700+ source destination pairs, and estimated the tail parameter of *on-periods*. Value usually in the range $(1, 2)$.
- (Theoretical) Mathematical analysis has been based on renewal theory. Without a finite mean, stationary versions of renewal processes do not exist and (uncontrolled) buffer content stochastic processes would not be stable.

BUT

Need for the case $0 < \beta < 1$:

- BU study (Crovella, Bestavros, ...) of file sizes downloaded in www session over 4 months in 2 labs. In November, 1994 in room 272: $\beta \approx .66$.
- Calgary study of file lengths downloaded from various servers found β's in the range of 0.4 to 0.6.

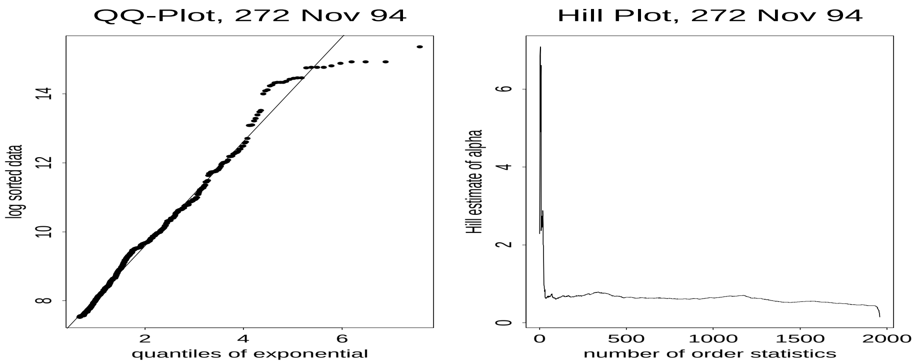


Figure 1: BU File Lengths, Nov 1994, Room 272

1.1. Summary of the difficulties with the case $0 < \beta < 1$.

- Models will not be stable in the conventional senses; normalization necessary to keep control.
- Models will not have stationary versions.
- Some common performance measures which are expressed in terms of moments, may not be applicable.
- Nervousness about models where moments do not exist.
- Confusion between concepts of unbounded support and infinite moments. (Normal, exponential, gamma, weibull, … have unbounded support.)

2. Infinite source Poisson model=M/G/∞ input model

Notation and concepts:

$$
\begin{aligned}
\{\Gamma_n, n \geq 1\} &= \text{times of session initiations; homogeneous Poisson} \\
&\quad\ \text{points on } [0,\infty) \text{ with rate } \lambda. \\
\{L_n\} &= \text{session durations; iid non-negative rv's} \\
&\quad\ \text{with common distribution } F_L(x) \text{ satisfying} \\
&\quad\ \bar{F}_L(x) \sim x^{-\beta}\ell(x), \quad 0<\beta<1; \\
m(t) &= \int_0^t \bar{F}_L(s)ds \sim ct^{1-\beta}\ell(t) \uparrow \infty;\ \text{truncated mean function,} \\
M(t) &= \text{number of active sessions at } t, \\
&= \sum_{n=1}^{\infty} 1_{[\Gamma_n \leq t < \Gamma_n + L_n]}; \quad E(M(t)) = m(t); \\
&= \#\text{ of servers in M/G/}\infty\text{ telephone model.} \\
A(t) &= \int_0^t M(s)ds, \\
&= \text{cumulative work inputted in } [0,t] \text{ assuming unit rate input.}
\end{aligned}
$$

$$
\begin{aligned}
r &= \text{server work rate,}\\
X(t) &= \text{content process,}\\
dX(t) &= dA(t) - r1_{[X(t)>0]}dt\\
\tau(\gamma) &= \inf\{t > 0 : X(t) \geq \gamma\}
\end{aligned}
$$

3. Sessions initiated at renewal times. (Mikosch)

See Mikosch and Resnick (2006).

3.1. The model.

$$\begin{aligned}
\{S_n, n \geq 1\} &= \text{times of session initiations; ordinary renewal process;} \\
&\quad S_0 = 0,\ S_n = \sum_{i=1}^{n} X_i;\ \{X_n\} \text{ iid;} \\
&\quad X_i \sim F_X(x);\ \bar{F}_X(x) = 1 - F_X(x) \in RV_{-\alpha} \\
\{L_n\} &= \text{session durations; iid non-negative rv's} \\
&\quad \text{with common distribution } F_L(x) \text{ satisfying} \\
&\quad \bar{F}_L(x) \in RV_{-\beta}, \\
&\quad \{L_n\} \perp\!\!\!\perp \{S_n\}. \\
M(t) &= \text{number of active sessions at } t, \\
&= \sum_{n=1}^{\infty} 1_{[S_n \leq t < S_n + L_n]} \\
A(t) &= \int_0^t M(s)ds, \\
&= \text{cumulative work inputted in } [0,t], \text{ assuming unit rate input.}
\end{aligned}$$

3.2. Cases.

1. COMPARABLE TAILS: $\beta = \alpha$ and $\bar{F}_X(x) \sim c\,\bar{F}_L(x)$, $c > 0$, as $x \to \infty$.
   (a) The distribution tails of X_1 and L_1 are essentially the same.
   (b) For simplicity, we assume $c = 1$.
   (c) Kind of stability for M which converges weakly w/o normalization.
2. F_L HEAVIER-TAILED:
   (a) $0 < \beta < \alpha < 1$ or, if $\beta = \alpha$, then $\bar{F}_X(x)/\bar{F}_L(x) \to 0$ as $x \to \infty$ so that the distribution tail of X_1 is lighter than the distribution tail of L_1.
   (b) $0 = \beta < \alpha < 1$ so that the distribution tail of L_1 is slowly varying and thus again heavier than that of X_1.
   (a) Implies buildup in the M process.
3. F_X HEAVIER-TAILED: $\beta > \alpha$ so that the distribution tail of X_1 is heavier than the distribution tail of L_1.
   - Renewal epochs sparse relative to session lengths.
   - Of less interest.

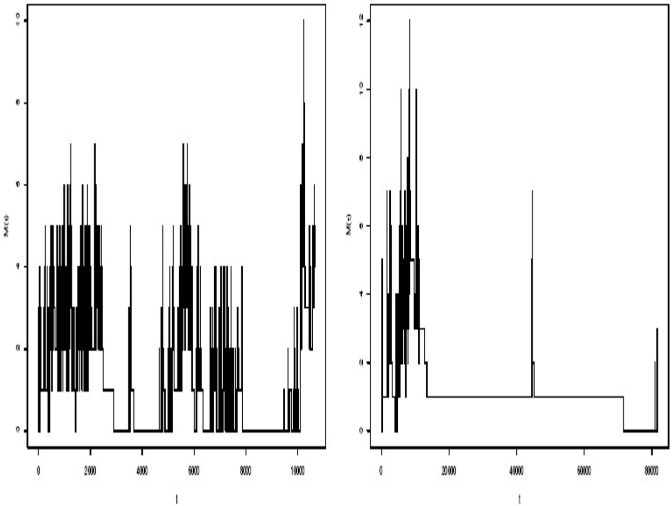


Figure 2: Paths of M; $\alpha = \beta = 0.9$ (left); $\alpha = \beta = 0.6$ (right).

