Multivariate Heavy Tails, Asymptotic Independence and Beyond

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

http://www.orie.cornell.edu/~sid
sir1@cornell.edu

April 21, 2005

Work with: K. Maulik, J. Heffernan, S. Marron, ...
1. Multidimensional Heavy Tails.

Consider a vector $\mathbf{X} = (X^{(1)}, \ldots, X^{(d)})$ where

- The components may be dependent.
- The components are each univariate heavy tailed.

Big issue: How to model the dependence?

- The tail indices ($\alpha$’s) for each component are typically different in practice.
- Parametric (use MLE) vs semi-parametric (use asymptotic theory).
  - Parametric will fail goodness of fit with large data sets.
  - Semi-parametric will have difficult asymptotic theory.
- Stable and max-stable distributions indexed by measures on the unit sphere—large classes and why should even the marginals be correct? Parametric sub-families may be ad hoc.
- Copula methods.
1.1. Example.

Internet traffic:
Consider

\[ F = \text{file size,} \]
\[ L = \text{duration of transmission,} \]
\[ R = \text{throughput} = F/L. \]

All three, are seen empirically to be heavy tailed.

Two studies:
- BU
- UNC

What is the dependence structure of \((F, R, L)\)?
Since \(F = LR\), the tail parameters \((\alpha_F, \alpha_R, \alpha_L)\) cannot be arbitrary.
Note for BU measurements, we have the following empirical estimates:

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$\hat{\alpha}_F$</th>
<th>$\hat{\alpha}_R$</th>
<th>$\hat{\alpha}_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>estimated value</td>
<td></td>
<td>1.15</td>
<td>1.13</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Two theoretical possibilities:

- If $(L, R)$ have a joint distribution with multivariate regularly varying tail but are NOT asymptotically independent then (Maulik, Resnick, Rootzen (2002))

  $$\hat{\alpha}_F = \frac{\hat{\alpha}_L \hat{\alpha}_R}{\hat{\alpha}_L + \hat{\alpha}_R} = .625 \neq 1.15.$$ 

- If $(L, R)$ obey a form (not the EVT version) of asymptotic independence, (Maulik+Resnick+Rootzen; Heffernan+Resnick)

  $$tP[(L, \frac{R}{b(t)}) \in \cdot] \overset{v}{\to} G \times x^{-\alpha-1}dx$$

  then

  $$\alpha_F = \alpha_R \land \alpha_L$$

  and in our example

  $$1.15 \approx 1.13 \land 1.4.$$
For two examples

- BU: Evidence seems to support some form of independence for $\{R, L\}$.

- UNC: Conclusions from Campos, Marron, Resnick, Jeffay (2005);
  - Large values of $F$ tend to be independent of large values of $R$.

  $\Rightarrow$ Large files do not seem to receive any special consideration when rates are assigned.
BuL vs BuR:

Data processed from the original 1995 Boston University data; 4161 file sizes (F) and download times (L) noted and transmission rates (R) inferred. The data consists of bivariate pairs (R,L).
2. Multivariate Regular Variation.

2.1. Standard Case

A fct $U : \mathbb{R}_+^d \mapsto \mathbb{R}_+$ is mult reg varying if

$$\frac{U(tx)}{U(t1)} \to \lambda(x) \neq 0,$$

for $x \geq 0$, $x \neq 0$. Then $\exists \rho$ and

$$\lambda(tx) = t^\rho \lambda(x),$$

and $U(t1) \in RV_\rho$.

Usually there is a sequential equivalent version: $\exists b_n \to \infty$ such that

$$\frac{U(b_n x)}{n} \to \lambda(x).$$
Application to distributions: For simplicity, let \( Z, Z_n, n \geq 1 \) be iid, range=\( \mathbb{R}^d_+ \) and common df \( F \). A regularly varying tail means

\[
\frac{1 - F(tx)}{1 - F(t1)} \rightarrow \nu([0, x]^c),
\]

for some Radon measure \( \nu \). However, it is awkward to deal with mult df’s and better to deal with measures.

