Conditional limit laws for multivariate excesses

Anne-Laure Fougères and Philippe Soulier

Université Paris-Ouest Nanterre

Rennes, August 27, 2008

I. Statement of the problem

II. Models

III. Estimation

I. Statement of the problem

- Quantity of interest:

$$\theta(\mathbf{x},\mathbf{y}) = \mathrm{P}(\mathbf{Y} \leq \mathbf{y} \mid \mathbf{X} > \mathbf{x}) \; .$$

when x is *extreme*, i.e. out of the range of the observations, so that the empirical distribution is useless:

$$\hat{\theta}_{\mathbf{EMP}}(\mathbf{x},\mathbf{y}) = \frac{\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbb{1}_{\{\mathbf{X_i}>\mathbf{x}\}} \mathbb{1}_{\{\mathbf{Y_i}\leq\mathbf{y}\}}}{\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbb{1}_{\{\mathbf{X_i}>\mathbf{x}\}}} = \frac{\mathbf{0}}{\mathbf{0}} \; .$$

This is an extreme value problem even though the probability $\theta(\mathbf{x},\mathbf{y})$ may not be extreme.

- Motivations
  - Financial contagion: $\mathbf{P}(\mathbf{Y} > \mathbf{y} \mid \mathbf{X} > \mathbf{x}) > \mathbf{P}(\mathbf{Y} > \mathbf{y})$.
  - Extreme regression (conditional) quantiles.

Multivariate case, Cf. Balkema and Embrechts (2007).

Goals

Find necessary and sufficient conditions for existence of norming sequences $\mathbf{a}(\mathbf{x})$ and $\mathbf{m}(\mathbf{x})$ and non degenerate proper limiting distribution $\boldsymbol{\Psi}$ such that

$$\lim_{\mathbf{x}\to\infty} \mathrm{P}(\mathbf{Y} \leq \mathbf{m}(\mathbf{x}) + \mathbf{a}(\mathbf{x})\mathbf{z} \mid \mathbf{X} > \mathbf{x}) = \boldsymbol{\Psi}(\mathbf{z})\,. \tag{1}$$

Estimate a, m, $\boldsymbol{\Psi}$.

Apply this to estimate $\theta(x, y)$ for extremely large x.

What can Extreme Value Theory do?

- This is not a standard extreme value problem: there is no characterization of the possible limiting probabilities.
  If X and Y are independent, then $\theta(\mathbf{x}, \mathbf{y}) = \mathrm{P}(\mathbf{Y} \leq \mathbf{y})$.
  All probability laws are possible limits.

- If (X, Y) belongs to the domain of attraction of a bivariate extreme value distribution, does condition (1) always hold?

  If (X, Y) is in the domain of attraction of a bivariate EVD with *dependent* marginals: **yes**.
  If (X, Y) is in the domain of attraction of a bivariate EVD with *independent* marginals: **no**.

- Related concept: hidden regular variation (Maulik and Resnick, 2003).

II. Models

There is not a single characterization of bivariate distributions that satisfy condition (1). Two kinds of results are known.

- Particular examples: most bivariate extreme value distributions satisfy condition (1). (Heffernan and Tawn, 2004).
- Geometric conditions.
  Eddy and Gale (1981), Berman (1992): spherical distributions.
  Abdous et al. (2005), Hashorva (2007): elliptical distributions.
  Generalisations: Barbe (2003), Balkema and Embrechts (2007).
- We will describe some geometric conditions.

Geometric conditions

Elliptic distributions

- A bivariate random vector (X, Y) has a standard elliptic distribution if it can be expressed as

$$(\mathbf{X}, \mathbf{Y}) = \mathbf{R}(\cos(\boldsymbol{\Theta}), \boldsymbol{\rho}\cos(\boldsymbol{\Theta}) + \sqrt{\mathbf{1} - \boldsymbol{\rho}^{\mathbf{2}}}\sin(\boldsymbol{\Theta}))\ .$$

  with Θ uniformly distributed on $[0, 2\pi]$ and R independent of Θ.
- If R is in the domain of attraction of an EVD, then (X, Y) is in the domain of attraction of a bivariate EVD.
  1. If R has *regularly varying* upper tails then (X, Y) is in the domain of attraction of a bivariate EVD with *dependent* marginals.
  2. If R has *rapidly varying* upper tails then (X, Y) is in the domain of attraction of a bivariate EVD with *independent* marginals.

