A Framework for Seeking Hidden Risks; Origins of Power Laws in Networks

Sidney Resnick
School of Operations Research and Information Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA

http://people.orie.cornell.edu/∼sid
sir1@cornell.edu sr2382@columbia.edu

Columbia April MURI

April 25, 2013

1. Part I: Risk Estimation–Basic Problem.

(Joint with Filip Lindskog, Joyjit Roy.)
Given a *risk vector*

$$\boldsymbol{X} = (X_1, \dots, X_d), \qquad d \leq \infty,$$

estimate the probability of a *risk region* $\mathcal{R}$

$$P[\boldsymbol{X} \in \mathcal{R}]$$

where $\mathcal{R}$ is beyond the range of observed data. Solution based on asymptotic assumption of heavy tails: In $M_+(\mathbb{E})$.

$$nP\Big[\frac{\boldsymbol{X}}{b(n)} \in \cdot\Big] \to \nu(\cdot) \qquad \text{(DOA)}$$

Issues:

- What is "$\to$"?
- What is $\mathbb{E}$? $M_+(\mathbb{E})$? Traditionally used one point uncompactification and vague convergence.
- Specify $\nu(\cdot)$.

Quasi-solution to estimation problem:

$$P[\boldsymbol{X} \in \mathcal{R}] \approx \frac{1}{n}\hat{\nu}(\mathcal{R}/\hat{b}(n)).$$

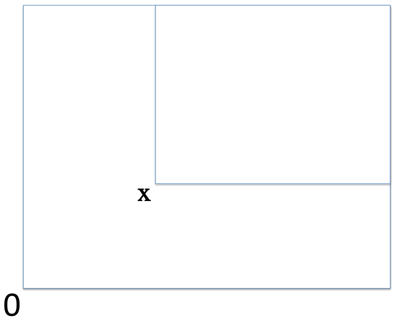
Example:

$d = 2$ and

$\mathcal{R} = (\mathbf{x}, \boldsymbol{\infty}] = (x_1, \infty] \times (x_2, \infty]$

and

$P[\boldsymbol{X} \in \mathcal{R}] = P[X_1 > x_1, X_2 > x_2].$



Risk contagion: Can two or more components of the risk vector $\boldsymbol{X}$ be simultaneously large?

Ambiguity: Should we do the approximation assuming $\mathbb{E} = [0, \infty)^2 \setminus \{\mathbf{0}\}$ or assuming $\mathbb{E} = (0, \infty)^2$.

2. Asymptotic independence

- If $\mathbb{E} = [0, \infty)^d \setminus \{\mathbf{0}\}$, asymptotic independence means

$$\nu(\{\mathbf{x} : x_i > y_i,\ x_j > y_j),\}) = 0, \quad y_i > 0,\, y_j > 0. \quad \text{(AsyIndep)}$$

for all $1 \le i < j \le d$ and thus such an *asymptotic* model has no risk contagion since we estimate

$$P[\text{ two or more components of } \boldsymbol{X} \text{ are large simultaneously }] \approx 0.$$

- Can we improve on this asymptotic method?

2.1. (AsyIndep) not uncommon:

- $\boldsymbol{X}$ has independent components.
- Gaussian copula model: Heavy tailed marginals but Gaussian dependence with
$$\operatorname{corr}(X_i, X_j) = \rho(i, j) < 1.$$
- Let $U \sim U(0, 1)$ and
$$\boldsymbol{X} = \Big(\frac{1}{U}, \frac{1}{1-U}\Big).$$

2.2. Standard Example.

For $d = 2$: If $\boldsymbol{X} = (X_1, X_2)$ and $X_1 \perp\!\!\!\perp X_2$, X_1, X_2 iid with

$$P[X_i > y] = y^{-1}, \quad y > 1.$$

- Then for $x_1 > 0$, $x_2 > 0$, as $n \to \infty$

$$nP[X_i > nx_i] \to x_i^{-1}, \qquad i = 1, 2,$$
$$nP[X_1 > nx_1, X_2 > nx_2] \to 0,$$

so $\boldsymbol{X}$ is regularly varying on $\mathbb{E} = [0, \infty)^2 \setminus \{\mathbf{0}\}$ with index 1 and limit measure concentrating on the axes.

- Also for $x_1 > 0, x_2 > 0$,

$$\sqrt{n}P[X_1 > \sqrt{n}x_1] \cdot \sqrt{n}P[X_2 > \sqrt{n}x_2] =$$
$$nP[X_1 > \sqrt{n}x_1, X_2 > \sqrt{n}x_2] \to \frac{1}{x_1 x_2},$$

so $\boldsymbol{X}$ is regularly varying on $\mathbb{E}_0 = (0, \infty)^2 \setminus \{\mathbf{0}\}$ with index 2 and limit measure giving positive mass to $(\mathbf{x}, \boldsymbol{\infty}]$.

Conclude for this example:

- $\boldsymbol{X}$ is regularly varying on $\mathbb{E} = [\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$ with index 1 (scale by n) and limit measure concentrating on lines through $\{\mathbf{0}\}$, and giving zero mass to $(\mathbf{0}, \boldsymbol{\infty})$.
- $\boldsymbol{X}$ is regularly varying on $\mathbb{E}_0 = (\mathbf{0}, \boldsymbol{\infty})$ with index 2 (scale by $\sqrt{n}$) and the limit measure gives positive mass to $(\mathbf{0}, \boldsymbol{\infty})$.

Summary:

Lesson: If the support (eg, axes) of the limit measure is disjoint from the risk region (eg, $(\mathbf{x}, \boldsymbol{\infty})$)

- peel away the support (axes);
- look for extreme value behavior on what's left (eg, $\mathbb{E} \setminus \{\text{axes}\} = \mathbb{E}_0$).

Some progress:

- Detection.
- Estimation.
- Non-parametric technique using rank methods detects hidden structure.

3. Regular variation on cones in CSMS (with Lindskog & Roy).

Context: Consider random element $\boldsymbol{X}$ of a complete separable metric space $\mathbb{S}$ with an origin and scaling operation: Examples of $\mathbb{S}$:

- $\mathbb{R}_+$; $\boldsymbol{X}$ is a random variable,
- $\mathbb{R}_+^d$; $\boldsymbol{X}$ is a random vector,
- $\mathbb{R}_+^\infty$; $\boldsymbol{X}$ is a random sequence,
- $D[0,1]$=càdlàg space; $\boldsymbol{X}$ is a càdlàg process such as a Lévy process.

Suppose $\mathbb{F}_1 \subset \mathbb{S}$ closed (cone) containing $\mathbf{0}$ and define the TABOF space

$$\mathbb{S}_{\mathbb{F}_1} = \mathbb{S} \setminus \mathbb{F}_1.$$

→ The random element $\boldsymbol{X} \in \mathbb{S}$ has a distribution with a regularly varying tail on $\mathbb{S}_{\mathbb{F}_1}$ if $\exists\, b(t) \uparrow \infty$ and measure $\nu \not\equiv 0$ on $\mathbb{S}_{\mathbb{F}_1}$ such that

$$tP\Big[\frac{\boldsymbol{X}}{b(t)} \in \cdot\,\Big] \to \nu(\cdot), \qquad \text{in } M^*(\mathbb{S}_{\mathbb{F}_1}).$$

[Must define topology on $M^*(\mathbb{S}_{\mathbb{F}_1})$, the measures on $\mathbb{S}_{F_1}$ that are finite on sets at positive distance from F_1; fairly routine.]

Let $\mathbb{F}_2$ be another closed (cone) containing $\mathbf{0}$ and set

$$\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2} = \mathbb{S} \setminus (\mathbb{F}_1 \cup \mathbb{F}_1).$$

→ The random $\boldsymbol{X}$ has a distribution with hidden regular variation on $\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}$ if there is regular variation on $\mathbb{S}_{\mathbb{F}_1}$ AND if $\exists\, b_1(t) \uparrow \infty$ and a measure $\nu_1(\cdot) \not\equiv 0$ on $\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}$ such that

$$tP\Big[\frac{\boldsymbol{X}}{b_1(t)} \in \cdot\,\Big] \to \nu_1(\cdot), \qquad \text{in } M^*(\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}),$$

AND

$$b(t)/b_1(t) \to \infty$$

(which makes the behavior on $\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}$ hidden).

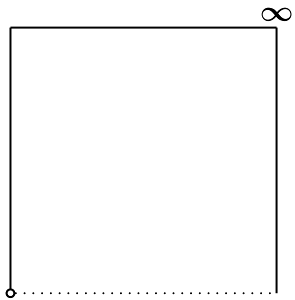
3.0.1. Examples for $d = 2$:

1. Regular variation on the positive quadrant with conditional extreme value (CEV) model:

$$\mathbb{S} = [\mathbf{0}, \infty), \ \mathbb{F}_1 = \{\mathbf{0}\},$$
$$\mathbb{S}_{\mathbb{F}_1} = [\mathbf{0}, \infty) \setminus \{\mathbf{0}\}.$$

CEV on $\mathbb{D}_{\sqcap}$:

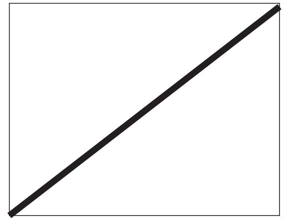
$$\mathbb{F}_2 = \{(x, 0) : x > 0\},$$
$$\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2} = \mathbb{S} \setminus (\mathbb{F}_1 \cup \mathbb{F}_2)$$
$$= [0, \infty) \times (0, \infty) =: \mathbb{D}_{\sqcap}.$$



2. Asymptotic full dependence:

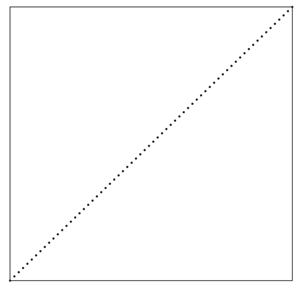
   Regular variation on $[\mathbf{0},\infty)\backslash\{\mathbf{0}\}$ with limit measure concentrating on diagonal.

$$\mathbb{S} = [\mathbf{0},\infty),\ \mathbb{F}_1 = \{\mathbf{0}\},$$
$$\mathbb{S}_{\mathbb{F}_1} = \mathbb{S}\setminus\mathbb{F}_1 = [\mathbf{0},\infty)\setminus\{\mathbf{0}\}.$$



   Remove diagonal:

$$\begin{aligned}\mathbb{F}_2 =& \{(x,x) : x > 0\},\\ \mathbb{S}_{\mathbb{F}_1\cup\mathbb{F}_2} =& \mathbb{S}\setminus(\mathbb{F}_1\cup\mathbb{F}_2)\\ =& [\mathbf{0},\infty)\setminus\{(x,x) : x \geq 0\}\end{aligned}$$



CORNELL
I. Risk Estimation
Asy Indep
RV on Cones; CSMS
Challenges
Visits
II. (in,out)
Help
Title Page
Page 10 of 24
Go Back
Full Screen
Close
Quit

3.0.2. Example: iid on $\mathbb{R}_+^\infty$.

$\mathbb{S} = \mathbb{R}_+^\infty$, $\mathbb{F} = \mathbb{F}^{(j)}$, $j \geq 0$, where

$$\begin{aligned}\mathbb{F}^{(j)} =& \{\mathbf{x} := (x_1, x_2, \dots) \in \mathbb{R}_+^\infty : \sum_{j=1}^\infty \epsilon_{x_j}(0,\infty) \leq j\} \\ =& \{\mathbf{x} : \text{ at most } j \text{ components} > 0\}.\end{aligned}$$

So

$$\begin{aligned}\mathbb{F}^{(0)} =& \{\mathbf{0}_\infty\} \\ \mathbb{F}^{(1)} =& \text{axes in } \mathbb{R}_+^\infty \text{ through } \mathbf{0}, \text{ including } \mathbf{0} \\ =& \bigcup_{j=1}^\infty \{0\}^{j-1} \times (0,\infty) \times \{0\}^\infty \cup \{\mathbf{0}_\infty\}, \\ & \vdots\end{aligned}$$

Leads to a sequence of spaces:

$$\underbrace{\mathbb{R}_+^\infty \setminus \mathbb{F}^{(0)}}_{\text{remove } \mathbf{0}_\infty} \supset \underbrace{\mathbb{R}_+^\infty \setminus \mathbb{F}^{(1)}}_{\text{remove axes}} \supset \underbrace{\mathbb{R}_+^\infty \setminus \mathbb{F}^{(2)}}_{\text{remove 2-dim faces}} \supset \dots \quad \dots .$$

Suppose,

- $\mathbb{S} = \mathbb{R}_+^\infty$;
- $\boldsymbol{X} = (X_1, X_2 \dots)$ has iid components with each X_i having a regularly varying tail with scaling function $b(t)$. Means:

$$tP[X_j > b(t)x] \to \nu_\alpha(x,\infty) = x^{-\alpha}, \quad t \to \infty,\ \alpha > 0,$$

and $b(t)$ satisfies,

$$b(t) = \Big(\frac{1}{P[X_j > \cdot]}\Big)^{-1}(t), \quad P[X_j > b(t)] \sim \frac{1}{t}.$$

Then as $t \to \infty$, for $j \geq 1$

$$tP[\boldsymbol{X}/b(t^{1/j}) \in \cdot] \to \mu^{(j)} \text{ in } \mathbb{M}(\mathbb{R}_+^\infty \setminus \mathbb{F}^{(j-1)})$$

and $\mu^{(j)}$ concentrates on $\mathbb{F}^{(j)} \setminus \mathbb{F}^{(j-1)}$, the sequences with *exactly* j positive components.

The summary for Dr. Spock types.

The limit measure for $j = 1, 2, \ldots$.

j	remove	scaling	$\mu^{(j)}$	support
1	$\{0\}$	$b(t)$	$\sum_{l=1}^{\infty} \nu_\alpha(dx_l)\left[\prod_{i\neq l}\epsilon_0(dx_i)\right]$	axes
2	axes	$b(\sqrt{t})$	$\sum_{l,m} \nu_\alpha(dx_l)\nu_\alpha(dx_m)\left[\prod_{i\notin\{l,m\}} \epsilon_0(dx_i)\right]$	2d faces
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	$\mathbb{F}^{(m-1)}$	$b(t^{\frac{1}{m}})$	$\sum_{(l_1,\ldots,l_m)} \prod_{p=1}^{m} \nu_\alpha(dx_{l_p})\left[\prod_{i\notin\{l_1,\ldots,l_m\}} \epsilon_0(dx_i)\right]$	$\mathbb{F}^{(m)} \setminus \mathbb{F}^{(m-1)}$

3.0.3. Example continued: CUMSUM on $\mathbb{R}_+^\infty$.

Define

$$\text{CUMSUM} : \mathbb{R}_+^\infty \mapsto \mathbb{R}_+^\infty$$

by

$$\text{CUMSUM}(\mathbf{x}) = (x_1, x_1 + x_2, \dots)$$

and then, because CUMSUM is *uniformly continuous*,

$$tP[\text{CUMSUM}(\boldsymbol{X})/b(t^{1/j}) \in \cdot] \to \mu^{(j)} \circ \text{CUMSUM}^{-1}$$

in $\mathbb{M}\big(\mathbb{R}_+^\infty \setminus \text{CUMSUM}(\mathbb{F}^{(j-1)})\big)$ and $\mu^{(j)} \circ \text{CUMSUM}^{-1}$ concentrates on $\text{CUMSUM}(\mathbb{F}^{(j)} \setminus \mathbb{F}^{(j-1)}))$, the set of sequences with j positive jumps. Applying a projection

$$\mathbb{R}_+^\infty \mapsto \mathbb{R}_+; \quad \mathbf{x} \mapsto x_k,$$

to the $j = 1$ case, yields the *one big jump principle*:

$$P[X_1 + \cdots + X_k > b(t)x] \to kx^{-\alpha}, \quad x > 0.$$

or

$$P[X_1 + \cdots + X_k > t] \sim kP[X_1 > t], \quad t \to \infty.$$

3.0.4. Extension to Lévy processes

- $\{X(t),\, t \geq 0\}$: Lévy process with Lévy measure $\nu(\cdot)$.
- As $x \to \infty$,
$$\nu(x,\infty) \sim x^{-\alpha} L(x).$$
- $\nu_\alpha(x,\infty) = x^{-\alpha}$, $x > 0$.
- For $m \geq 0$, $\mathbb{D}_{\leq m} \subset \mathbb{D}([0,1])$ are nondecreasing step functions with at most m jumps.
- $\{U_1, U_2, \dots\}$ are iid $U(0,1)$.

Then for $j \geq 1$,

$$nP\Big[\frac{X(\cdot)}{b(n^{1/j})} \in \cdot\,\Big] \to \mathbb{E}\Big[\nu_\alpha^j\Big\{\boldsymbol{y} = (y_1,\dots,y_j) \in (0,\infty)^j : \sum_{i=1}^{j} y_i 1_{[U_i,1]} \in \cdot\Big\}\Big]$$

in $M^*(\mathbb{D}[0,1] \setminus \mathbb{D}_{\leq j-1})$ as $n \to \infty$.

For $j = 1$ recover Hult/Lindskog

$$nP\Big[\frac{X(\cdot)}{b(n)} \in \cdot\,\Big] \to \mathbb{E}\Big[\nu_\alpha\Big\{y \in (0,\infty)^j : y 1_{[U_1,1]} \in \cdot\Big\}\Big],$$

where the limit measure concentrates on functions with one upward jump occurring at uniform time.

4. Challenges.

- Practical?
  - Limitations of asymptotic methods: rates of convergence?
  - Instead of estimating a risk probability as 0, estimate is a very small number.
- Need for more formal inference for estimation including confidence statements.
- General HRV technique requires knowing the support of the limit measure. Estimate support?
- High dimension problems? How to sift through different possible subcones? There could be a sequence of cones with regular variation on each. How to teach a computer to find the cones?
- How to go from standard to more realistic non-standard case; still some inference problems.

5. Visits (that I know about) post Natick

- January 8: Don $\to$ Sid at Columbia.
- Febrary 5: Sid & Richard $\to$ Don & Weibo & group at UMass.
- March 22: Sid $\to$ John Nolan at AmericanU.
- April 8: Bo Jiang (UMass) $\to$ Sid & Richard at Columbia.
- Conference call: John Nolan & Gena with Edan Ben-Ari & Mech Eng re ballistics.
- May 6: John Lavery $\to$ John Nolan at AmericanU.

> "*This visit is a site visit for informal discussion (not an official review of the MURI). Nevertheless, the visit is something that the MURI team can mention in its next interim progress report. Site visits are not large items to report but are positive and it is worthwhile to record this small item in the interim progress report.*"

6. Properties of joint distribution of (in-degree, out-degree) of nodes in preferential attachment models

Discussions with Don Towsley, Bo Jiang, Richard Davis.

- Preferential attachment growth model of directed graph.
  - V_n = set of vertices at time (stage, event) n.
  - E_n = set of edges.
  - in(v,n) = in-degree of node $v \in V_n$ at time n.
  - out(v,n) = out-degree of node $v \in V_n$ at time n.
  - $\mathcal{F}_n$ = known information from observing growth up through time n.
  - $N_{i,j}(n)$ = # number of nodes $v \in V_n$ such that

$$\big(\text{in}(v,n), \text{out}(v,n)\big) = (i,j).$$

- Abstract the growth dynamic for Krapivsky or reciprocity model or similar models.

  Either:

  - With prob p new node $v \notin V_n$ appears and attaches to $u \in V_n$ with prob $p_u^{(n+1)} \in \mathcal{F}_n$;

or something else happens (eg, in Krapivsky):

– With prob $q = 1 - p$ no new node appears but a new directed edge $e \notin E_n$ appears from $u \in V_n \to v \in V_n$ with prob $\tilde{p}_{u,v}^{(n+1)} \in \mathcal{F}_n$.

For example: in Krapivsky, for a parameters $\lambda \geq 0$, $\mu \geq 0$,

$$
\begin{aligned}
p_u^{(n+1)} &= \frac{\mathrm{in}(u,n) + \lambda}{\sum_{u' \in V_n} \mathrm{in}(u',n) + \lambda} && \in \mathcal{F}_n, \\
\tilde{p}_{u,v}^{(n+1)} &= \frac{(\mathrm{out}(u,n) + \mu)(\mathrm{in}(v,n) + \lambda)}{\sum_{u',v' \in V_n} (\mathrm{out}(u',n) + \mu)(\mathrm{in}(v',n) + \lambda)} && \in \mathcal{F}_n.
\end{aligned}
$$

- Summary statistic (Richard): Point measure,

$$
\mathrm{CT}_n := \sum_{v \in V_n} \epsilon_{(\mathrm{in}(v,n),\mathrm{out}(v,n))} = \sum_{(i,j)} N_{i,j}(n) \epsilon_{(i,j)}
$$

7. Don't know!

- What is the origin of heavy tails? Presumably, as $n \to \infty$,

$$\frac{N_{i,j}(n)}{n} \to f_{i,j}$$
$$\frac{\mathrm{CT}_n}{n} \to \mathrm{CT}_\infty =: \sum_{(i,j)} f_{i,j}\epsilon_{(i,j)}.$$

  - Find a sexy martingale argument to prove the limits exist.
  - Find broad circumstances under which $f_{i,j}$ exhibit power law behavior.
  - Find good definition of *power law behavior*. Understand how this relates to current formulation of multivariate heavy tails. In what space do the power laws exist; see first part of the talk.
  - In one dimension, a regularly varying sequence $\{c_n\}$ means

$$\lim_{n\to\infty} \frac{c_{[nt]}}{c_n} = g(t), \quad t > 0$$

    and this implies there exists a regularly varying function of a continuous variable $c(t)$ such that

$$c_n = c(n).$$

Is there a comparable result in dimensions bigger than one?

- In one dimension, *Karamata's theorem* says the indefinite integral of a regularly varying function is still regularly varying. This is known to sometimes fail in higher dimensions. If we generate a measure from $\{f_{i,j}\}$, is it regularly varying?
- (Richard) Double limit vs single limit: The argument for power law behavior in preferential attachment models lets $n \to \infty$ to get $f_{i,j}$ and then effectively takes another limit in i, j to observe power law behavior. Does there exist $a_n, b_n \to \infty$ such that
$$\sum_{v \in V_n} \epsilon_{(\mathrm{in}(v,n)/a_n, \mathrm{out}(v,n)/b_n)}$$
converges to something identifiable as heavy tailed. Does this single limit relate nicely to existing theory?
- MLE or other forms of estimation: Based on what data? How to sample? (Don? Zhi-Li?)

Contents

References

B. Das, A. Mitra, and S. Resnick. Living on the multi-dimensional edge: Seeking hidden risks using regular variation. *ArXiv e-prints 1108.5560*, August 2011.

G. Draisma, H. Drees, A. Ferreira, and L. de Haan. Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli*, 10 (2):251–280, 2004.

J.E. Heffernan and S.I. Resnick. Hidden regular variation and the rank transform. *Adv. Appl. Prob.*, 37(2):393–414, 2005.

H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)*, 80(94):121–140, 2006. ISSN 0350-1302. doi: 10.2298/PIM0694121H. URL http://dx.doi.org/10.2298/PIM0694121H.

A.W. Ledford and J.A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187, 1996. ISSN 0006-3444.

A.W. Ledford and J.A. Tawn. Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B*, 59(2):475–499, 1997. ISSN 0035-9246.

K. Maulik and S.I. Resnick. Characterizations and examples of hidden regular variation. *Extremes*, 7(1):31–67, 2005.

A. Mitra and S.I. Resnick. Modeling multiple risks: Hidden domain of attraction. Technical report, Cornell University, 2011a. URL http://arxiv.org/abs/1110.0561.

A. Mitra and S.I. Resnick. Hidden regular variation and detection of hidden risks. *Stochastic Models*, 27(4):591–614, 2011b.

S.I. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4):303–336 (2003), 2002. ISSN 1386-1999.

M. Sibuya. Bivariate extreme statistics. *Ann. Inst. Stat. Math.*, 11: 195–210, 1960.