Let

\[
\mathbb{E} = [0, \infty]^d \setminus \{0\},
\]

\[
\mathbb{N} = \{x \in \mathbb{E} : \|x\| = 1\},
\]

\[
R = \|Z\|, \quad \Theta = \frac{Z}{\|Z\|} \in \mathbb{N}.
\]

The following are equivalent and define multivariate heavy tails or regularly varying tails.
1. ∃ a Radon measure ν on E such that
\[ \lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t1)} = \lim_{t \to \infty} \frac{\mathbb{P}[\frac{Z_1}{t} \in [0, x]^c]}{\mathbb{P}[\frac{Z_1}{t} \in [0, 1]^c]} = c\nu([0, x]^c), \]
where some \( c > 0 \) and for all points \( x \in [0, \infty) \setminus \{0\} \) which are continuity points of \( \nu([0, \cdot]^c) \).

2. ∃ a function \( b(t) \to \infty \) and a Radon measure ν on E such that in \( M_+(E) \)
\[ t\mathbb{P}[\frac{Z_1}{b(t)} \in \cdot] \xrightarrow{v} \nu, \quad t \to \infty. \]

3. ∃ a pm \( S(\cdot) \) on \( \mathbb{N} \) and \( b(t) \to \infty \) such that
\[ t\mathbb{P}[\left(\frac{R_1}{b(t)}, \Theta_1\right) \in \cdot] \xrightarrow{v} c\nu_\alpha \times S \]
in \( M_+((0, \infty] \times \mathbb{N}) \), where \( c > 0 \) and
\[ \nu_\alpha(x, \infty] = x^{-\alpha}. \]

4. ∃ \( b_n \to \infty \) such that in \( M_p(E) \)
\[ \sum_{i=1}^{n} \epsilon_i Z_i/b_n \Rightarrow \text{PRM}(\nu). \]
5. \( \exists \) a sequence \( b_n \rightarrow \infty \) such that in \( \mathbb{M}_p((0, \infty] \times \mathbb{N}) \)

\[
\sum_{i=1}^{n} \epsilon_{(R_i/b_n, \Theta_i)} \Rightarrow \text{PRM}(c\nu \times S).
\]

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

6. In \( \mathbb{M}_+(\mathbb{E}) \),

\[
\frac{1}{k} \sum_{i=1}^{n} \epsilon_{Z_i/b(n/k)} \Rightarrow \nu \quad (*)
\]

\[
\frac{1}{k} \sum_{i=1}^{n} \epsilon_{(R_i/b(n/k), \Theta_i)} \Rightarrow (c\nu \times S). \quad (**)
\]

and (6) is equivalent to any of (1)–(5), provided \( k(\cdot) \) satisfies \( k(n) \sim k(n + 1) \).

Ignore fact \( b(\cdot) \) unknown:
→ LHS of Eqn (*) is a consistent estimator of \( \nu \).
→ From (**), consistent estimator of \( S \) is

\[
\frac{\sum_{i=1}^{n} \epsilon_{(R_i/b(n/k), \Theta_i)}[1, \infty] \times \cdot}{\sum_{i=1}^{n} \epsilon_{R_i/b(n/k)}[1, \infty]}.
\]
But:

- This theoretical formulation is for the *standard case*.
  - Problematic for applications. If we norm each component by the same $b(t) \Rightarrow$ marginal tails same; ie components on the same scale:
    \[ \mathbb{P}[Z^{(i)} > x] \sim c_{ij}\mathbb{P}[Z^{(j)} > x], \quad c_{ij} > 0, \quad x \to \infty. \]
  - Standard case almost never happens in practice.

- How to transform to the standard case in practice?
  - Simple minded: Hope $1 - F^{(i)}(x) \sim x^{-\alpha_i}$ for all $i$ and then power up. **BUT:** Must estimate $\alpha$’s. YECH!
  - Use ranks method (Huang, 1992; de Haan & de Ronde). **BUT:** Lose independence among observations.
The ranks method:

Given $d$-dimensional random vectors \( \{ \mathbf{X}_1, \ldots, \mathbf{X}_n \} \) where

\[
\mathbf{X}_i = (X_i^{(1)}, \ldots, X_i^{(d)}), \quad i = 1, \ldots, n,
\]
define the (anti)-ranks for each component: Comparing the $j$th components, $X_1^{(j)}, \ldots, X_n^{(j)}$, the anti-rank of $X_i^{(j)}$ is

\[
r_i^{(j)} = \sum_{l=1}^{n} 1 \left[ X_l^{(j)} \geq X_i^{(j)} \right] = \# j\text{th components} \geq X_i^{(j)}.
\]

Replace each $\mathbf{X}_i$ by

\[
\mathbf{X}_i \mapsto (1/r_i^{(j)}, \; j = 1, \ldots, d).
\]
Rank method–UNC

Steps:

- Transform (F,R) data using rank method.
- Convert to polar coordinates.
- Keep 2000 pairs with biggest radius vector.
- Compute density estimate for angular measure $S$.