- If R has a density h then (X, Y) has a density

$$\frac{\mathbf{h}(\sqrt{\mathbf{x}^2 + (\mathbf{y} - \rho\mathbf{x})^2/\sigma^2})}{\sqrt{\mathbf{x}^2 + (\mathbf{y} - \rho\mathbf{x})^2/\sigma^2}}$$

  with $\sigma^2 = 1 - \rho^2$.

  The level lines of the density are homothetic.

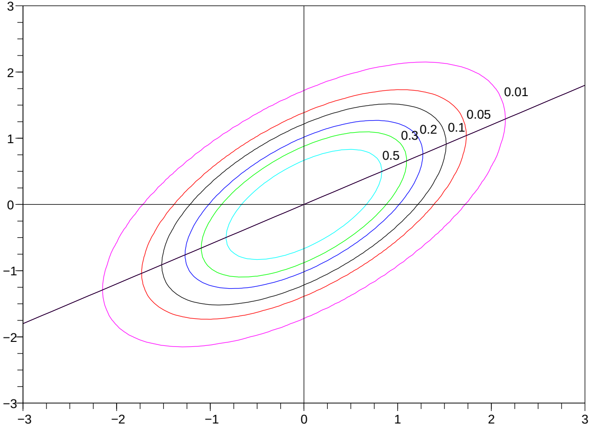


Figure: Level lines of the density of an elliptical distribution. The slope of the straight line is $\rho = .6$.

Theorem

Eddy and Gale (1981), Berman (1992), Abdous et al. (2005), Hashorva (2007).

If R has *rapidly varying* tails, i.e.

$$\lim_{\mathbf{x}\to\infty} \frac{\mathrm{P}(\mathbf{R} > \mathbf{x} + \psi(\mathbf{x})\mathbf{t})}{\mathrm{P}(\mathbf{R} > \mathbf{x})} = \mathrm{e}^{-\mathbf{t}}$$

with $\psi(x)/x \to 0$;

then

$$\lim_{\mathbf{x}\to\infty} \mathrm{P}(\mathbf{Y} \le \boldsymbol{\rho}\mathbf{x} + \boldsymbol{\sigma}\sqrt{\mathbf{x}\psi(\mathbf{x})}\mathbf{z} \mid \mathbf{X} > \mathbf{x}) = \boldsymbol{\Phi}(\mathbf{z}) \; .$$

where $\boldsymbol{\Phi}$ is the cdf of the standard Gaussian distribution.

This result can be generalised in two directions.

1. By weakening the conditions on the level lines of the density.
2. By extending the class of functions u and v in the representation $(X, Y) = R(u(T), v(T))$ where T is independent of R and has values in $[0, 1]$.

Geometric conditions

Generalisation 1. (heuristics)

Let (X, Y) be a random vector that admits a density and such that:

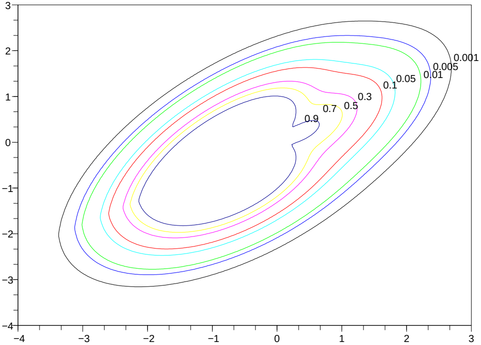
1. The level lines of the density
   - are (asymptotically) homothetical;
   - have (asymptotically) a vertical tangent at their intersection with the line $y = \rho x$;
   - have (asymptotically) positive finite curvature σ;
2. X has a rapidly varying distribution with auxiliary function ψ.

Then

$$\lim_{\mathbf{x}\to\infty} \mathrm{P}(\mathbf{Y} \leq \rho\mathbf{x} + \sigma\sqrt{\mathbf{x}\psi(\mathbf{x})}\mathbf{z} \mid \mathbf{X} > \mathbf{x}) = \mathbf{\Phi}(\mathbf{z}) ,$$

where $\mathbf{\Phi}$ is the cdf of the standard Gaussian distribution.