CORNELL
Intro
Poisson sources
Renewal: sessions
Renewal: work
Title Page
Page 9 of 23
Go Back
Full Screen
Close
Quit

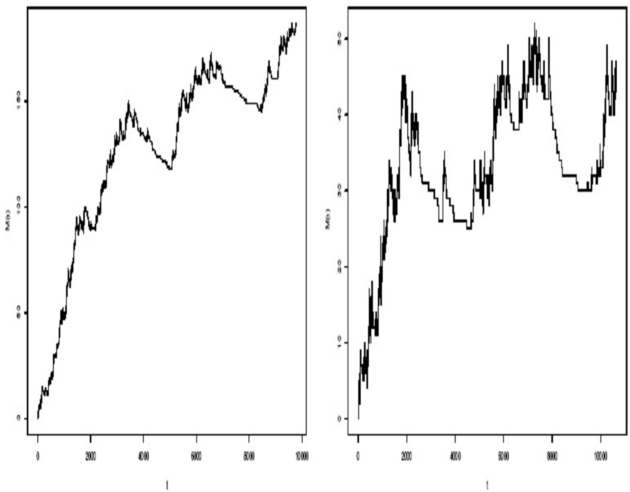


Figure 3: Paths of M; $(\alpha, \beta) = (0.9, 0.2)$ (left); $(\alpha, \beta) = (0.9, 0.4)$ (right).

3.3. Warm-up: Mean value analysis for $\alpha, \beta < 1$.

Obtain asymptotic behavior of $E\bigl(M(t)\bigr)$ from Karamata's Tauberian theorem. Let

$$U(x) = \sum_{n=0}^{\infty} F_X^{n*}(x), \quad x > 0, \quad = \text{renewal function.}$$

Since $0 < \alpha < 1$, well known (eg Feller, 1971); as $x \to \infty$,

$$U(x) \sim \bigl(\Gamma(1-\alpha)\,\Gamma(1+\alpha)\,\bar F_X(x)\bigr)^{-1} \sim c(\alpha)\,x^{\alpha}/\ell_F(x).$$

Therefore ($t \to \infty$),

$$\begin{aligned}\mathbb{E}M(t) &= \int_0^t U(dx)\,\bar F_L(t-x) = \int_0^1 \frac{\bar F_L(t(1-s))}{\bar F_L(t)}\frac{U(tds)}{U(t)}\bigl(\bar F_L(t)U(t)\bigr)\\ &\sim c(\alpha)\int_0^1 (1-s)^{-\beta}\alpha s^{\alpha-1}ds\,\frac{\bar F_L(t)}{\bar F_X(t)} = c'(\alpha)\frac{\bar F_L(t)}{\bar F_X(t)}.\end{aligned}$$

$$\mathbb{E}M(t) \sim c'(\alpha)\frac{\bar{F}_L(t)}{\bar{F}_X(t)}, \quad (t \to \infty).$$

Conclusions

- Case (1): Comparable tails.

$$E\big(M(t)\big) \text{ converges to a constant.}$$

- Case (2): F_L is more heavy-tailed than F.

$$\mathbb{E}M(t) \to \infty.$$

- Case (3): F_X is more heavy-tailed than F_L.

$$\mathbb{E}M(t) \to 0.$$

  and hence

$$M(t) \xrightarrow{L_1} 0.$$

  so Case (3) may be of lesser interest. (Renewals are sparse relative to event durations that at any time there is not likely to be an event in progress.)

3.4. Summary of results.

Table 1: Limiting behavior of $M(t)$ as $t \to \infty$.

Conditions	Limit behavior of $M(t)$ as $t \to \infty$
$0 < \alpha < 1$ $\bar{F}_X \sim \bar{F}_L$	$M(t) \Rightarrow$ random limit.
$0 \le \beta < \alpha < 1$ or $0 < \alpha = \beta < 1$ and $\bar{F}_X = o(\bar{F}_L)$	$\frac{\bar{F}_X(t)}{\bar{F}_L(t)} M(t) \Rightarrow$ random limit.
$0 < \beta < 1$ $\mathbb{E}(X_1) < \infty$	$\frac{M(t)}{t\bar{F}_L(t)} \Rightarrow$ constant $\frac{M(t)- \text{ random centering}}{\sqrt{t\bar{F}_L(t)}} \Rightarrow$ Gaussian rv
$0 < \beta \le \alpha = 1$ $\mathbb{E}(X_1) = \infty$	$\frac{M(t)}{t\bar{F}_L(t)\mu(t)} \Rightarrow$ constant $\mu(t)=$ truncated 1st moment X
$\mathbb{E}(X_1) < \infty$ $\mathbb{E}(L_1) < \infty$	Stationary version of $M(\cdot)$ exists

Focus on the first 2 rows corresponding to $\alpha,\ \beta < 1$.

3.5. Renewal: $\alpha, \beta < 1$, comparable tails.

A kind of stability exists for this case since fidi's of $M(t\cdot)$ converge in distribution to a limit.

3.5.1. Preliminaries

Define

$$N(x) = \sum_{n=0}^{\infty} 1_{[S_n \le x]} = \inf\{n : S_n > x\} = S^{\leftarrow}(x), \quad x \ge 0.$$

$$= \text{renewal counting function.}$$

$$\sum_k \epsilon_{(t_k, j_k)} = N_\infty = \text{ PRM(Leb} \times \nu_\alpha) \text{ on } [0,\infty) \times (0,\infty] := \mathbb{E}$$

$$\nu_\alpha(x,\infty] = x^{-\alpha}$$

$$X_\alpha(t) = \sum_{t_k \le t} j_k, \quad t \ge 0,$$

$$=\text{non-decreasing } \alpha\text{-stable Lévy motion with Lévy measure } \nu_\alpha.$$

$$b(t) \sim \left(\frac{1}{1-F_X}\right)^{\leftarrow}(t), \quad t \to \infty, \quad t\bar{F}_X(b(t)) \sim 1;$$

$$= \text{quantile function of } F_X$$

$$X^{(s)}(t) = \frac{S_{[st]}}{b(s)} \Rightarrow X_\alpha(t), \quad (s \to \infty),$$

$$= \text{renewal epochs are asymptotically stable.}$$

$$(X^{(s)})^{\leftarrow} \Rightarrow X_\alpha^{\leftarrow}, \quad \bar{F}_X(s) N(s\cdot) \Rightarrow X_\alpha^{\leftarrow}(\cdot)$$

or, setting $u^{-1} = \bar{F}_X(s)$,

$$\frac{1}{u} N(b(u)\cdot) \Rightarrow X_\alpha^{\leftarrow}(\cdot), \quad u \to \infty,$$

$$\frac{1}{u} \sum_{n=0}^{\infty} \epsilon_{\frac{S_n}{b(u)}} \Rightarrow X_\alpha^{\leftarrow} \quad \text{in } M_+[0,\infty).$$

Have not yet referenced L.