Plot: Density estimates with various amounts of smoothing+jitter plot (green) of angles.

Full disclosure: These types of plots can be rather sensitive to choice of threshold.
Heavy Tails—≥ 2 dim
Mult Reg Var
Asymptotic Indep
Hidden reg var (HRV)
Characterize HRV
Detecting HRV
Conditional models
Medical Care

EDM = 0.23155

Angle, \( \theta \cdot (2 / \pi) \)
2.2. Simplifying assumptions

For theory, proceed assuming

- Standard case.
- One dimensional marginals $F_{(i)}$, $i = 1, \ldots, d$ are the same.
- $d = 2$ (just for ease of explanation).
3. **Significance of limit measure**

The limit measure $\nu$ controls the (asymptotic) dependence structure:

The distribution $F$ of $Z_1$ possesses *asymptotic independence* if either

1. $\nu((0, \infty)) = 0$ so that $\nu$ concentrates on the axes;

    \textbf{OR}

2. $S$ concentrates on $\{(1, 0), (0, 1)\}$.

This definition designed to yield

- As $n \to \infty$

  \[
  \sqrt[n]{\prod_{i=1}^{n} \frac{Z_i}{b_n}} \Rightarrow (Y^{(1)}, Y^{(2)}),
  \]

  where $(Y^{(1)}, Y^{(2)})$, are independent Frechet distributed.

- Probability of 2 components being simultaneously large is negligible: For $d = 2$:

  \[
  \lim_{t \to \infty} \mathbb{P}[Z^{(2)} > t | Z^{(1)} > t] \to 0.
  \]
3.1. Why asymptotic independence creates problems.

- Estimators of various parameters may behave badly under asymptotic independence; eg, estimator of the spectral measure $S$. Estimators may be asymptotically normal with an asymptotic variance of 0 (oops!).

- Estimators of probabilities given by asymptotic theory may be uninformative.
Scenario: Estimate the probability of simultaneous non-compliance.

Suppppose $Z = (Z^{(1)}, Z^{(2)}) = \text{concentrations of different pollutants}$. Environmental agencies set critical levels $t_0 = (t_0^{(1)}, t_0^{(2)})$ which not be exceeded. Imagine simultaneous non-compliance creates a health hazard. Worry about

$$[\text{health hazard}] = [Z > t_0] = [Z^{(j)} > t_0^{(j)}; \ j = 1, 2].$$

Assume only regular variation with unequal components. Then for the probability of non-compliance, we estimate

$$P[Z^{(1)} > t_0^{(1)}, Z^{(2)} > t_0^{(2)}] = P\left[\frac{Z^{(j)}}{b(j)(\frac{n}{k})} > \frac{t_0^{(j)}}{b(j)(\frac{n}{k})}; \ j = 1, 2\right]$$

$$\approx \frac{k}{n} \nu \left( \left( \left( \frac{t_0^{(1)}}{b^{(1)}(\frac{n}{k})}, \frac{t_0^{(2)}}{b^{(2)}(\frac{n}{k})} \right), \infty \right) \right) = 0$$

since $\nu$ has empty interior by asymptotic independence.

This is not helpful!!
4. Hidden Regular Variation.

A submodel of asymptotic independence.

The random vector $Z$ has a distribution possessing hidden regular variation if

1. Regular variation on the big cone $E = [0, \infty]^2 \setminus \{0\}$:

$$t \mathbb{P}\left[\frac{Z}{b(t)} \in \cdot \right] \overset{v}{\to} \nu,$$

AND

2. Regular variation on the small cone $(0, \infty]^2$: $\exists$ a non-decreasing function $b^*(t) \uparrow \infty$ such that

$$b(t)/b^*(t) \to \infty$$

and $\exists$ a measure $\nu^* \neq 0$ which is Radon on $E^0 = (0, \infty]^2$ and such

that

$$t \mathbb{P}\left[\frac{Z}{b^*(t)} \in \cdot \right] \overset{v}{\to} \nu^* = \text{hidden measure}$$

on the cone $E^0$.