Rigorous results: Balkema and Embrechts (2004,2007).



Level lines of the density

$$\mathbf{f(x,y)} = \mathbf{c}\exp\{-\mathbf{x}^2 - (\mathbf{y} - \rho\mathbf{x})^2/(\mathbf{1} - \rho^2)\}$$
$$\times[\mathbf{1} + \arctan^2\{(\mathbf{y} - \rho\mathbf{x})/\mathbf{x}\sqrt{\mathbf{1} - \rho^2}\} - (\pi/\mathbf{4})^2]^2 \ .$$

In Balkema and Embrechts (2007), the function $\arctan^2\{(\mathbf{y} - \rho\mathbf{x})/\mathbf{x}\sqrt{\mathbf{1} - \rho^2}\} - (\pi/\mathbf{4})^2$ *is refered to as a* flat function*.*

Generalisation 2.

- The random vector (X, Y) can be expressed as $(X, Y) = R(u(T), v(T))$, where
  1. $u : [0, 1] \to [0, 1]$ is continuous, has a unique maximum 1 at a point $t_0 \in (0, 1)$ and has an expansion

     $$\mathbf{u(t_0 + t) = 1 + \ell(t)}$$

     where ℓ is increasing in $[-\epsilon, 0]$, decreasing in $[0, \epsilon]$ for some $\epsilon > 0$ and regularly varying at zero with index $\boldsymbol{\kappa} > 0$.
  2. v is strictly increasing in a neighborhood of t_0, $v(t_0) = \boldsymbol{\rho}$; the function $t \to v(t_0 + t) - \rho$ is regularly varying with index $\boldsymbol{\delta} \in (0, \boldsymbol{\kappa})$.
  3. R has a rapidly varying distribution with auxiliary function $\boldsymbol{\psi}$.
  4. T is independent of R and has a locally bounded density f on $[0, 1] \setminus \{t_0\}$, regularly varying at t_0 with index $\boldsymbol{\tau} > -1$.

Theorem

- There exists a function g regularly varying at 0 with index $(1+\boldsymbol{\tau})/\boldsymbol{\kappa}$ such that

$$\mathrm{P}(\mathbf{X} > \mathbf{x}) = \mathbf{g}(\psi(\mathbf{x})/\mathbf{x})\mathrm{P}(\mathbf{R} > \mathbf{x}) \; .$$

- There exists a function h regularly varying at 0 with index $\boldsymbol{\delta}/\boldsymbol{\kappa}$ such that, for all y,

$$\lim_{\mathbf{x}\to\infty} \mathrm{P}(\mathbf{Y} \leq \boldsymbol{\rho}\mathbf{x} + \mathbf{x}\mathbf{h}(\psi(\mathbf{x})/\mathbf{x})\mathbf{y} \mid \mathbf{X} > \mathbf{x}) = \mathbf{H}_{\boldsymbol{\delta},\boldsymbol{\kappa},\boldsymbol{\tau}}(\mathbf{y}) \; ,$$

with

$$\mathbf{H}_{\delta,\kappa,\tau}(\mathbf{y}) = \frac{\mathbf{1}}{\mathbf{c}(\delta,\kappa,\tau)} \int_{-\infty}^{\mathbf{y}} \mathrm{e}^{-\frac{|\mathbf{s}|^{\kappa/\delta}}{\kappa/\delta}} \, |\mathbf{s}|^{\frac{\mathbf{1}+\tau}{\delta}-\mathbf{1}} \, \mathrm{d}\mathbf{s} \; .$$

Examples

- If $\kappa = 2$ and $\delta = 1$ (smooth curve) and $\tau = 0$ (the density f of T is continuous at t_0), then the result of Balkema and Embrechts (2007) can also be applied. The limiting distribution is the standard Gaussian and the normalisation is $\sqrt{x\psi(x)}$.
- L_p-spherical distributions. Hashorva et al. (2007)

$$\mathbf{u}(\mathbf{t}) = (\mathbf{1} - \mathbf{t}^{\mathbf{p}})^{\mathbf{1}/\mathbf{p}} \,, \quad \mathbf{v}(\mathbf{t}) = \mathbf{t} \,.$$