3.6. Comparable tails; $\alpha,\ \beta < 1$.

Define the time change map by

$$T : \mathbb{D}^{\uparrow}[0,\infty) \times M_+(\mathbb{E}) \mapsto M_+(\mathbb{E}), \qquad (\mathbb{E} = [0,\infty) \times (0,\infty])$$

by

$$T(x,m) = \tilde{m}$$

where $\tilde{m}$ is defined by

$$\tilde{m}(f) = \iint f(x(u),v)\, m(du,dv)\,, \quad f \in C_K^+(\mathbb{E})\,.$$

If m is a point measure with representation $m = \sum_k \epsilon_{(\tau_k, y_k)}$, then

$$T(x,m) = \sum_k \epsilon_{(x(\tau_k),y_k)}\,.$$

Steps for analysis: $\bar F_L \sim \bar F_X$; $\alpha=\beta$.

1. in $M_p(\mathbb{E})$, as $s\to\infty$, regular variation of $\bar F_L$ equiv to
$$\sum_{k=0}^{\infty}\epsilon_{(\frac{k}{s},\frac{L_k}{b(s)})}\Rightarrow N_\infty = PRM(\text{Leb}\times\nu_\alpha).$$
2. Since $\{S_k\}$ is independent of $\{L_k\}$, get joint convergence in $D[0,\infty)\times M_p(\mathbb{E})$,
$$\Big(\frac{S_{[s\cdot]}}{b(s)},\sum_{k=0}^{\infty}\epsilon_{(\frac{k}{s},\frac{L_k}{b(s)})}\Big)\Rightarrow\big(X_\alpha,N_\infty\big)\,.$$
3. Apply the a.s. continuous function T:
$$T\Big(\frac{S[s\cdot]}{b(s)},\sum_{k=0}^{\infty}\epsilon_{(\frac{k}{s},\frac{L_k}{b(s)})}\Big)=\sum_{k=0}^{\infty}\epsilon_{(\frac{S_{[sk/s]}}{b(s)},\frac{L_k}{b(s)})}=$$
$$\sum_{k=0}^{\infty}\epsilon_{(\frac{S_k}{b(s)},\frac{L_k}{b(s)})}\Rightarrow T\big(X_\alpha,N_\infty\big)\,.$$
4. Evaluate on $\{(u,v): u\le t\le u+v\}$ to get result for M: The fidi's of $M(t)$ satisfy as $s\to\infty$,
$$M(s\,t)=\sum_{k=0}^{\infty}1_{[\frac{S_k}{s}\le t<\frac{S_k+L_k}{s}]}\Rightarrow M_\infty(t)=\sum_k 1_{[X_\alpha(t_k)\le t<X_\alpha(t_k)+j_k]}\,.$$

3.7. Case 2: $\bar{F}_L$ heavier; $\alpha, \beta < 1$.

Ingredients for analysis:

1. Recall $b(t)$ is the quantile function of F_X and satisfies
$$s\bar{F}_X(b(s)) \to 1, \quad (s \to \infty).$$
2. Since $\bar{F}_L \in RV_{-\beta}$,
$$\frac{s\,\bar{F}_X(b(s))}{\bar{F}_L(b(s))} F_L(b(s)\cdot) \xrightarrow{v} \nu_\beta.$$
in $M_+(0,\infty]$, where $\xrightarrow{v}$ denotes vague convergence.
3. Equivalent to previous convergence is (variant of basic convergence)
$$\frac{\bar{F}_X(b(s))}{\bar{F}_L(b(s))} \sum_{k=0}^{[s]} \varepsilon_{\frac{L_k}{b(s)}} \Rightarrow \nu_\beta\,.$$
4. Extend by adding time component:
$$\frac{\bar{F}_X(b(s))}{\bar{F}_L(b(s)} \sum_{k=0}^{\infty} \varepsilon_{(\frac{k}{s}, \frac{L_k}{b(s)})} \Rightarrow \mathrm{Leb} \times \nu_\beta\,.$$

5. Augment using independence of $\{L_n\}$ and $\{S_n\}$:

$$\left(\frac{S_{[s\cdot]}}{b(s)}\,,\, \frac{\bar{F}_X(b(s))}{\bar{F}_L(b(s))} \sum_{k=0}^{\infty} \varepsilon_{\left(\frac{k}{s}\,,\,\frac{L_k}{b(s)}\right)} \right) \Rightarrow (X_\alpha\,, \mathrm{Leb} \times \nu_\beta)\,.$$

6. Apply the a.s. continuous map T; evaluate the result on the correct region to get result for M: The fidi' s of M satisfy

$$\frac{\bar{F}_X(s)}{\bar{F}_L(s)}\, M(s\,t) \Rightarrow \int_0^t (t-u)^{-\beta}\, dX_\alpha^{\leftarrow}(u). \quad s \to \infty.$$

4. Cumulative work process.

Sample results for workload process

$$A(t) = \int_0^t M(s)ds.$$

4.1. The case $\mu_X < \infty$, $\beta \in (1,2)$

Define the quantile function of F_L:

$$\sigma(t) \sim \left(\frac{1}{1-F_L}\right)^{\leftarrow}(t)\,, \quad t \to \infty.$$

Assume either $\bar{F}_X \in RV_{-\alpha}, 1 < \alpha \leq 2$ or $\sigma_X^2 < \infty$.

1. Suppose $\bar{F}_X \in RV_{-\alpha}$ and either
   (a) $\alpha > \beta$ or
   (b) $\alpha = \beta$ and $\bar{F}_X(x) = o(\bar{F}_L(x))$ or
   (c) $\sigma_X^2 < \infty$.

   Set

   $$A_s(u) = \sigma(s)^{-1}\Big(A(su) - su\mu_L/\mu_X\Big), \quad u \geq 0.$$

   Then $(s \to \infty)$,

   $$A_s(\cdot) \Rightarrow \mu_X^{-1/\beta} X_\beta(\cdot)\,, \tag{1}$$

where X_β is a β-stable spectrally positive Lévy motion on $[0,\infty)$.