Then there exists $\alpha^* \geq \alpha$ such that $b^* \in RV_{1/\alpha^*}$. 
Consequences:

- With the right formulation,
  
  Second order regular variation + asy indep
  \[ \Rightarrow \text{hidden regular variation} \]
  \[ \Rightarrow \text{asymptotic independence}. \]

- Means for every \( s \geq 0, s \neq 0, \bigvee_{i=1}^{d} s(i)Z(i) \) has distribution with a regularly varying tail of index \( \alpha \) and for every \( a \geq 0, a \neq 0, \bigwedge_{i=1}^{d} a(i)Z(i) \) has a regularly varying distribution tail of index \( \alpha^* \).

- In particular, hidden regular variation means both \( Z^{(1)} \lor Z^{(2)} \) and \( Z^{(1)} \land Z^{(2)} \) have regularly varying tail probabilities with indices \( \alpha \) and \( \alpha^* \). Note
  \[ \eta = 1/\alpha^* = \text{coefficient of tail dependence} \]
  (Ledford and Tawn (1996,1997)).

- Define on \( \mathbb{R} \cap \mathbb{E}^0 \)
  
  \[ S^*(\Lambda) = \nu^*\{x \in \mathbb{E}^0 : |x| \geq 1, \frac{x}{|x|} \in \Lambda\} \]

  called the hidden angular measure.
Sub-model (cont)—Two Examples:

Example 1: $d = 2$; independent random quantities $B, Y, U$ with

$$P[B = 0] = P[B = 1] = 1/2$$

and $Y = (Y^{(1)}, Y^{(1)})$ is iid with

$$P[Y^{(1)} > x] \in RV_{-1}$$

and

$$b(t) = \left(\frac{1}{P[Y^{(1)} > \cdot]}\right)(t) \in RV_1.$$

Let $U$ have multivariate regularly varying distribution on $E$ and

$\exists \alpha^* > 1, b^*(t) \in RV_{1/\alpha^*}, \nu^* \neq 0,$

$$tP\left[\frac{U}{b^*(t)} \in \cdot\right] \to \nu^* \neq 0.$$

Define

$$Z = BY + (1 - B)U$$

which has hidden regular variation, and the property

$$S^*(\mathbb{R}^0) := \nu^*\{x \in \mathbb{R}^0 : \|x\| > 1\} < \infty.$$

Example 2: \(d = 2\), define
\[
\nu^*([x, \infty]) = (x^{(1)}x^{(2)})^{-1}.
\]
Define \(Z = (Z^{(1)}, Z^{(2)})\) iid and Pareto distributed with
\[
P[Z^{(i)} > x] = x^{-1}, \quad x > 1, \quad i = 1, 2.
\]
Set
\[
b(t) = t, \quad b^*(t) = \sqrt{t},
\]
so that \(b(t)/b^*(t) \rightarrow \infty\). Then on \(\mathbb{E}\)
\[
tP\left[\frac{Z}{b(t)} \in \cdot \right] \rightarrow^v \nu,
\]
\(\nu(\mathbb{E}^0) = 0\), and on \(\mathbb{E}^0\)
\[
tP\left[\frac{Z}{b^*(t)} \in \cdot \right] \rightarrow^v \nu^*,
\]
and
\[
S^*(\mathbb{N}^0) := \nu^*\{x \in \mathbb{E}^0 : \|x\| > 1\} = \infty.
\]
How dense are these 2 examples?

Need for a concept of multivariate tail equivalence: Suppose

\[ 0 \leq Y \sim F; \quad 0 \leq Z \sim G. \]

Say \( F, G \) (or \( Y \) and \( Z \)) are tail equivalent on cone \( \mathcal{C} \) if there exists \( b(t) \uparrow \infty \) such that

\[
tP \left[ \frac{Y}{b(t)} \in \cdot \right] = tF(b(t)\cdot) \xrightarrow{\nu} \nu
\]

and

\[
tP \left[ \frac{Z}{b(t)} \in \cdot \right] = tG(b(t)\cdot) \xrightarrow{cv} c\nu
\]

for \( c > 0 \), Radon \( \nu \neq 0 \) on \( \mathcal{C} \).

Write

\[ Y^{te(\mathcal{C})} \sim Z. \]
5. Characterizations.