Then $\kappa = p$ and $\delta = 1$. If T is uniform or if its density is bounded above and away from zero ($\tau = 0$), then the normalisation is $x^{1-1/p}\psi(x)^{1/p}$ and the limit distribution is

$$\mathbf{H}_{\mathbf{p}}(\mathbf{y}) = \frac{\mathbf{1}}{\mathbf{2p}^{\mathbf{1}/\mathbf{p}-\mathbf{1}}\boldsymbol{\Gamma}(\mathbf{1}/\mathbf{p})} \int_{-\infty}^{\mathbf{y}} \mathrm{e}^{-\frac{|\mathbf{s}|^{\mathbf{p}}}{\mathbf{p}}} \,\mathrm{d}\mathbf{s} \,.$$

A by-product Lemma

- Let R be a nonnegative r. v. whose cdf H is rapidly varying with auxiliary function ψ.
  Let U be a nonnegative r. v., independent of R, such that $U \leq b < \infty$ a.s. and that admits a density g in a neighborhood of b which is regularly varying with index $\tau > -1$.

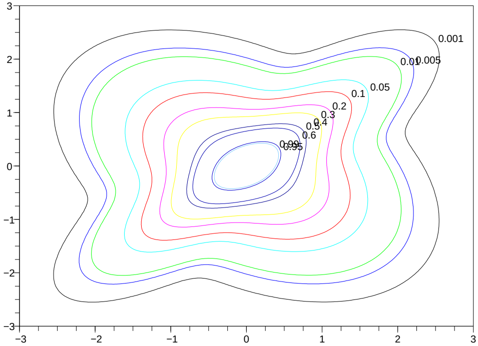
  Then RU has a rapidly varying upper tail with auxiliary function $\psi(x/b)$ and

  $$\mathrm{P}(\mathbf{RU} > \mathbf{x}) \sim \mathbf{b}^2 \boldsymbol{\Gamma}(\tau + \mathbf{1}) \frac{\psi(\mathbf{x}/\mathbf{b})}{\mathbf{x}} \mathbf{g}(\{\mathbf{1}/\mathbf{b} + \psi(\mathbf{x}/\mathbf{b})/\mathbf{x}\}^{-\mathbf{1}}) \bar{\mathbf{H}}(\mathbf{x}/\mathbf{b}) \,.$$

  (already known?)

Remark

In the case where X has *rapidly varying* tails, these results are **local**. It is possible to condition on (X, Y) to remain in a cone around the line $y = \rho x$. This allows to deal with mixtures.



Level lines of the density of a mixture of bivariate Gaussian distributions.

III. Estimation

- In order to define and study estimation procedures, we need to have a deeper understanding of Condition 1:

$$\lim_{\mathbf{x}\to\infty} \mathrm{P}(\mathbf{Y} \leq \mathbf{m}(\mathbf{x}) + \mathbf{a}(\mathbf{x})\mathbf{z} \mid \mathbf{X} > \mathbf{x}) = \mathbf{\Psi}(\mathbf{z}) \; .$$

- Heffernan and Resnick (2007).

  Assumption 1

  There exist functions a, b, m and ψ such that the measure $\boldsymbol{\nu}_n$ defined by

$$\boldsymbol{\nu}_{\mathbf{n}}(\cdot) = \mathbf{nP}\left(\left\{\frac{\mathbf{X} - \mathbf{b}(\mathbf{n})}{\psi \circ \mathbf{b}(\mathbf{n})}, \frac{\mathbf{Y} - \mathbf{m} \circ \mathbf{b}(\mathbf{n})}{\mathbf{a} \circ \mathbf{b}(\mathbf{n})}\right\} \in \cdot\right)$$

  converges vaguely to a Radon measure $\boldsymbol{\nu}$ on $(-\infty, \infty] \times [-\infty \times \infty]$ such that $\boldsymbol{\nu}([0, \infty] \times [-\infty \times \infty]) = 1$.

Consequences. 1

Assumption 1 is a standard vague convergence assumption. It implies that

- X is in the max domain of attraction of an extreme value distribution, with auxiliary function ψ;
- b is the quantile function of X;
- The first marginal of $\boldsymbol{\nu}$ is a GPD:

$$\int_{\mathbf{0}}^{\mathbf{t}} \int_{-\infty}^{\infty} \boldsymbol{\nu}(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}) = \mathbf{1} - (\mathbf{1} + \gamma \mathbf{t})^{-\mathbf{1}/\gamma} \quad (\gamma \geq \mathbf{0}) \; .$$

- The second marginal of $\boldsymbol{\nu}$ is $\boldsymbol{\Psi}$.
- The random measure

$$\frac{\mathbf{1}}{\mathbf{k}} \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \delta_{\left\{ \frac{\mathbf{X}_{(\mathbf{i})} - \mathbf{b}(\mathbf{n}/\mathbf{k})}{\psi \circ \mathbf{b}(\mathbf{n}/\mathbf{k})}, \frac{\mathbf{Y}_{[\mathbf{i}]} - \mathbf{m} \circ \mathbf{b}(\mathbf{n}/\mathbf{k})}{\mathbf{a} \circ \mathbf{b}(\mathbf{n}/\mathbf{k})} \right\}} ,$$

  where $X_{(i)}$ is the i-th order statistics (in decreasing order) and $Y_{[i]}$ is its concomittant, converges weakly to $\boldsymbol{\nu}$.

Consequences. 2

- What can be said of the functions a and m?
  Not much!

- What can be said of the limiting measure ν?
  Not much either!

  Two cases: ν is or is not a product measure.

Consequences. 3

- Is it possible to standardize both X and Y and still have a conditional limit as in Assumption 1, with linear normalisations? **Not in general**.
- No copulas!
- Heffernan and Resnick (2007):
  It is possible if and only if the limiting measure ν is not a product measure.

Conclusions?

- Examples lead to assume
  - $m(x) = \rho x$ for some $\rho \in \mathbb{R}$.
  - $a(x) = O(x)$.

  Heuristically, this means that X and Y are commensurate.

- We will focus on the case where
  - X has rapidly varying tails;
  - X and Y are asymptotically independent (in the tails);
  - $\boldsymbol{\nu}$ is in product form:

$$\boldsymbol{\nu}((\mathbf{t}, \infty) \times (-\infty, \mathbf{y}]) = \mathrm{e}^{-\mathbf{t}} \boldsymbol{\Psi}(\mathbf{y}) .$$

Estimation

- Goal: estimate $\theta(\mathbf{x},\mathbf{y}) = \mathbf{P}(\mathbf{Y} \le \mathbf{y} \mid \mathbf{X} < \mathbf{x})$ for very large x by means of the approximation

$$\lim_{\mathbf{x}\to\infty} \mathbf{P}(\mathbf{Y} \le \rho\mathbf{x} + \mathbf{a}(\mathbf{x})\mathbf{z} \mid \mathbf{X} > \mathbf{x}) = \mathbf{\Psi}(\mathbf{z}) \ .$$

- If $\mathbf{\Psi}$ is known (e.g. $\mathbf{\Psi} = \mathbf{\Phi}$ the standard Gaussian law), estimate ρ and a and define

$$\hat{\theta}(\mathbf{x},\mathbf{y}) = \mathbf{\Psi}\left(\frac{\mathbf{y} - \hat{\rho}\mathbf{x}}{\hat{\mathbf{a}}(\mathbf{x})}\right) \ .$$

- If $\mathbf{\Psi}$ is unknown, estimate it also and define

$$\hat{\theta}(\mathbf{x},\mathbf{y}) = \hat{\mathbf{\Psi}}\left(\frac{\mathbf{y} - \hat{\rho}\mathbf{x}}{\hat{\mathbf{a}}(\mathbf{x})}\right) \ .$$

Assumptions

1. There exist monotone functions a, b, ψ and a real number ρ such that the measure $\boldsymbol{\nu}_n$ defined by

$$\nu_{\mathbf{n}}(\cdot) = \mathbf{nP}\left(\left\{\frac{\mathbf{X}-\mathbf{b}(\mathbf{n})}{\psi\circ\mathbf{b}(\mathbf{n})}, \frac{\mathbf{Y}-\rho\mathbf{b}(\mathbf{n})}{\mathbf{a}\circ\mathbf{b}(\mathbf{n})}\right\}\in\cdot\right)$$

converges vaguely to a Radon measure ν on $(-\infty,\infty]\times[-\infty\times\infty]$ such that $\nu([0,\infty]\times[-\infty\times\infty]) = 1$.

2. There exist $\zeta^* > 0$, $\kappa^* > 0$ such that

$$\lim_{\mathbf{n}\to\infty}\int_{\mathbf{0}}^{\infty}\int_{-\infty}^{\infty}\mathbf{x}^{\zeta^*}|\mathbf{y}|^{\kappa^*}\nu_{\mathbf{n}}(\mathrm{d}\mathbf{x},\mathrm{d}\mathbf{y}) = \int_{\mathbf{0}}^{\infty}\int_{-\infty}^{\infty}\mathbf{x}^{\zeta^*}|\mathbf{y}|^{\kappa^*}\nu(\mathrm{d}\mathbf{x},\mathrm{d}\mathbf{y})\ .$$

3. The function ψ satisfies $\lim_{x\to\infty}\psi(x)/x = 0$, and the measure $\boldsymbol{\nu}$ is in product form:

$$\boldsymbol{\nu}((\mathbf{x},\infty)\times[-\infty,\mathbf{y}]) = \mathrm{e}^{-\mathbf{x}}\boldsymbol{\Psi}(\mathbf{y})\ ,$$

where $\boldsymbol{\Psi}$ is some probability distribution function.

Comments

- Assumption 2 is a moment assumption that allows to use bounded convergence arguments.
- Assumption 3 holds for the previous examples. It implies that $\psi(x) = o(a(x))$ and

  $$(\mathbf{X} - \mathbf{x})/\mathbf{a}(\mathbf{x}) \to \mathbf{0} \text{ conditionally on } \mathbf{X} > \mathbf{x}$$

  and most of the limiting results hold with X or x.

  Some results can be obtained without Assumption 3. For brevity, we present only the results under Assumption 3.

Estimation of ρ

- Let k be an intermediate sequence, i.e.

$$\lim_{n\to\infty} \mathbf{k} = \lim_{n\to\infty} \mathbf{n}/\mathbf{k} = \infty \ .$$

- If Assumptions 1, 2,and 3 hold, then

$$\hat{\rho} = \frac{\sum_{i=1}^{k} \mathbf{Y}_{[i]} \{\mathbf{X}_{(i)} - \mathbf{X}_{(k+1)}\}}{\sum_{i=1}^{k} \mathbf{X}_{(i)} \{\mathbf{X}_{(i)} - \mathbf{X}_{(k+1)}\}} \longrightarrow_{\mathbf{P}} \rho \ .$$

Non parametric estimation of ψ and a

- Define

$$\hat{\psi}_{\mathbf{NP}}(\mathbf{X}_{(\mathbf{k}+1)}) = \frac{1}{\mathbf{k}} \sum_{\mathbf{i}=1}^{\mathbf{k}} (\mathbf{X}_{(\mathbf{i})} - \mathbf{X}_{(\mathbf{k}+1)}) \; .$$

  Under Assumption 1, $\hat{\psi}_{\mathbf{NP}}(\mathbf{X}_{(\mathbf{k}+1)})/\psi(\mathbf{X}_{(\mathbf{k}+1)}) \rightarrow_{\mathbf{P}} \mathbf{1}$.

- Define

$$\hat{\mathbf{a}}^2_{\mathbf{NP}}(\mathbf{X}_{(\mathbf{k}+1)}) = \frac{1}{\mathbf{k}} \sum_{\mathbf{i}=1}^{\mathbf{k}} \{\mathbf{Y}_{[\mathbf{i}]} - \hat{\rho}\mathbf{X}_{(\mathbf{i})}\}^2 \; .$$

  Under Assumptions 1, 2, and 3,

$$\hat{\mathbf{a}}_{\mathbf{NP}}(\mathbf{X}_{(\mathbf{k}+1)})/\mathbf{a}(\mathbf{X}_{(\mathbf{k}+1)}) \rightarrow_{\mathbf{P}} \mathbf{1} \; .$$

Non parametric estimation of Ψ

- Define

$$\hat{\Psi}(z) = \frac{1}{k} \sum_{i=1}^{k} \mathbb{1}_{\{Y_{[i]} \leq \hat{\rho} X_{(i)} + \hat{a}(X_{(k+1)}) z\}} \ .$$

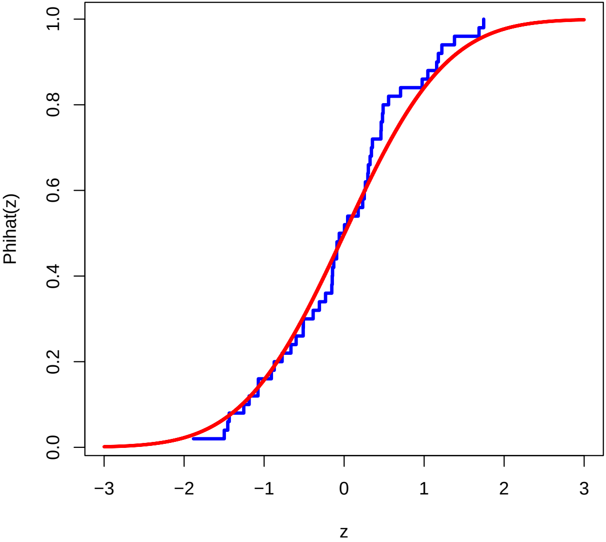
  Under Assumptions 1, 2, and 3, $\hat{\Psi}(z) \to_P \hat{\Psi}(z)$.

- Under a second order condition and an adequate choice of k,

$$k^{1/2}(\hat{\Psi} - \Psi)$$

  converges in $\mathcal{D}$ to a Brownian bridge.
  If Ψ is continuous this allows to derive a K-S test.

Non parametric estimation of $\mathbf{\Psi}$ for a sample of size $n = 1000$ of elliptical vectors with non uniform angular distribution. $k = 50$.



Semi-parametric estimation of ψ and a

- We want to estimate $\theta(x, y) = P(Y \leq y \mid X > x)$ for large x, using the asymptotic approximation

$$\hat{\theta}(\mathbf{x}, \mathbf{y}) = \mathbf{\Psi}\left(\frac{\mathbf{y} - \hat{\boldsymbol{\rho}}\mathbf{x}}{\hat{\mathbf{a}}(\mathbf{x})}\right) \quad \text{or} \quad \hat{\theta}(\mathbf{x}, \mathbf{y}) = \hat{\mathbf{\Psi}}\left(\frac{\mathbf{y} - \hat{\boldsymbol{\rho}}\mathbf{x}}{\hat{\mathbf{a}}(\mathbf{x})}\right).$$

  If x is outside the range of the observations, then the previous nonparametric estimation of a is useless.

- Solution: extrapolation by semiparametric estimation.

Assumptions:

(i) X has a rapidly varying tail, with auxiliary function ψ of Weibull type, i.e. $\psi(\mathbf{x}) = \mathbf{x}^{\mathbf{1}-\boldsymbol{\beta}}/(\mathbf{c}\boldsymbol{\beta})$, for $c, \beta > 0$;

(ii) the limiting distribution $\mathbf{\Psi}$ is the gaussian cdf $\mathbf{\Phi}$;

(iii) $\mathbf{a}(\mathbf{x}) = \boldsymbol{\sigma}\sqrt{\mathbf{x}\psi(\mathbf{x})}$, for some positive $\boldsymbol{\sigma}$;

Semi-parametric estimation of ψ and a

- Semi-parametric estimation of $\psi(x) = x^{1-\beta}/(c\beta)$.

$$\hat{\beta} = \frac{\sum_{i=1}^{k} \log\log(n/i) - \log\log(n/(k+1))}{\sum_{i=1}^{k} \log(X_{(i)}) - \log(X_{(k+1)})} ,$$

$$\hat{c} = \frac{1}{k} \sum_{i=1}^{k} \frac{\log(n/i)}{X_{(i)}^{\hat{\beta}}} ,$$

$$\hat{\psi}_{SP}(x) = x^{1-\hat{\beta}}/(\hat{c}\hat{\beta})$$

  Beirlant et al (1996).
- If $a(x) = \sigma\sqrt{x\psi(x)}$, this yields a semi-param. estimator of a:

$$\hat{a}_{SP}(x) = \hat{\sigma}(\hat{c}\hat{\beta})^{-1/2} x^{(2-\hat{\beta})/2} ;$$

  $\hat{\sigma}$ is defined via the *non parametric* estimators of ψ and a:

$$\hat{\sigma} = \frac{\hat{a}_{NP}(X_{(k+1)})}{\sqrt{X_{(k+1)}\hat{\psi}_{NP}(X_{(k+1)})}} .$$

Under Assumptions (i)–(iii), which imply Assumptions 1 and 3, and under Assumption 2,

$$\hat{\theta}(\mathbf{x}, \mathbf{y}) = \boldsymbol{\Phi} \left\{ \frac{\mathbf{y} - \hat{\boldsymbol{\rho}}\mathbf{x}}{\hat{\boldsymbol{\sigma}} \sqrt{\mathbf{x}\hat{\psi}(\mathbf{x})}} \right\}$$

is a consistent estimator of $\boldsymbol{\Phi} \left\{ \frac{y - \boldsymbol{\rho} x}{\boldsymbol{\sigma} \sqrt{x \boldsymbol{\psi}(x)}} \right\}$.

Numerical illustration

Under Assumptions (i)–(iii).

Three types of distributions:

(a) an elliptical distribution (the bivariate normal distribution);

(b) an asymptotically elliptic distribution (a distribution with radial representation $R(\cos(T), \sin(T))$, where the r.v. T has a non uniform concave density function, and the radial r.v. R has survival function $\bar{H}(t) = \mathrm{e}^{-t^2/2}$);

(c) a locally elliptic distribution (a mixture of two gaussian distributions).

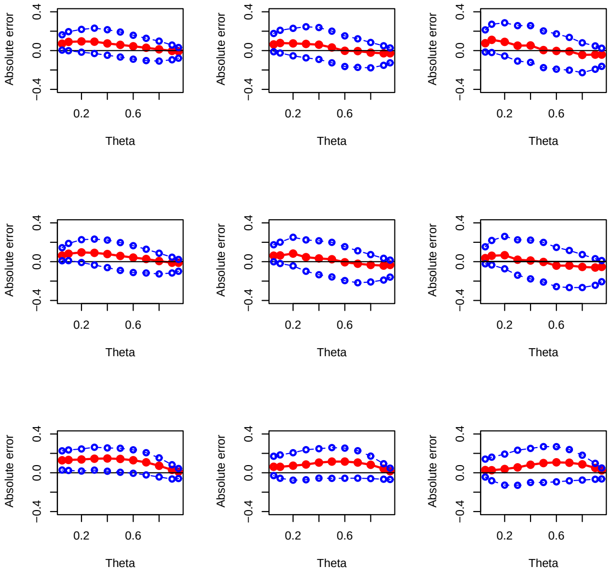
In each case, 100 samples of size 1000 are simulated. A proportion of 5% of the observations is used, which are the 50 observations with biggest first component.

For each sample, the semi-parametric estimate of $\theta(\mathbf{x}, \mathbf{y})$ is calculated for three values of x corresponding to the theoretical X-quantiles of order $1 - p$, where $p = 10^{-3}, 10^{-5}, 10^{-7}$, and different values of y chosen to form a grid of possible values of the probability $\theta(\mathbf{x}, \mathbf{y})$ from 0 to 1.

Median, 2.5%- and 97.5%-quantiles of a sample of size 100 of absolute errors $\hat{\theta}(x, y_p) - \theta(x, y_p)$ in terms of $p \in [0, 1]$, where y_p is the theoretical conditional quantile of (X, Y) given that $X > x$.

In column 1 (resp. 2, 3) the value x is the theoretical X-quantile of order $1 - 10^{-3}$ (resp. $1 - 10^{-5}$, $1 - 10^{-7}$).

The distribution of (X, Y) is **(a)** the bivariate normal distribution on Row 1; **(b)** an asymptotically elliptic distribution on Row 2; **(c)** a locally elliptic distribution on Row 3.


Absolute error
Theta
0.4
0.0
−0.4
0.2
0.6