2. If $\bar{F}_X \in RV_{-\alpha}$, $\alpha = \beta$ and $\bar{F}_X(x) \sim c\,\bar{F}_L(x)$, then (1) holds, where X_β is β-stable Lévy motion with a skewness parameter.

3. If $\bar{F}_X \in RV_{-\alpha}$ and $\alpha < \beta$ or $\alpha = \beta$ and $\bar{F}_L(x) = o(\bar{F}_X(x))$, then, as $s \to \infty$
$$(b(s))^{-1}\left[A(\cdot\, s) - s\,(\cdot)\,\mu_L/\mu_X\right] \Rightarrow \mu_X^{-1/\alpha}\,X_\alpha(\cdot)\,,$$
where X_α is spectrally negative α-stable Lévy motion.

4.2. $\alpha, \beta < 1$.

Case (2) assumptions hold: $0 \le \beta \le \alpha < 1$ and if $\alpha = \beta$, then $\bar{F}_X(s)/\bar{F}_L(s) \to 0$, as $s \to \infty$. Then
$$\frac{\bar{F}_X(s)}{s\bar{F}_L(s)}A(st) \Rightarrow \int_0^t \frac{(t-u)^{1-\beta}}{1-\beta}dX_\alpha^{\leftarrow}(u), \quad t \ge 0\,,$$
in $C[0,\infty)$.

NB: This is the integrated version of the result for M.

Contents

References

T. Mikosch and S.I. Resnick. Activity rates with very heavy tails. *Stochastic Process. Appl.*, 116:131–155, 2006.

Applied Probability Modeling of Data Networks;

5. Some (More) Stylized Applied Probability

Network Models: Large Time Scale Input Model.

Sidney Resnick
School of Operations Research and Industrial Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA
http://legacy.orie.cornell.edu/∼sid
sir1@cornell.edu

August 5, 2007

1. Infinite Source Poisson Fluid Input Model

Suppose *sessions* characterized by

- Session initiation times are $\{\Gamma_k\}$ where

$$\{\Gamma_k\} \sim \text{ homogeneous Poisson on } (-\infty, \infty), \text{ rate } \lambda.$$

- Sequence of iid marks independent of $\{\Gamma_k\}$: Each Poisson point Γ_k receives a *mark* which characterizes input characteristics:

$$(F_k, L_k, R_k) = (\text{file, duration, rate }),$$

  where

$$F_k = L_k R_k.$$

  Assume

$$R_k \equiv 1 \text{ so } F_k = L_k \text{ and } 1 < \alpha_F = \alpha < 2.$$

- Number of sessions in progress at t

$$M(t) = \sum_{k=1}^{\infty} 1_{[\Gamma_k \le t < \Gamma_k + L_k]}$$

- Cumulative input in $[0,t]$,

$$A(t)=\int_0^t M(s)ds.$$

- $\{\Gamma_k\} \perp\!\!\!\perp \{L_k\}$
- Session length distribution has regularly varying tail:

$$\mathbb{P}(L_k>x)=\bar{F}_L(x)=x^{-\alpha}\ell(x), \quad x>0,\ 1<\alpha<2\,, \tag{1}$$

  where ℓ is a slowly varying function.
- $\alpha\in(1,2)$ means the variance of L_k is infinite and its mean μ_L is finite.
- Quantile function

$$b(t)=\left(\frac{1}{1-F_L}\right)^{\leftarrow}(t)=:\inf\{x:\frac{1}{1-F_L(x)}\geq t\}, \quad t>0, \tag{2}$$

  which is regularly varying with index $1/\alpha$.
- Will scale cumulative input by $T\to\infty$ and let $\lambda=\lambda(T)\to\infty$ (At what rate?) and consider for the Tth model:

$$A_T(t)=A(Tt),$$

  and

$$M_T(t)=\ \#\text{ active sources in }T\text{th model}.$$

- Family of models indexed by T. As we move through the family by letting $T \to \infty$, is there an informative limit?

2. Growth conditions

- $\lambda = \lambda(T)$: the parameter governing the initiation of sessions in the Tth model;
- Suppose $\lambda = \lambda(T)$ is non-decreasing function of T.
- Asymptotic behavior of $A_T(\cdot)$ depends on whether *slow growth* or *fast growth* holds for $\lambda(T)$.

We phrase our condition first in terms of the quantile function b defined in (2).

Lemma 1 (Slow/fast growth) *Sppse*

- $\bar{F}_L \in RV_{-\alpha}$, $1 < \alpha < 2$.
- *Session initiations form a stationary PRM on* $\mathbb{R}$, *rate* $\lambda = \lambda(T)$.
- $M_T(t)$, *the number of active sources at time* t *in the* T*th model when* $\lambda = \lambda(T)$, *is a stationary process on* $\mathbb{R}$.

Then with $\lambda = \lambda(T)$, the following are euivalent and define slow growth:

1.
$$\lim_{T\to\infty} \frac{b(\lambda T)}{T} = 0,$$

2.
$$\lim_{T\to\infty} \lambda T\, \bar{F}_L(T) = 0,$$

3.
$$\lim_{T\to\infty} \mathrm{Cov}\,(M_T(0), M_T(T)) = 0.$$

Fast growth is defined by replacing 0 *by* ∞.

3. Two very different approximations

Theorem 1 (Mikosch et al. (2002)) *If slow growth holds, then the process* $(A(Tt), t \geq 0)$ *describing the total cumulative input in* $[0, Tt]$, $t \geq 0$, *satisfies the limit relation*

$$X^{(T)}(\cdot) := \frac{A(T\cdot) - T\lambda\mu_L(\cdot)}{b(\lambda T)} \stackrel{fidi}{\rightarrow} X_\alpha(\cdot), \tag{3}$$

where

- $X_\alpha(\cdot)$ *is* α*-stable Lévy motion;*
- $\stackrel{fidi}{\rightarrow}$ *denotes convergence of the finite dimensional distributions.*

Theorem 2 *If fast growth holds, then the process* $(A(Tt), t \geq 0)$ *describing the total accumulated input in* $[0, Tt]$, $t \geq 0$, *satisfies the limit relation*

$$\frac{A(T\cdot) - \lambda\mu_L T(\cdot)}{[\lambda T^3 \bar{F}_L(T)\sigma^2]^{1/2}} \stackrel{d}{\rightarrow} B_H(\cdot)\,,$$

where

- $\stackrel{d}{\rightarrow}$ *denotes weak convergence in* $(\mathbb{D}[0,\infty), J_1)$,
- B_H *is standard fractional Brownian motion,*
- $H = (3-\alpha)/2$

- σ^2 *is a constant.*

See Gaigalas and Kaj (2003), Kaj and Taqqu (2004), Kaj (1999, 2002), Konstantopoulos and Lin (1998), Mikosch and Stegeman (1999), Mikosch et al. (2002), Resnick and van den Berg (2000), Resnick (2003).

4. Why $A(T)$ should be asymptotically stable under slow growth.

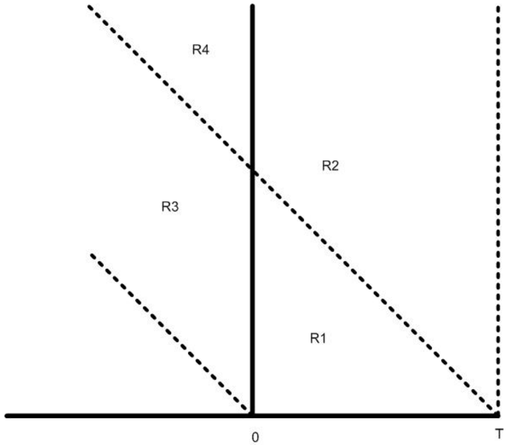
Recall using augmentation

$$\xi := \sum_k \epsilon_{(\Gamma_k, L_k)} = \mathrm{PRM}(\lambda \mathbb{LEB} \times F_L) \tag{4}$$

is a two dimensional Poisson random measure on $\mathbb{R} \times [0, \infty)$ with mean measure

$$\mu := \lambda dt \times F_L(dx).$$

Decompose $A(T)$ by decomposing (initiation time, session length) in $(-\infty, T] \times [0, \infty)$ into 4 regions $R_1, \dots, R_4$:



$$
\begin{aligned}
R_1 :=& \{(s,y) : 0 < s \le T,\ 0 < y, \\
&\qquad s + y \le T\}, \\
R_2 :=& \{(s,y) : 0 < s \le T,\ T < s + y\}, \\
R_3 :=& \{(s,y) : s \le 0,\ 0 < s + y \le T\}, \\
R_4 :=& \{(s,y) : s \le 0,\ T < s + y\},
\end{aligned}
$$

Rewrite $A(T)$

$$
\begin{aligned}
A(T) =& \sum_k L_k 1_{[(\Gamma_k, L_k) \in R_1]} + \sum_k (T - \Gamma_k) 1_{[(\Gamma_k, L_k) \in R_2]} \\
&+ \sum_k (L_k + \Gamma_k) 1_{[(\Gamma_k, L_k) \in R_3]} + \sum_k T 1_{[(\Gamma_k, L_k) \in R_4]} \\
=:& A_1(T) + A_2(T) + A_3(T) + A_4(T).
\end{aligned}
$$

Restriction of $\xi(\cdot)$ to regions R_i:

Recall

$$\xi = \sum_k \epsilon_{(\Gamma_k, L_k)}.$$

- The mean measure $\mu := E\xi(\cdot)$ restricted to R_i is finite for $i = 1, \ldots, 4$.
- Therefore, the points of $\xi\big|_{R_i}$ can be represented as a Poisson number of iid singleton point measures:

$$\xi\big|_{R_i} \stackrel{d}{=} \sum_{k=1}^{P_i} \epsilon_{(t_{k,i}, j_{k,i})}\,, \quad i = 1, \ldots, 4\,,$$

  where

  - P_i is a Poisson random variable which is independent of the iid pairs $(t_{k,i}, j_{k,i})$, $k \geq 1$,
  - with common distribution

$$\frac{\lambda \mathbb{LEB}(ds) F_L(dy)}{\mu(R_i)}\bigg|_{R_i},$$

    for $i = 1, \ldots, 4$.
  - P_i has mean

$$\mu(R_i).$$

Note the mean of P_1 is

$$\begin{aligned}\mu(R_1) =&E\xi(R_1) == \iint_{\substack{0\le s\le T\\ s+y\le T}} \lambda ds F_L(dy)\\ =&\int_{s=0}^{T} \lambda ds \int_{y=0}^{T-s} F_L(dy) = \lambda \int_0^T F_L(s)\, ds\\ \sim&\lambda T, \qquad (T\to\infty).\end{aligned}$$

Reminder

$$A_1(T) = \sum_k L_k 1_{[(\Gamma_k, L_k)\in R_1]} \stackrel{d}{=} \sum_{k=1}^{P_1} j_{k,1}$$

where

$$\{(t_{k,1}, j_{k,1}), k\ge 1\} \perp\!\!\!\perp P_1$$

and

$$\{(t_{k,1}, j_{k,1}), k\ge 1\} \text{ iid } \sim \frac{\lambda ds F_L(dy)}{\mu(R_1)} \text{ on } R_1.$$

So

- $A_1(T)$ is a Poissonized sum;
- Poissonized sums behave asymptotically like ordinary sums if you replace P_1 by $E(P_1) = \mu(R_1) \sim \lambda T$.
- Therefore, to show

$$\frac{A_1(T) - \text{centering}(T)}{b(\lambda T)}$$

is asymptotically stable, it should be enough (modulo the truncated variance condition) to show that

$$\lambda T P[j_{k,1} > b(\lambda T)x] \to x^{-\alpha}, \quad T \to \infty.$$

This is the analogue of

$$\sum_{k=1}^{[n\cdot]} (\boldsymbol{X}_{n,k} - \text{ centering }) \Rightarrow X_0(\cdot),$$

iff

$$\begin{cases} nP[\boldsymbol{X}_{n,1} \in \cdot\,] \xrightarrow{v} \nu(\cdot) \\ \text{truncated 2nd moment condition.} \end{cases}$$

Some blood and guts.

$$\begin{aligned}
\lambda T \mathbb{P}(j_{k,1} >& b(\lambda T)\, x) \\
=& \lambda T \iint_{\substack{0\le s\le T\\ 0\le s+y\le T\\ y>b(\lambda T)x}} \frac{\lambda ds F_L(dy)}{\mu_1(R_1)} \qquad (\text{use } \mu_1(R_1)\sim\lambda T), \\
\sim& \lambda \int_{s=0}^{T-b(\lambda T)x} \Big(\int_{y=b(\lambda T)x}^{T-s} F_L(dy) \Big) ds \\
=& \lambda \int_{s=0}^{T-b(\lambda T)x} [\bar F_L(b(\lambda T)x) - \bar F_L(T-s)] ds \\
=& \lambda \Big[\bar F_L(b(\lambda T)x)(T - b(\lambda T)x) \\
& \quad - \int_{s=0}^{T-b(\lambda T)x} \bar F_L(T-s) ds \Big] \\
\sim& \left(1 - \frac{b(\lambda T)x}{T}\right) \lambda T\, \bar F_L(b(\lambda T)\, x) \\
& \quad - \frac{b(\lambda T)}{T} \int_x^{T/b(\lambda T)} \lambda T\, \bar F_L(b(\lambda T)\, s)\, ds \\
\sim& x^{-\alpha},
\end{aligned} \tag{5}$$

provided $b(\lambda T)/T \to 0 \Leftrightarrow$ slow growth.

Final steps:

- Show truncated variance condition (Karamata theorem)

$$\lim_{\delta\to 0}\limsup_{T\to\infty}\lambda T\mathbf{E}\left(\left(\frac{j_{k,1}}{b(\lambda T)}\right)^2 1_{[\texttt{j\_\{k,1\}}\le b(\lambda T)\delta]}\right)=0.$$

- Show

$$\frac{A_i(T)}{b(\lambda T)}\xrightarrow{P}0,\qquad i=2,3,4.$$

- Show for any fixed t:

$$\frac{A_1(Tt)-Tt\lambda\mu_L}{b(\lambda T)}\Rightarrow X_\alpha(t),\qquad \text{in } \mathbb{R}.$$

- Show fi di convergence.

Contents

References

R. Gaigalas and I. Kaj. Convergence of scaled renewal processes and a packet arrival model. *Bernoulli*, 9(4):671–703, 2003. ISSN 1350-7265.

I. Kaj and M.S. Taqqu. Convergence to fractional brownian motion and to the telecom process: the integral representation approach. Available at Department of Mathematics, Uppsala University, U.U.D. M. 2004:16, 2004.

I. Kaj. Convergence of scaled renewal processes o fractional brownian motion. Preprint: Department of Mathematics, Uppsala University, Box 480, S-751 06 Uppsala, Sweden, 1999.

I. Kaj. *Stochastic Modeling in Broadband Communications Systems*. SIAM Monographs on Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. ISBN 0-89871-519-9.

T. Konstantopoulos and S.J. Lin. Macroscopic models for long-range dependent network traffic. *Queueing Systems Theory Appl.*, 28(1-3): 215–243, 1998. ISSN 0257-0130.

T. Mikosch and A. Stegeman. The interplay between heavy tails and rates in self-similar network traffic. Technical Report. University of Groningen, Department of Mathematics, Available via www.cs.rug.nl/˜eke/iwi/preprints/, 1999.

T. Mikosch, S.I. Resnick, H. Rootzén, and A.W. Stegeman. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.*, 12(1):23–68, 2002.

S.I. Resnick and E. van den Berg. Weak convergence of high-speed network traffic models. *J. Appl. Probab.*, 37(2):575–597, 2000.

S.I. Resnick. Modeling data networks. In B. Finkenstadt and H. Rootzén, editors, *SemStat: Seminaire Europeen de Statistique, Extreme Values in Finance, Telecommunications, and the Environment*, pages 287–372. Chapman-Hall, London, 2003.

CORNELL

Applied Probability Modeling of Data Networks;

5.5 Some (More More) Stylized Applied Probability Models:

Heavy Traffic and Heavy Tails.

Sidney Resnick
School of Operations Research and Industrial Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA
http://legacy.orie.cornell.edu/∼sid
sir1@cornell.edu

August 5, 2007

1. Heavy Traffic Approximation to Negative Drift Random Walks

- Aim: understanding of the heavy traffic approximation for stationary waiting time distribution in GI/G/1 queueing model when service times heavy tailed.
- Stationary waiting time distribution in a stable Lindley queue expressed as the distribution of the supremum of a negative drift random walk.
- Begin with negative drift random walk.

Asmussen (2003), Boxma and Cohen (1999), Cohen (1998), Furrer (1997, 1998), Resnick and Samorodnitsky (2000)

1.1. Approximation to a negative drift random walk.

Setup:

- For each $k = 1, 2, \ldots$, $\{X_i^{(k)}, i \geq 1\}$ are iid.
- The kth random walk is

$$S_0^{(k)} = 0, \quad S_n^{(k)} = \sum_{i=1}^{n} X_i^{(k)}, \quad n \geq 1,$$

so that the kth random walk has steps $X_i^{(k)}$, $i = 1, 2, \ldots$.

Assumptions:

Assumption 1. $\exists$ sequence of integers $0 < d(k) \to \infty$ such that

$$\nu_k(\cdot) := d(k)\mathbb{P}[X_1^{(k)} \in \cdot] \xrightarrow{v} \nu(\cdot), \tag{1}$$

vaguely in $[-\infty, \infty] \setminus \{0\}$, where ν is a measure on $[-\infty, \infty] \setminus \{0\}$ satisfying

(a) ν is a Lévy measure

(b) $\nu(-\infty, 0) = 0$, (spectrally positive)

(c) $\int_1^\infty x\nu(dx) < \infty$.

Assumption 2. Control mass near 0: for any $M > 0$

$$\lim_{\epsilon \downarrow 0} \limsup_{k \to \infty} d(k)\mathbf{E}\Big(\big(X_1^{(k)}\big)^2 1_{[|X_1^{(k)}| \le \epsilon]}\Big) = 0. \tag{2}$$

Assumption 3. Each $X_i^{(k)}$ has finite negative mean:

$$E(X_i^{(k)}) = \mu^{(k)} < 0, \qquad i \ge 1$$

satisfying

$$\lim_{k \to \infty} d(k)\mu^{(k)} = -1,$$

which implies $0 > \mu^{(k)} \to 0$ as $k \to \infty$.

Assumption 4. Additional control on ν near ∞:

$$\lim_{M \to \infty} \limsup_{k \to \infty} d(k)E\left(|X_i^{(k)}| 1_{[|X_i^{(k)}| > M]}\right) = 0. \tag{3}$$

How a sequence of negative drift random walks are approximated by a negative drift Lévy process.

Proposition 1 *Assume Assumptions 1–4 hold. Define the random element of* $D[0,\infty)$

$$Y^{(k)}(t) = S^{(k)}_{[d(k)t]}, \quad t \geq 0$$

for $k = 1, 2, \ldots$. *Define*

- $\{\xi^{(\infty)}(t), t \geq 0\}$ *is a totally skewed to the right zero mean Lévy process with Lévy measure* ν
- *and define*
$$Y^{(\infty)}(t) = \xi^{(\infty)}(t) - t, \; t \geq 0.$$

Then in $D[0,\infty)$
$$Y^{(k)}(\cdot) \Rightarrow Y^{(\infty)}(\cdot).$$

Proof: Step 1. Use standard convergence to Lévy method:

$$X^{(k)}(\cdot) := \sum_{i=1}^{[d(k)\cdot]} \left(X_i^{(k)} - E\left(X_i^{(k)} 1_{[|X_k^{(k)}|\le 1]}\right)\right) \Rightarrow X^{(\infty)}(\cdot) \qquad (4)$$

in $D[0,\infty)$ where

$$\begin{aligned} X^{(\infty)}(\cdot) :=& \lim_{\delta\downarrow 0}\Big(\sum_{t_k \le (\cdot)} j_k 1_{[j_k>\delta]} - (\cdot)\int_{\delta<x\le 1} x\nu(dx)\Big) \\ =& \lim_{\delta\downarrow 0}\Bigg(\sum_{t_k\le(\cdot)} j_k 1_{\left[j_k\in(\delta,1]\right]} - (\cdot)\int_{\delta<x\le 1} x\nu(dx) \\ & + \sum_{t_k\le(\cdot)} j_k 1_{\left[j_k>1\right]} - (\cdot)\int_{x>1} x\nu(dx) + (\cdot)\int_{x>1} x\nu(dx)\Bigg) \\ =& \xi^{(\infty)}(\cdot) + (\cdot)\int_{x>1} x\nu(dx), \end{aligned}$$

and $\xi^{(\infty)}(t)$ is

- totally skewed to the right,
- has Lévy measure ν,
- zero mean.

STEP 2. Center (4) to zero expectations:

$$d(k)\left(\mu^{(k)} - E\big(X_1^{(k)} 1_{[|X_1^{(k)}|\le 1]}\big)\right) = d(k)E\big(X_1^{(k)} 1_{[|X_1^{(k)}|>1]}\big)$$
$$\to \int_{x>1} x\nu(dx).$$

Use vague convergence (1) [Assumption 1], (3) of Assumption 4 and Assumption 1c.

STEP 3. Conclude:

$$\sum_{i=1}^{[d(k)t]} X_i^{(k)} - [d(k)t]\mu^{(k)} \Rightarrow X^{(\infty)}(t) - t\int_1^\infty x\nu(dx)$$

in $D[0,\infty)$ and since (Assumption 3)

$$\lim_{k\to\infty} d(k)\mu^{(k)} = -1,$$

we get

$$Y^{(k)}(t) = S^{(k)}_{[d(k)t]} \Rightarrow X^{(\infty)}(t) - t\int_1^\infty x\nu(dx) - t$$
$$= \xi^{(\infty)}(t) - t =: Y^{(\infty)}(t)$$

in $D[0,\infty)$. $\square$

2. Approximation to the supremum of a negative drift random walk.

Reminder: The supremum of negative drift random walk related to the equilibrium waiting time of GI/G/1 queue.

Proposition 2 *Assume Assumptions 1–4 of Section 1.1 hold. Define*

$$W^{(k)} := \bigvee_{t=0}^{\infty} Y^{(k)}(t) = \bigvee_{n=0}^{\infty} S_n^{(k)}$$

$$= \textit{all-time max of } k\textit{th random walk.}$$

Then in $\mathbb{R}$*, we have the convergence in distribution, as* $k \to \infty$*,*

$$W^{(k)} \Rightarrow W^{(\infty)} := \bigvee_{t=0}^{\infty} Y^{(\infty)}(t) = \bigvee_{t=0}^{\infty} \left(\xi^{(\infty)}(t) - t\right),$$

where recall $\xi^{(\infty)}(\cdot)$ *is the zero mean Lévy process of Proposition 1.*

All time max of kth random walk $\Rightarrow$ all time max of zero mean Lévy process with negative drift.

Proof. Use a method from Asmussen (2003) for the finite variance case: For $M > 0$, the map

$$x \mapsto \bigvee_{s=0}^{M} x(s)$$

is continuous from $D[0,\infty) \mapsto \mathbb{R}$ so from Proposition 1 $\Rightarrow$

$$\bigvee_{s=0}^{M} Y^{(k)}(s) \Rightarrow \bigvee_{s=0}^{M} Y^{(\infty)}(s)$$

in $\mathbb{R}$. Get rid of M: The desired result follows from Second Converging Together Theorem if for any $\eta > 0$ that

$$\lim_{T\to\infty} \limsup_{k\to\infty} \mathbb{P}[\bigvee_{j \geq d(k)T} S_j^{(k)} > \eta] = 0. \tag{5}$$

Use reverse MG argument. $\square$

3. Heavy traffic approximation for queues with heavy tailed services.

Construct sequence of Lindley queuing models.
Setup:

- $\{\tau_i^{(k)}, i \geq 1\}$–non-negative iid sequence of service lengths with common distribution $B^{(k)}(x)$; finite mean;
- $\{\sigma_i^{(k)}, i \geq 1\}$–iid sequence of non-negative interarrival times with common distribution $A^{(k)}(x)$; finite mean;
- For each k:
$$\{\tau_i^{(k)}, i \geq 1\} \perp\!\!\!\perp \{\sigma_i^{(k)}, i \geq 1\}.$$
- Traffic intensity for the kth model:
$$\rho^{(k)} = E(\tau_1^{(k)})/E(\sigma_1^{(k)}).$$
- The delay or waiting time process of the kth Lindley queue:
$$W_0^{(k)} = 0, \quad W_{n+1}^{(k)} = (W_n^{(k)} + \tau_n^{(k)} - \sigma_{n+1}^{(k)})^+, \quad n \geq 0.$$

Heavy traffic approximation requires the following conditions.

CONDITION (A).

- $\exists$ df F on $[0, \infty)$ such that
$$\bar{F} := 1 - F \in RV_{-\alpha}, \qquad 1 < \alpha < 2.$$
The quantile function:
$$b(t) = \left(\frac{1}{1-F}\right)^{\leftarrow}(t) \in RV_{1/\alpha}. \tag{6}$$
- Further $B^{(k)}(x)$ tail approximated by F in the sense that
$$\lim_{x\to\infty} \frac{\bar{B}^{(k)}(x)}{\bar{F}(x)} = 1, \tag{7}$$
uniformly in $k = 1, 2, \dots$; ie, given $\delta > 0$, there exists $x_0 = x_0(\delta)$ independent of k such that for $x > x_0$ and all k
$$1 - \delta < \frac{\bar{B}^{(k)}(x)}{\bar{F}(x)} \leq 1 + \delta. \tag{8}$$

CONDITION (B). The tails of the distribution of $\sigma_1^{(k)}$ are always lighter than the tail of F. A convenient way we ensure this is by supposing that there exists $\eta > \alpha$ such that
$$c^{\vee} := \sup_{k\geq 1} E(\sigma_1^{(k)})^{\eta} < \infty. \tag{9}$$

Condition (C). Assume heavier loading as k gets large:

$$0 > m(k) := E(\tau_1^{(k)}) - E(\sigma_1^{(k)}) \to 0, \tag{10}$$

as $k \to \infty$. Set

$$X_i^{(k)} := \frac{\tau_i^{(k)} - \sigma_{i+1}^{(k)}}{b(d(k))}, \quad i \geq 1, \tag{11}$$

where $d(k)$ still to be specified. The step mean $E(X_1^{(k)}$ is

$$\mu^{(k)} = \frac{E(\tau_1^{(k)}) - E(\sigma_1^{(k)})}{b(d(k))} = \frac{m(k)}{b(d(k))}.$$

Interpret (10):

- The kth random walk has negative drift
- $\Rightarrow$ the kth Lindley queue is stable;
- As k increases, the random walk drift becomes more and more negligible
- $\Rightarrow$ the associated Lindley models become less and less stable.
- Hence the need for scaling by $b(d(k))$.

Definition of $d(k)$.

In order for the random walks with negative drift to be approximated by stable Lévy motion with drift -1, the function $d(k)$ must satisfy

$$d(k)\mu^{(k)} := \frac{d(k)m(k)}{b(d(k))} \to -1,$$

as $k \to \infty$. The function

$$H(t) := \frac{t}{b(t)} \in RV_{1-\frac{1}{\alpha}}, \quad 1-\alpha^{-1} > 0 \tag{12}$$

grows like a power function and has an asymptotic inverse

$$H^{\leftarrow} \in RV_{\alpha/(\alpha-1)}.$$

The sequence $d(k)$ must satisfy

$$H(d(k)) \sim \frac{1}{|m(k)|}.$$

Therefore, we choose the sequence $\{d(k)\}$ to be any sequence satisfying

$$d(k) \sim H^{\leftarrow}\left(\frac{1}{|m(k)|}\right), \tag{13}$$

where H is specified in (12).
We now state the approximation theorem.

Theorem 1 *Assume Conditions (A)–(C) hold. With* $\{X_i^{(k)}, i \geq 1\}$ *defined by* (11) *and* $\{d(k)\}$ *satisfying* (13) *we have in* $D[0,\infty)$

$$Y^{(k)}(t) := \sum_{i=1}^{[d(k)t]} X_i^{(k)} = \frac{1}{b(d(k))} \sum_{i=1}^{[d(k)t]} \left(\tau_i^{(k)} - \sigma_{i+1}^{(k)}\right) \Rightarrow Y^{(\infty)}(t), \quad (14)$$

where the limit $Y^{(\infty)}(t) = \xi^{(\infty)}(t) - t$ *and* $\xi^{(\infty)}(t)$ *is a totally skewed to the right, zero mean,* α*-stable Lévy motion.*
Furthermore, the sequence of stationary waiting times indexed by k *converges in distribution in* $\mathbb{R}$*:*

$$\frac{1}{b(d(k))} W^{(k)} = \bigvee_{n=0}^{\infty} \frac{1}{b(d(k))} \sum_{i=1}^{n} \left(\tau_i^{(k)} - \sigma_{i+1}^{(k)}\right) \Rightarrow W^{(\infty)} = \bigvee_{t=0}^{\infty} Y^{(\infty)}(t). \quad (15)$$

Remark

- The distribution of the maximum $W^{(\infty)}$ of a negative drift α-stable Lévy motion computed in Furrer (1997, 1998) using Zolotarev (1964).
- The limit distribution is a Mittag–Leffler distribution.

We have the following corollary.

Corollary 1 *Suppose the assumptions of Theorem 1 hold. Then for every* $t > 0$,

$$P(W^{(k)}/b(d(k)) \le t) \to P(W^{(\infty)} \le t) = 1 - \sum_{n=0}^{\infty} \frac{(-a)^n}{\Gamma(1+n(\alpha-1))} t^{n(\alpha-1)}, \tag{16}$$

where $a = (\alpha-1)/\Gamma(2-\alpha)$, *and for every* $\lambda \ge 0$,

$$Ee^{-\lambda W^{(k)}} \to Ee^{-\lambda W^{(\infty)}} = \frac{a}{a+\lambda^{\alpha-1}}. \tag{17}$$

4. Example

Consider

- GI/G/1 queue with service times $\{\tau_i, i \geq 1\}$ having the Pareto distribution
$$F(x) = 1 - x^{-3/2}, \ x \geq 1\,,$$
- Interarrival times $\{\sigma_i, i \geq 1\}$ having the Gamma(β, λ) distribution.
- Assume
$$\rho = E(\tau_1)/E(\sigma_1) < 1\,,$$

Approximate values of the probabilities that the stationary waiting time in the system exceeds a given level: the scaling setup suggests associating big k with ρ close to 1. With this association

$$\rho \leftrightarrow k$$

we have

- $b(t) = t^{2/3}$ for $t \geq 1$,
- $m(\rho) = -(1-\rho)E(\tau_1) = -2(1-\rho)$.
- The function H in (12) is
$$H(t) = t^{1/3} \text{ for } t \geq 1$$

- Its inverse is

$$H^{\leftarrow}(u) = u^3 \text{ for } u \geq 1\,.$$

- Supposing the relation (13) defining $d(k)$ is an equality so

$$d(k) = \frac{1}{8}(1-\rho)^{-3}\,.$$

With $W^{(\infty)}$ having the Mittag-Leffler distribution (16), our approximation is then

$$P(W > t) \approx P(W^{(\infty)} > t/b(d(k))) = P\Big(W^{(\infty)} > 4(1-\rho)^2 t\Big)\,.$$

Contents

References

S. Asmussen. *Applied Probability and Queues*, volume 51 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. ISBN 0-387-00211-1. Stochastic Modelling and Applied Probability.

O. J. Boxma and J. W. Cohen. Heavy-traffic analysis for the $GI/G/1$ queue with heavy-tailed distributions. *Queueing Systems Theory Appl.*, 33(1-3):177–204, 1999. ISSN 0257-0130. Queues with heavy-tailed distributions.

J. W. Cohen. A heavy-traffic theorem for the $GI/G/1$ queue with a Pareto-type service time distribution. *J. Appl. Math. Stochastic Anal.*, 11(3):247–254, 1998. ISSN 1048-9533.

H. Furrer. *Risk Theory and Heavy-Tailed Lévy Processes*. PhD thesis, ETH, Zurich, 1997. Diss. ETH No. 12408.

H. Furrer. Risk processes perturbed by α-stable Lévy motion. *Scand. Actuar. J.*, 1998(1):59–74, 1998. ISSN 0346-1238.

S.I. Resnick and G. Samorodnitsky. A heavy traffic approximation for workload processes with heavy tailed service requirements. *Management Science*, 46:1236–1248, 2000.

V. M. Zolotarev. The first passage time of a level and the behaviour at infinity of a class of processes with independent increments. *Theory Probab. Appl.*, 9:653–661, 1964.