Mixture Characterization; $S^*$ is Finite

Assume finite hidden angular measure: Sppse $Z \sim F$ is multivariate regularly varying on

$$E := [0, \infty]^d \setminus \{0\}, \quad \text{scaling } b(t),$$

$$E_0 := (0, \infty]^d, \quad \text{scaling } b^*(t), \quad b(t)/b^*(t) \to \infty,$$

$$b \in RV_{1/\alpha}, \quad b^* \in RV_{1/\alpha^*}, \quad \alpha \leq \alpha^*.$$

Then $F$ is tail equivalent on both the cones $E$ and $E_0$ to a mixture distribution

$$Z \overset{te(C)}{\sim} 1_{[I=0]} V + \sum_{i=1}^{d} 1_{[I=i]} X_i e_i.$$

Here $e_i; i = 1, \ldots, d$ are the usual basis vectors.
Remarks on the characterization:

\[ Z \overset{te(\mathbb{C})}{\sim} 1_{[I=0]} V + \sum_{i=1}^{d} 1_{[I=i]} X_i e_i. \]

- \( \sum_{i=1}^{d} 1_{[I=i]} X_i e_i \) concentrates on the axes, has no hidden regular variation, and the marginal distributions (of the \( X_i \)) have scaling function \( b(t) \),

- \( V \) mult reg varying on \( \mathbb{E} \) (not \( \mathbb{E}_0 \)–this is the effect of finite \( \nu^* \)) with scaling function \( b^*(t) \); tails of \( V \) are lighter than those of the completely asymptotically independent distribution \( \sum_{i=1}^{d} 1_{[I=i]} X_i e_i \).

- Conversely: if \( F \) tail equivalent to a mixture as above, \( b(t)/b^*(t) \to \infty \), then \( F \) is multivariate reg varying on \( \mathbb{E} \) and \( \mathbb{E}_0 \) with finite hidden angular measure and with scaling functions \( b, b^* \).
Mixture Characterization; $S^*$ is Infinite

Assume infinite hidden angular measure. Suppose $Z \sim F$ mult regularly varying on

$$E := [0, \infty]^d \setminus \{0\}, \quad \text{scaling } b(t),$$
$$E_0 := (0, \infty]^d, \quad \text{scaling } b^*(t), \quad b(t)/b^*(t) \to \infty,$$
$$b \in RV_{1/\alpha}, \quad b^* \in RV_{1/\alpha^*}, \quad \alpha \leq \alpha^*.$$

Then $F$ is tail equivalent on both the cones $E$ and $E_0$ to a mixture distribution

$$Z = 1_{[I=0]} V + \sum_{i=1}^{d} 1_{[I=i]} X_i e_i.$$

Remarks and notes on the infinite case:

- $V$ is only guaranteed to be reg varying on $E_0$; index is $\alpha^*$.
- If the reg variation of $V$ can be extended to $E$, then the 1-dim marginals have heavier tails of index $\leq \alpha^*$.
- BUT: do not have a useful criterion for when reg var on $E_0$ can be extended to $E$. 
6. Can We Detect Hidden Regular Variation?

Example 1: Simulation.

5000 pairs of iid Pareto, $\alpha = 1; \alpha_* = 2$. Hillplot for rank transformed data taking minima of components.
Example 2: UNC Wed (F,R).

QQ plot of rank transformed data using 1000 upper order statistics for UNC Wed (F,R); \( \alpha = 1 \) and \( \hat{\alpha}_* = 1.6 \).
6.1. Estimating $\nu^*$. 

The hidden measure $\nu^*$ has a spectral measure $S^*$ defined on $\mathbb{R}_0$, the unit sphere in $E_0$:

$$S^*(\Lambda) := \nu^* \{ x \in E_0 : \| x \| > 1, \frac{x}{\| x \|} \in \Lambda \}.$$

$S^*$ may not necessarily be finite.

We estimate $S^*$ rather than $\nu^*$. 

Estimation procedure (Heffernan & Resnick) for estimating $\nu^*$:

1. Replace the heavy tailed multivariate sample $Z_1, \ldots, Z_n$ by the $n$ vectors of reciprocals of anti-ranks $1/r_1, \ldots, 1/r_n$, where

$$r_i^{(j)} = \sum_{l=1}^{n} 1_{\{Z_l^{(j)} \geq Z_i^{(j)}\}}; \quad j = 1, \ldots, d; \quad i = 1, \ldots, n.$$ 

2. Compute normalizing factors

$$m_i = \bigwedge_{j=1}^{d} \frac{1}{r_i^{(j)}}; \quad i = 1, \ldots, n,$$

and their order statistics

$$m(1) \geq \cdots \geq m(n).$$

3. Compute the polar coordinates $\{(R_i, \Theta_i); i = 1, \ldots, n\}$ of $\{(1/r_i^{(j)}; j = 1, \ldots, d); \quad i = 1, \ldots, n\}$.

4. Estimate $S^*$ using the $\Theta_i$ corresponding to $R_i \geq m(k)$.
Details:

• If $\nu^*$ is infinite, let $\mathcal{N}_0(K)$ be compact subset of $\mathcal{N}_0$.
  
  – For $d = 2$ where $\mathcal{N}$ can be parameterized as $\mathcal{N} = [0, \pi/2]$ and $\mathcal{N}_0 = (0, \pi/2)$, set $\mathcal{N}_0(K) = [\delta, \pi/2 - \delta]$ for some small $\delta > 0$.

• Then
  
  \[
  \frac{\sum_{i=1}^{n} 1[R_i \geq m(k), \Theta_i \in \mathcal{N}_0(K)] \epsilon \Theta_i}{\sum_{i=1}^{n} 1[R_i \geq m(k), \Theta_i \in \mathcal{N}_0(K)]} \Rightarrow S_0(\cdot \cap \mathcal{N}_0(K)).
  \]

• If $\nu^*$ is finite, we can replace $\mathcal{N}_0(K)$ with $\mathcal{N}_0$. 
Example.

UNC (F,R), April 26. Asymptotic independence present. Since $S^*$ may be infinite, we restricted estimation to the angular interval interval $[0.1,0.9]$ instead of all of $[0,1]$. All plots show the hidden measure to be bimodal with peaks around 0.2 and 0.85.
7. **Conditional models.**

Other form of asymptotic independence (Maulik, Resnick, Rootzen):

\[ nP\left[ \left( \frac{X}{b(n)}, \frac{Y}{b(n)} \right) \in \cdot \right] \xrightarrow{v} G \times \nu_\alpha \]  \hspace{1cm} (1)

on \([0, \infty] \times (0, \infty]\) where \(G\) is a pm on \([0, \infty)\) and

\[ \nu_\alpha(x, \infty) = x^{-\alpha}, \quad x > 0. \]

Equivalent: \(Y\) has a regularly varying tail and

\[ P[X \leq x | Y = t] \xrightarrow{t \to \infty} G(x). \]

Heffernan & Tawn models:

\[ P\left[ \frac{X - \beta(t)}{\alpha(t)} \leq x | Y = t \right] \xrightarrow{t \to \infty} G(x). \]

With Jan Heffernan: Meld 2 approaches. Reformulate as

\[ tP\left[ \left( \frac{X - \beta(t)}{\alpha(t)}, \frac{Y - b(t)}{a(t)} \right) \in \cdot \right] \xrightarrow{v} \mu \]

where \(\mu\) satisfies non-degeneracy assumptions.
7.1. Basic Convergence

Assume 2 dimensions and

\[ tP \left[ \left( \frac{X - \beta(t)}{\alpha(t)}, \frac{Y - b(t)}{a(t)} \right) \in \cdot \right] \xrightarrow{v} \mu(\cdot), \quad (2) \]

in \( M_+([-\infty, \infty] \times (-\infty, \infty)) \), and non-degeneracy assumptions:

1. for each fixed \( y \), \( \mu((-\infty, x] \times (y, \infty]) \) is not a degenerate distribution function in \( x \);

2. for each fixed \( x \), \( \mu((-\infty, x] \times (y, \infty]) \) is not a degenerate distribution function in \( y \).

Observations:

- The Basic Convergence (2) implies

\[ tP \left[ \frac{Y - b(t)}{a(t)} \in \cdot \right] \xrightarrow{v} \mu([-\infty, \infty] \times (\cdot)), \]

so \( P[Y \in \cdot] \in D(G_{\gamma}) \), for some \( \gamma \in \mathbb{R} \).

- The Basic Convergence (2) implies the conditioned limit

\[ tP \left[ \frac{X - \beta(t)}{\alpha(t)} \leq x \mid Y > b(t) \right] \rightarrow \mu([-\infty, x] \times (0, \infty]). \]
• WLOG can assume $Y$ is heavy tailed and reduce the basic convergence to standard form:

$$tP\left[\left(\frac{X - \beta(t)}{\alpha(t)}, \frac{Y}{t}\right) \in \cdot \right] \rightarrow \mu$$

in $M_+([-\infty, \infty] \times (0, \infty])$ (with a modified $\mu$).

• Suppose $(X, Y)$ are regularly varying on $[0, \infty]^2 \setminus \{0\}$.
  - With no asymptotic independence, Basic Convergence automatically holds.
  - With asymptotic independence, Basic Convergence is an extra assumption.

More remarks:

- A convergence to types argument implies variation properties of $\alpha(\cdot)$ and $\beta(\cdot)$: Suppose $(X, Y)$ satisfy the standard form condition (3). \exists two functions $\psi_1(\cdot), \psi_2(\cdot)$, such that for all $c > 0$,

$$\lim_{t \to \infty} \frac{\alpha(tc)}{\alpha(t)} = \psi_1(c), \quad \lim_{t \to \infty} \frac{\beta(tc) - \beta(t)}{\alpha(t)} \to \psi_2(c).$$

locally uniformly.

- \exists important cases where $\psi_2 \equiv 0$ (bivariate normal).
• Can sometimes also standardize the $X$ variable so that
\[ tP \left[ \frac{\beta^{-}(X)}{t} \leq x, \frac{Y}{t} > y \right] \to \mu([-\infty, \psi_2(x)] \times (y, \infty]). \quad (4) \]

When?? Short version: When $\mu$ is not a product measure.

- $\mu = H \times \nu_1$ iff $\psi_1 \equiv 1 (\alpha(\cdot) \text{ is sv})$ and $\psi_2 \equiv 0$.
- If $\beta(t) \geq 0$ and $\beta^{-}$ is non-decreasing on the range of $X$, then (4) is possible iff $\mu$ is NOT a product.
- A transformation of $X$ allows one to bring the problem to the previous case.

• If we have $X \geq 0$ and both regular variation on $C_2 = [0, \infty]^2 \setminus \{0\}$
\[ tP \left[ \left( \frac{X}{a'(t)}, \frac{Y}{t} \right) \in \cdot \right] \to \nu_* \]
and (4):
\[ tP \left[ \frac{\beta^{-}(X)}{t} \leq x, \frac{Y}{t} > y \right] \to \mu([-\infty, \psi_2(x)] \times (y, \infty]) \]
on $C_1 = [0, \infty] \times (0, \infty]$, then we have a form of hidden regular variation since
\[ C_1 \subset C_2. \]
7.3. Form of the limit.

Assume $\mu$ is not a product and can standardize $X$

$$tP\left[\frac{\beta^- (X)}{t} \leq x, \frac{Y}{t} > y\right] \rightarrow \mu([0, \psi_2(x)] \times (y, \infty)) = \mu_*([0, x] \times (y, \infty))$$

on $C_1 = [0, \infty] \times (0, \infty]$. This is standard regular variation on the cone $C_1$ so

$$\mu_* (c\Lambda) = c^{-1} \mu_* (\Lambda).$$

$\exists$ spectral form: Let

$$\| (x, y) \| = x + y, \quad \mathcal{N} = \{ (w, 1 - w) : 0 \leq w < 1 \}$$

and

$$\mu_* \{ \mathbf{x} : \| \mathbf{x} \| > r, \frac{\mathbf{x}}{\| \mathbf{x} \|} \in A \} = r^{-1} S(A),$$

where $S$ is a measure on $[0, 1]$.

Conclude: Can write $\mu_* [0, x] \times (y, \infty]$ as function of $S$ and get characterization of the class of limit measures.
7.4. Random norming.

When both variables can be standardized

\[ tP \left[ \left( \frac{\beta^{\leftarrow}(X)}{Y}, \frac{Y}{t} \right) \in \cdot \right] \rightarrow G \times \nu_1 \]

in \( M_+([0, \infty] \times (0, \infty]) \) where

\[ \nu_1(x, \infty) = x^{-1}, \quad G(x) = \int_{[0, \frac{x}{1+x}]} (1 - w)S(dw). \]
8. **Medical Care in Copenhagen**

What to expect if you have a knee problem in Copenhagen: