Semi-parametric Estimation of Multivariate Extreme Expectiles

Elena Di Bernardino

Laboratoire J.A. Dieudonné, Université Côte d'Azur

Joint work with Nicholas Beck (HEC Montréal) and Mélina Mailhot (Concordia University, Montréal)

RESIM 2021 : 13th International Workshop on Rare-Event Simulation

May 18th, 2021

Contents

Risk measures : crucial tool in a multitude of fields relating to mathematics and statistics, constantly evolving.

There have been numerous developments in this field :

- ↪ Establishing ideal properties
- ↪ Extensions of univariate measures to higher dimension
- ↪ Development of new measures
- ↪ Estimating these measures non-parametrically or semi-parametrically

Risk measures : crucial tool in a multitude of fields relating to mathematics and statistics, constantly evolving.

There have been numerous developments in this field :

↪ Establishing ideal properties

↪ Extensions of univariate measures to higher dimension

↪ Development of new measures

↪ Estimating these measures non-parametrically or semi-parametrically

Risk measures : crucial tool in a multitude of fields relating to mathematics and statistics, constantly evolving.

There have been numerous developments in this field :

- ↪ Establishing ideal properties
- ↪ Extensions of univariate measures to higher dimension
- ↪ Development of new measures
- ↪ Estimating these measures non-parametrically or semi-parametrically

Risk measures : crucial tool in a multitude of fields relating to mathematics and statistics, constantly evolving.

There have been numerous developments in this field :

- ↪ Establishing ideal properties
- ↪ Extensions of univariate measures to higher dimension
- ↪ Development of new measures
- ↪ Estimating these measures non-parametrically or semi-parametrically

Risk measures : crucial tool in a multitude of fields relating to mathematics and statistics, constantly evolving.

There have been numerous developments in this field :

- ↪ Establishing ideal properties
- ↪ Extensions of univariate measures to higher dimension
- ↪ Development of new measures
- ↪ Estimating these measures non-parametrically or semi-parametrically

Risk measures : crucial tool in a multitude of fields relating to mathematics and statistics, constantly evolving.

There have been numerous developments in this field :

- ↪ Establishing ideal properties
- ↪ Extensions of univariate measures to higher dimension
- ↪ Development of new measures
- ↪ Estimating these measures non-parametrically or semi-parametrically

Elicitability

Gneiting (2011) defines elicitability as the ability to express a risk measure $T_\alpha(X)$ in the form of an optimization problem:

$$T_\alpha(X) = \operatorname*{argmin}_{x \in \mathbb{R}} \mathbb{E}\left\{S(x, \alpha)\right\},$$

for a risk level $\alpha \in (0, 1)$, a random risk X, where S is the score function associated to the risk measure T_α.

Two of the most well-known elicitable univariate risk measures:

- Value-at-risk (VaR) $\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$

$$\mathrm{VaR}_\alpha(X) = \operatorname*{argmin}_{x \in \mathbb{R}} \mathbb{E}\{\alpha(X - x)_+ + (1 - \alpha)(X - x)_-\}$$

  where $x_+ = \max\{0, x\}$ and $x_- = \max\{0, -x\}$,

- Expectiles ("expectation+quantiles")

$$e_\alpha(X) = \operatorname*{argmin}_{x \in \mathbb{R}} \mathbb{E}\left\{\alpha(X - x)_+^2 + (1 - \alpha)(X - x)_-^2\right\}.$$

Expectiles are *coherent* when $\alpha \geq 0.5$; this is quite advantageous.

Elicitability

Gneiting (2011) defines elicitability as the ability to express a risk measure $T_\alpha(X)$ in the form of an optimization problem:

$$T_\alpha(X) = \underset{x \in \mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\left\{S(x, \alpha)\right\},$$

for a risk level $\alpha \in (0,1)$, a random risk X, where S is the score function associated to the risk measure T_α.

Two of the most well-known elicitable univariate risk measures:

- Value-at-risk (VaR) $\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$

$$\mathrm{VaR}_\alpha(X) = \underset{x \in \mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\{\alpha(X-x)_+ + (1-\alpha)(X-x)_-\}$$

  where $x_+ = \max\{0, x\}$ and $x_- = \max\{0, -x\}$,

- Expectiles ("expectation+quantiles")

$$e_\alpha(X) = \underset{x \in \mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\left\{\alpha(X-x)_+^2 + (1-\alpha)(X-x)_-^2\right\}.$$

Expectiles are *coherent* when $\alpha \geq 0.5$; this is quite advantageous.

Elicitability

Gneiting (2011) defines elicitability as the ability to express a risk measure $T_\alpha(X)$ in the form of an optimization problem:

$$T_\alpha(X) = \operatorname*{argmin}_{x \in \mathbb{R}} \mathbb{E}\left\{S(x, \alpha)\right\},$$

for a risk level $\alpha \in (0,1)$, a random risk X, where S is the score function associated to the risk measure T_α.

Two of the most well-known elicitable univariate risk measures:

Value-at-risk (VaR) $\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$

$$\mathrm{VaR}_\alpha(X) = \operatorname*{argmin}_{x \in \mathbb{R}} \mathbb{E}\{\alpha(X - x)_+ + (1-\alpha)(X - x)_-\}$$

where $x_+ = \max\{0, x\}$ and $x_- = \max\{0, -x\}$,

Expectiles (“expectation+quantiles”)

$$e_\alpha(X) = \operatorname*{argmin}_{x \in \mathbb{R}} \mathbb{E}\left\{\alpha(X - x)_+^2 + (1-\alpha)(X - x)_-^2\right\}.$$

Expectiles are *coherent* when $\alpha \geq 0.5$; this is quite advantageous.

Elicitability

Gneiting (2011) defines elicitability as the ability to express a risk measure $T_\alpha(X)$ in the form of an optimization problem:

$$T_\alpha(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \mathbb{E}\left\{S(x,\alpha)\right\},$$

for a risk level $\alpha \in (0,1)$, a random risk X, where S is the score function associated to the risk measure T_α.

Two of the most well-known elicitable univariate risk measures:

Value-at-risk (VaR) $\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$

$$\mathrm{VaR}_\alpha(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \mathbb{E}\{\alpha(X-x)_+ + (1-\alpha)(X-x)_-\}$$

where $x_+ = \max\{0,x\}$ and $x_- = \max\{0,-x\}$,

Expectiles (“expectation+quantiles”)

$$e_\alpha(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \mathbb{E}\left\{\alpha(X-x)_+^2 + (1-\alpha)(X-x)_-^2\right\}.$$

Expectiles are *coherent* when $\alpha \geq 0.5$; this is quite advantageous.

Expectiles

Expectiles are uniquely identified by the first-order condition,

$$\alpha \, \mathbb{E}\left[\{X - e_\alpha(X)\}_+\right] = (1-\alpha) \, \mathbb{E}\left[\{X - e_\alpha(X)\}_-\right].$$

The above equation can also be written as

$$\frac{1-\alpha}{\alpha} = \frac{\mathbb{E}\left[\{X - e_\alpha(X)\}_+\right]}{\mathbb{E}\left[\{X - e_\alpha(X)\}_-\right]}.$$

This makes the economic interpretation of expectiles as risk measures clearer:

Expectiles can be seen as the value of X that provides a profits/loss ratio of $\frac{1-\alpha}{\alpha}$.

Note that both expectiles and VaR fall into the family of generalized quantiles (Bellini et al. 2014), defined by

$$q_\alpha(X) = \underset{x \in \mathbb{R}}{\operatorname{argmin}} \left(\alpha \mathbb{E}\left[\Phi_1\{(X-x)_+\}\right] + (1-\alpha)\mathbb{E}\left[\Phi_2\{(X-x)_-\}\right]\right),$$

where $\Phi_1, \Phi_2 : [0,\infty) \mapsto [0,\infty)$ are strictly increasing convex functions satisfying $\Phi_i(0) = 0$ and $\Phi_i(1) = 1$ for $i \in \{1,2\}$.

Expectiles

Expectiles are uniquely identified by the first-order condition,

$$\alpha \, \mathbb{E}\left[\{X - e_\alpha(X)\}_+\right] = (1-\alpha) \, \mathbb{E}\left[\{X - e_\alpha(X)\}_-\right].$$

The above equation can also be written as

$$\frac{1-\alpha}{\alpha} = \frac{\mathbb{E}\left[\{X - e_\alpha(X)\}_+\right]}{\mathbb{E}\left[\{X - e_\alpha(X)\}_-\right]}.$$

This makes the economic interpretation of expectiles as risk measures clearer:

Expectiles can be seen as the value of X that provides a profits/loss ratio of $\frac{1-\alpha}{\alpha}$.

Note that both expectiles and VaR fall into the family of generalized quantiles (Bellini et al. 2014), defined by

$$q_\alpha(X) = \underset{x \in \mathbb{R}}{\operatorname{argmin}} \left(\alpha \mathbb{E}\left[\Phi_1\{(X-x)_+\}\right] + (1-\alpha)\mathbb{E}\left[\Phi_2\{(X-x)_-\}\right]\right),$$

where $\Phi_1, \Phi_2 : [0,\infty) \mapsto [0,\infty)$ are strictly increasing convex functions satisfying $\Phi_i(0) = 0$ and $\Phi_i(1) = 1$ for $i \in \{1,2\}$.

Expectiles

Expectiles are uniquely identified by the first-order condition,

$$\alpha\,\mathbb{E}\left[\{X - e_\alpha(X)\}_+\right] = (1-\alpha)\,\mathbb{E}\left[\{X - e_\alpha(X)\}_-\right].$$

The above equation can also be written as

$$\frac{1-\alpha}{\alpha} = \frac{\mathbb{E}\left[\{X - e_\alpha(X)\}_+\right]}{\mathbb{E}\left[\{X - e_\alpha(X)\}_-\right]}.$$

This makes the economic interpretation of expectiles as risk measures clearer:

Expectiles can be seen as the value of X that provides a profits/loss ratio of $\frac{1-\alpha}{\alpha}$.

Note that both expectiles and VaR fall into the family of generalized quantiles (Bellini et al. 2014), defined by

$$q_\alpha(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \left(\alpha\mathbb{E}\left[\Phi_1\{(X-x)_+\}\right] + (1-\alpha)\mathbb{E}\left[\Phi_2\{(X-x)_-\}\right]\right),$$

where $\Phi_1, \Phi_2 : [0,\infty) \mapsto [0,\infty)$ are strictly increasing convex functions satisfying $\Phi_i(0) = 0$ and $\Phi_i(1) = 1$ for $i \in \{1,2\}$.

Elicitability in multivariate context

Ignoring **potential dependence between risks** can provide inaccurate inference and induce prohibitive losses. As such, our interest lies in exploring multivariate expectiles as these dependence structures can be incorporated directly into the measure.

For any d-dimensional random vector $\boldsymbol{X} \in \mathbb{R}^d$ an associated risk measure $\boldsymbol{T}_\alpha$ is elicitable if

$$\boldsymbol{T}_\alpha(\boldsymbol{X}) = \underset{x \in \mathbb{R}^d}{\operatorname{argmin}} \left\{ \boldsymbol{S}(\boldsymbol{x}, \alpha) \right\}.$$

Previous literature on Multivariate Expectiles :

↪ Multivariate geometric definition of expectiles (Herrmann et al. 2018)

↪ (Maume-Deschamps et al. 2017) define two notions of multivariate expectiles: L^p-**expectiles** and Σ-expectiles.

Elicitability in multivariate context

Ignoring **potential dependence between risks** can provide inaccurate inference and induce prohibitive losses. As such, our interest lies in exploring multivariate expectiles as these dependence structures can be incorporated directly into the measure.

For any d-dimensional random vector $\boldsymbol{X} \in \mathbb{R}^d$ an associated risk measure $\boldsymbol{T}_\alpha$ is elicitable if

$$\boldsymbol{T}_\alpha(\boldsymbol{X}) = \underset{x \in \mathbb{R}^d}{\operatorname{argmin}} \left\{ \boldsymbol{S}(\boldsymbol{x}, \alpha) \right\}.$$

Previous literature on Multivariate Expectiles :

- ↪ Multivariate geometric definition of expectiles (Herrmann et al. 2018)
- ↪ (Maume-Deschamps et al. 2017) define two notions of multivariate expectiles: L^p-**expectiles** and Σ-expectiles.

Elicitability in multivariate context

Ignoring **potential dependence between risks** can provide inaccurate inference and induce prohibitive losses. As such, our interest lies in exploring multivariate expectiles as these dependence structures can be incorporated directly into the measure.
For any d-dimensional random vector $\boldsymbol{X} \in \mathbb{R}^d$ an associated risk measure $\boldsymbol{T}_\alpha$ is elicitable if

$$\boldsymbol{T}_\alpha(\boldsymbol{X}) = \underset{x \in \mathbb{R}^d}{\operatorname{argmin}} \left\{ \boldsymbol{S}(\boldsymbol{x}, \alpha) \right\}.$$

Previous literature on Multivariate Expectiles :

↪ Multivariate geometric definition of expectiles (Herrmann et al. 2018)

↪ (Maume-Deschamps et al. 2017) define two notions of multivariate expectiles: L^p**-expectiles** and Σ-expectiles.

L^1-expectiles

Definition (L^1-expectile)

Define the L^1-expectile of a random vector $\boldsymbol{X}$ by

$$\boldsymbol{e}_\alpha(\boldsymbol{X}) = \operatorname*{argmin}_{\boldsymbol{x}\in\mathbb{R}^d} \mathbb{E}\left\{\alpha\left(\sum_{i=1}^d |X_i - x_i|_+\right)^2 + (1-\alpha)\left(\sum_{i=1}^d |X_i - x_i|_-\right)^2\right\}.$$

Analogously to the univariate case, the L^1-expectile is the unique solution in $\boldsymbol{R}^d$ of

$$\frac{1-\alpha}{\alpha} = \frac{\mathbb{E}[\|(X-x)_+\|_1 \mathbb{1}\{X_k > x_k\}]}{\mathbb{E}[\|(X-x)_-\|_1 \mathbb{1}\{X_k < x_k\}]}, \quad k \in \{1,\dots,d\}.$$

Thus, it can be interpreted as a ratio of expected positive scenarios over negative ones.

Our aim:

We aim to explore semi-parametric estimation of the L^1-expectile for elevated risk levels $\alpha \approx 1$

(i) when the underlying dependence structure and marginal distributions are unknown;

(ii) via the approximated optimization problem

$$\underset{\boldsymbol{\Theta} \in \mathbb{R}^d}{\operatorname{argmin}} L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta})$$

for some (asymptotic) loss function L and consistently estimated parameter set $\hat{\boldsymbol{\Lambda}}$.

In Maume-Deschamps et al. 2017 it was shown that multivariate expectiles could be consistently estimated using Robbins-Monro's stochastic optimization for moderate levels of α.

However, for elevated levels of α this approach, without any asymptotic extrapolation techniques, fails.

Our aim:

We aim to explore semi-parametric estimation of the L^1-expectile for elevated risk levels $\alpha \approx 1$

(i) when the underlying dependence structure and marginal distributions are unknown;

(ii) via the approximated optimization problem

$$\underset{\boldsymbol{\Theta} \in \mathbb{R}^d}{\operatorname{argmin}} \, L_{\hat{\boldsymbol{\Lambda}}} (\boldsymbol{\Theta})$$

for some (asymptotic) loss function L and consistently estimated parameter set $\hat{\boldsymbol{\Lambda}}$.

In Maume-Deschamps et al. 2017 it was shown that multivariate expectiles could be consistently estimated using Robbins-Monro's stochastic optimization for moderate levels of α.

However, for elevated levels of α this approach, without any asymptotic extrapolation techniques, fails.

For moderate levels of α (see Maume-Deschamps et al. 2017)

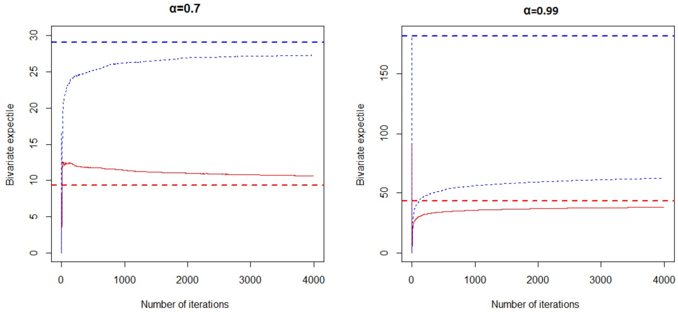


Figure: Difference in convergence between two different levels $\alpha = 0.7$ and $\alpha = 0.99$ for L_1-expectile, Pareto independent model $X_i \sim \mathrm{P}\{2, 10\}$ (red) $X_i \sim \mathrm{P}\{2, 20\}$ (blue).

- ↪ Convergence is not very satisfactory for values of α close to 1.
- ↪ The algorithm is not efficient to estimate the asymptotic expectile.
- ↪ A study of asymptotic behavior of the expectile seems necessary, *particularly in cases where there is no analytical solution.*

For moderate levels of α (see Maume-Deschamps et al. 2017)

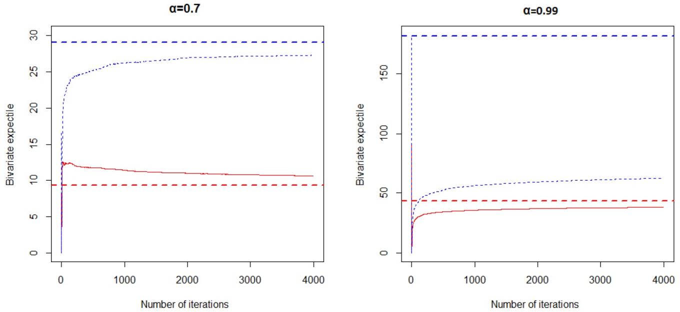


Figure: Difference in convergence between two different levels $\alpha = 0.7$ and $\alpha = 0.99$ for L_1-expectile, Pareto independent model $X_i \sim \mathrm{P}\{2, 10\}$ (red) $X_i \sim \mathrm{P}\{2, 20\}$ (blue).

- ↪ Convergence is not very satisfactory for values of α close to 1.
- ↪ The algorithm is not efficient to estimate the asymptotic expectile.
- ↪ A study of asymptotic behavior of the expectile seems necessary, *particularly in cases where there is no analytical solution.*

Proposition (Maume-Deschamps et al. (2018))

Assume that $\boldsymbol{X}$ has MRV distribution with index θ and, for all $i \in \{2, \ldots, d\}$,

$\lim_{x \to +\infty} \frac{\overline{F}_i(x)}{\overline{F}_1(x)} = c_i$, *(equivalent regularly varying marginal tails).*

Consider the L_1-expectile $\mathbf{e}_\alpha(\boldsymbol{X}) = (e^i_\alpha(\boldsymbol{X}))_{i \in \{1,\ldots,d\}}$. Then any limit vector $\boldsymbol{\Theta}$:

$$\boldsymbol{\Theta} := (\eta, \beta_2, \ldots, \beta_d) = \lim_{\alpha \to 1} \left(\frac{1-\alpha}{\overline{F}_1\{e^1_\alpha(\boldsymbol{X})\}}, \frac{e^2_\alpha(\boldsymbol{X})}{e^1_\alpha(\boldsymbol{X})}, \ldots, \frac{e^d_\alpha(\boldsymbol{X})}{e^1_\alpha(\boldsymbol{X})} \right)$$

satisfies the following system of equations

$$\frac{1}{\theta-1} - \eta \frac{\beta_k^\theta}{c_k} = - \sum_{i=1, i \neq k}^{d} \left\{ \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik}\left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) \mathrm{d}t - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right\}, \; k \in \{1, \ldots, d\}$$

where λ^{ik} is the upper tail dependence (UTD) function for the random pair (X_i, X_k).

In particular, explicit system solutions (see Maume-Deschamps et al. (2018))

- $\boldsymbol{\Theta}^\perp = (\eta^\perp, \beta_2^\perp, \ldots, \beta_d^\perp)$ (asympt. $\perp$ case) and
- $\boldsymbol{\Theta}^+ = (\eta^+, \beta_2^+, \ldots, \beta_d^+)$ (Comon. case).

Proposition (Maume-Deschamps et al. (2018))

Assume that $\mathbf{X}$ has MRV distribution with index θ and, for all $i \in \{2, \dots, d\}$,

$$\lim_{x \to +\infty} \frac{\overline{F}_i(x)}{\overline{F}_1(x)} = c_i, \text{(equivalent regularly varying marginal tails).}$$

Consider the L_1-expectile $\mathbf{e}_\alpha(\mathbf{X}) = (e^i_\alpha(\mathbf{X}))_{i \in \{1,\dots,d\}}$. Then any limit vector $\mathbf{\Theta}$:

$$\mathbf{\Theta} := (\eta, \beta_2, \dots, \beta_d) = \lim_{\alpha \to 1} \left(\frac{1-\alpha}{\overline{F}_1\{e^1_\alpha(\mathbf{X})\}}, \frac{e^2_\alpha(\mathbf{X})}{e^1_\alpha(\mathbf{X})}, \dots, \frac{e^d_\alpha(\mathbf{X})}{e^1_\alpha(\mathbf{X})} \right)$$

satisfies the following system of equations

$$\frac{1}{\theta - 1} - \eta \frac{\beta_k^\theta}{c_k} = - \sum_{i=1, i \neq k}^{d} \left\{ \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) \mathrm{d}t - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right\}, \; k \in \{1, \dots, d\}$$

where λ^{ik} is the upper tail dependence (UTD) function for the random pair (X_i, X_k).

In particular, explicit system solutions (see Maume-Deschamps et al. (2018))

- $\mathbf{\Theta}^{\perp} = (\eta^{\perp}, \beta_2^{\perp}, \dots, \beta_d^{\perp})$ (asympt. $\perp$ case) and
- $\mathbf{\Theta}^{+} = (\eta^{+}, \beta_2^{+}, \dots, \beta_d^{+})$ (Comon. case).

Proposition (Maume-Deschamps et al. (2018))

Assume that $\boldsymbol{X}$ *has MRV distribution with index* θ *and, for all* $i \in \{2, \ldots, d\}$,

$\lim_{x\to+\infty} \frac{\overline{F}_i(x)}{\overline{F}_1(x)} = c_i$, *(equivalent regularly varying marginal tails).*

Consider the L_1*-expectile* $\boldsymbol{e}_\alpha(\boldsymbol{X}) = (e^i_\alpha(\boldsymbol{X}))_{i\in\{1,\ldots,d\}}$. *Then any limit vector* $\boldsymbol{\Theta}$:

$$\boldsymbol{\Theta} := (\eta, \beta_2, \ldots, \beta_d) = \lim_{\alpha\to 1} \left(\frac{1-\alpha}{\overline{F}_1\{e^1_\alpha(\boldsymbol{X})\}}, \frac{e^2_\alpha(\boldsymbol{X})}{e^1_\alpha(\boldsymbol{X})}, \ldots, \frac{e^d_\alpha(\boldsymbol{X})}{e^1_\alpha(\boldsymbol{X})} \right)$$

satisfies the following system of equations

$$\frac{1}{\theta-1} - \eta\frac{\beta_k^\theta}{c_k} = -\sum_{i=1, i\neq k}^{d} \left\{ \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik}\left(\frac{c_i}{c_k} t^{-\theta}, 1\right) \mathrm{d}t - \eta\frac{\beta_k^{\theta-1}}{c_k}\beta_i \right\}, \; k \in \{1, \ldots, d\}$$

where λ^{ik} *is the upper tail dependence (UTD) function for the random pair* (X_i, X_k).

In particular, explicit system solutions (see Maume-Deschamps et al. (2018))

- $\boldsymbol{\Theta}^\perp = (\eta^\perp, \beta_2^\perp, \ldots, \beta_d^\perp)$ (asympt. $\perp$ case) and
- $\boldsymbol{\Theta}^+ = (\eta^+, \beta_2^+, \ldots, \beta_d^+)$ (Comon. case).

Alternative optimization problem for MEEs

Definition

Let $\mathbf{\Theta} = (\eta, \beta_2, \ldots, \beta_d)$, $\mathbf{\Lambda} = (\theta, c_2, \ldots, c_d, \lambda(\cdot))$. Define the loss function

$$L_{\mathbf{\Lambda}}(\mathbf{\Theta}) := \frac{1}{2}\|F_{\mathbf{\Lambda}}(\mathbf{\Theta})\|_2^2,$$

where

$$F_{\mathbf{\Lambda}}(\mathbf{\Theta}) = \left(F^{(1)}_{\mathbf{\Lambda}}(\mathbf{\Theta}), \ldots, F^{(d)}_{\mathbf{\Lambda}}(\mathbf{\Theta})\right) = \left(g^{(1)}_{\mathbf{\Lambda}}(\mathbf{\Theta}) + f^{(1)}_{\mathbf{\Lambda}}(\mathbf{\Theta}), \ldots, g^{(d)}_{\mathbf{\Lambda}}(\mathbf{\Theta}) + f^{(d)}_{\mathbf{\Lambda}}(\mathbf{\Theta})\right),$$

with, for all $k \in \{1, \ldots, d\}$,

$$g^{(k)}_{\mathbf{\Lambda}}(\mathbf{\Theta}) = \frac{1}{\theta - 1} - \eta\frac{\beta_k^\theta}{c_k} \quad \text{and} \quad f^{(k)}_{\mathbf{\Lambda}}(\mathbf{\Theta}) = \sum_{i \neq k}^{d}\left\{\int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik}\left(\frac{c_i}{c_k}t^{-\theta}, 1\right)\mathrm{d}t - \eta\frac{\beta_k^{\theta-1}}{c_k}\beta_i\right\}.$$

Define an optimal vector $\mathbf{\Theta}^*$, obtained by optimizing the loss function $L_{\mathbf{\Lambda}}$, *i.e.*,

$$\mathbf{\Theta}^* = \operatorname{argmin}_{\mathbf{\Theta}} L_{\mathbf{\Lambda}}(\mathbf{\Theta}).$$

Furthermore, we know that, for $\alpha \to 1$,

$$\boldsymbol{e}_\alpha(\boldsymbol{X}) \sim \mathrm{VaR}_\alpha(X_1)\eta^{1/\theta}(1, \beta_2, \ldots, \beta_d)^\top.$$

Broyden–Fletcher–Goldfarb–Shanno (BFGS) descent algorithm

To solve our optimization problem the quasi-Newton BFGS descent algorithm will be used here:

- ↪ to avoid calculating second derivatives,
- ↪ to improve computation time.

see details here

Problem

In the loss function, we have several unknown parameters:

$$g_{\Lambda}^{(k)}(\Theta) = \frac{1}{\theta - 1} - \eta \frac{\beta_k^{\theta}}{c_k}; \qquad f_{\Lambda}^{(k)}(\Theta) = \sum_{i \neq k}^{d} \left\{ \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) \mathrm{d}t - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right\}.$$

Direct application of the BFGS algorithm for the optimization problem

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}}\, L_{\Lambda}(\Theta)$$

is not possible.

Broyden–Fletcher–Goldfarb–Shanno (BFGS) descent algorithm

To solve our optimization problem the quasi-Newton BFGS descent algorithm will be used here:

- ↪ to avoid calculating second derivatives,
- ↪ to improve computation time.

see details here

Problem

In the loss function, we have several unknown parameters:

$$g_{\Lambda}^{(k)}(\Theta) = \frac{1}{\theta - 1} - \eta \frac{\beta_k^{\theta}}{c_k}; \qquad f_{\Lambda}^{(k)}(\Theta) = \sum_{i \neq k}^{d} \left\{ \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) \mathrm{d}t - \eta \frac{\beta_k^{\theta - 1}}{c_k} \beta_i \right\}.$$

Direct application of the BFGS algorithm for the optimization problem

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}}\, L_{\Lambda}(\Theta)$$

is not possible.

Broyden–Fletcher–Goldfarb–Shanno (BFGS) descent algorithm

To solve our optimization problem the quasi-Newton BFGS descent algorithm will be used here:

- ↪ to avoid calculating second derivatives,
- ↪ to improve computation time.

see details here

Problem

In the loss function, we have several unknown parameters:

$$g_{\boldsymbol{\Lambda}}^{(k)}(\boldsymbol{\Theta}) = \frac{1}{\theta - 1} - \eta \frac{\beta_k^{\theta}}{c_k}; \qquad f_{\boldsymbol{\Lambda}}^{(k)}(\boldsymbol{\Theta}) = \sum_{i \neq k}^{d} \left\{ \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) \mathrm{d}t - \eta \frac{\beta_k^{\theta - 1}}{c_k} \beta_i \right\}.$$

Direct application of the BFGS algorithm for the optimization problem

$$\boldsymbol{\Theta}^* = \underset{\boldsymbol{\Theta}}{\operatorname{argmin}}\, L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta})$$

is not possible.

Approximated Optimization Problem

Instead, one can focus on the approximate optimum

$$\underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}}\, L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta}) \qquad \Rightarrow \qquad \underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}}\, L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta})$$

for some vector of estimators $\hat{\boldsymbol{\Lambda}} = (\hat{\theta}, \hat{c}_2, \ldots, \hat{c}_d, \hat{\lambda})$.

Specifically, convergence of the estimated optimum can be shown in the following way:

1. To show that $\hat{\boldsymbol{\Lambda}} \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Lambda}$,
2. To show that $L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta})$ and $\nabla L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} \nabla L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta})$
3. To show the consistency of every iteration of the BFGS algorithm

$$\widehat{\boldsymbol{\Theta}}^k \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Theta}^k, \quad k \in \{1, 2, \ldots\}.$$

Approximated Optimization Problem

Instead, one can focus on the approximate optimum

$$\underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}}\, L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta}) \qquad \Rightarrow \qquad \underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}}\, L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta})$$

for some vector of estimators $\hat{\boldsymbol{\Lambda}} = (\hat{\theta}, \hat{c}_2, \ldots, \hat{c}_d, \hat{\lambda})$.

Specifically, convergence of the estimated optimum can be shown in the following way:

1. To show that $\hat{\boldsymbol{\Lambda}} \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Lambda}$,
2. To show that $L_{\hat{\Lambda}}(\Theta) \xrightarrow[n\to\infty]{\mathbb{P}} L_{\Lambda}(\Theta)$ and $\nabla L_{\hat{\Lambda}}(\Theta) \xrightarrow[n\to\infty]{\mathbb{P}} \nabla L_{\Lambda}(\Theta)$
3. To show the consistency of every iteration of the BFGS algorithm

$$\widehat{\boldsymbol{\Theta}}^k \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Theta}^k, \quad k \in \{1, 2, \ldots\}.$$

Approximated Optimization Problem

Instead, one can focus on the approximate optimum

$$\underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}} L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta}) \qquad \Rightarrow \qquad \underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}} L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta})$$

for some vector of estimators $\hat{\boldsymbol{\Lambda}} = (\hat{\theta}, \hat{c}_2, \ldots, \hat{c}_d, \hat{\lambda})$.

Specifically, convergence of the estimated optimum can be shown in the following way:

1 To show that $\hat{\boldsymbol{\Lambda}} \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Lambda}$,

2 To show that $L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta})$ and $\nabla L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} \nabla L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta})$

3 To show the consistency of every iteration of the BFGS algorithm

$$\widehat{\boldsymbol{\Theta}}^k \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Theta}^k, \quad k \in \{1, 2, \ldots\}.$$

Approximated Optimization Problem

Instead, one can focus on the approximate optimum

$$\underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}} L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta}) \qquad \Rightarrow \qquad \underset{\boldsymbol{\Theta}\in\mathbb{R}^d}{\operatorname{argmin}} L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta})$$

for some vector of estimators $\hat{\boldsymbol{\Lambda}} = (\hat{\theta}, \hat{c}_2, \dots, \hat{c}_d, \hat{\lambda})$.

Specifically, convergence of the estimated optimum can be shown in the following way:

1 To show that $\hat{\boldsymbol{\Lambda}} \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Lambda}$,

2 To show that $L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta})$ and $\nabla L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} \nabla L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta})$

3 To show the consistency of every iteration of the BFGS algorithm

$$\widehat{\boldsymbol{\Theta}}^k \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Theta}^k, \quad k \in \{1, 2, \dots\}.$$

Considered estimators

Parameter	Estimator
θ (tail index)	$\hat{\theta} = \frac{1}{\hat{\gamma}}$ where $\hat{\gamma} = \frac{1}{\ell_\theta} \sum_{i=1}^{\ell_\theta} \ln\left(\frac{X_{1:n-i+1,n}}{X_{1:n-\ell_\theta,n}}\right)$
c_i (tail ratio)	$\hat{c}_i = \left(\frac{X_{i:n-\ell_i+1,n}}{X_{1:n-\ell_i+1,n}}\right)^{-\hat{\theta}}, \quad i \in \{2, \dots, d\}$
$\lambda(x_i, x_k)$ (UTD function)	$\hat{\lambda}^{ik}_{\mathrm{Beta}}(x_i, x_k) = \frac{n}{\ell_\lambda} \widehat{\overline{C}}\left(\frac{\ell_\lambda}{n} x_i, \frac{\ell_\lambda}{n} x_k\right)$ with $\widehat{\overline{C}}$ survival empirical Beta Copula

where $\ell_\theta = \ell_\theta(n)$, $\ell_i = \ell_i(n)$ and $\ell_\lambda = \ell_\lambda(n)$ intermediate integer sequences.

The consistency of $\hat{\theta}$ and $\hat{c}_i$ is established, *e.g.*, in Deheuvels et al. (1988) and Maume-Deschamps et al. (2018). Furthermore, one can show:

Proposition

Taking $\hat{\Lambda} = (\hat{\theta}, \hat{c}_2, \dots, \hat{c}_d, \hat{\lambda}^{ik}_{\mathrm{Beta}})$ as in Table above, one has

$$\int_{\frac{\beta_i}{\beta_k}}^{\infty} \hat{\lambda}^{ik}_{\mathrm{Beta}}\left(\frac{\hat{c}_i}{\hat{c}_k} t^{-\hat{\theta}}, 1\right) dt \xrightarrow[n\to\infty]{\mathbb{P}} \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik}\left(\frac{c_i}{c_k} t^{-\theta}, 1\right) dt.$$

Considered estimators

Parameter	Estimator
θ (tail index)	$\hat{\theta} = \frac{1}{\hat{\gamma}}$ where $\hat{\gamma} = \frac{1}{\ell_\theta} \sum_{i=1}^{\ell_\theta} \ln(\frac{X_{1:n-i+1,n}}{X_{1:n-\ell_\theta,n}})$
c_i (tail ratio)	$\hat{c}_i = \left(\frac{X_{i:n-\ell_i+1,n}}{X_{1:n-\ell_i+1,n}}\right)^{-\hat{\theta}}, \quad i \in \{2,\dots,d\}$
$\lambda(x_i, x_k)$ (UTD function)	$\hat{\lambda}^{ik}_{\mathrm{Beta}}(x_i, x_k) = \frac{n}{\ell_\lambda} \widehat{\overline{C}}\left(\frac{\ell_\lambda}{n} x_i, \frac{\ell_\lambda}{n} x_k\right)$ with $\widehat{\overline{C}}$ survival empirical Beta Copula

where $\ell_\theta = \ell_\theta(n)$, $\ell_i = \ell_i(n)$ and $\ell_\lambda = \ell_\lambda(n)$ intermediate integer sequences.

The consistency of $\hat{\theta}$ and $\hat{c}_i$ is established, *e.g.*, in Deheuvels et al. (1988) and Maume-Deschamps et al. (2018). Furthermore, one can show:

Proposition

Taking $\hat{\Lambda} = (\hat{\theta}, \hat{c}_2, \dots, \hat{c}_d, \hat{\lambda}^{ik}_{\mathrm{Beta}})$ as in Table above, one has

$$\int_{\frac{\beta_i}{\beta_k}}^{\infty} \hat{\lambda}^{ik}_{\mathrm{Beta}}\left(\frac{\hat{c}_i}{\hat{c}_k} t^{-\hat{\theta}}, 1\right) dt \xrightarrow[n\to\infty]{\mathbb{P}} \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik}\left(\frac{c_i}{c_k} t^{-\theta}, 1\right) dt.$$

Considered estimators

Parameter	Estimator
θ (tail index)	$\hat{\theta} = \frac{1}{\hat{\gamma}}$ where $\hat{\gamma} = \frac{1}{\ell_\theta} \sum_{i=1}^{\ell_\theta} \ln(\frac{X_{1:n-i+1,n}}{X_{1:n-\ell_\theta,n}})$
c_i (tail ratio)	$\hat{c}_i = \left(\frac{X_{i:n-\ell_i+1,n}}{X_{1:n-\ell_i+1,n}}\right)^{-\hat{\theta}}, \quad i \in \{2, \dots, d\}$
$\lambda(x_i, x_k)$ (UTD function)	$\hat{\lambda}^{ik}_{\mathrm{Beta}}(x_i, x_k) = \frac{n}{\ell_\lambda} \widehat{\overline{C}} \left(\frac{\ell_\lambda}{n} x_i, \frac{\ell_\lambda}{n} x_k\right)$ with $\widehat{\overline{C}}$ survival empirical Beta Copula

where $\ell_\theta = \ell_\theta(n)$, $\ell_i = \ell_i(n)$ and $\ell_\lambda = \ell_\lambda(n)$ intermediate integer sequences.

The consistency of $\hat{\theta}$ and $\hat{c}_i$ is established, *e.g.*, in Deheuvels et al. (1988) and Maume-Deschamps et al. (2018). Furthermore, one can show:

Proposition

Taking $\hat{\mathbf{\Lambda}} = (\hat{\theta}, \hat{c}_2, \dots, \hat{c}_d, \hat{\lambda}^{ik}_{\mathrm{Beta}})$ *as in Table above, one has*

$$\int_{\frac{\beta_i}{\beta_k}}^{\infty} \hat{\lambda}^{ik}_{\mathrm{Beta}} \left(\frac{\hat{c}_i}{\hat{c}_k} t^{-\hat{\theta}}, 1\right) dt \xrightarrow[n\to\infty]{\mathbb{P}} \int_{\frac{\beta_i}{\beta_k}}^{\infty} \lambda^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1\right) dt.$$

Theorem (Limit behaviour in n)

Let $\hat{\mathbf{\Lambda}} = (\hat{\theta}, \hat{c}_2, \ldots, \hat{c}_d, \hat{\lambda}^{ik}_{\mathrm{Beta}})$ *as in Table above. Then*

$$L_{\hat{\mathbf{\Lambda}}}(\mathbf{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} L_{\mathbf{\Lambda}}(\mathbf{\Theta}) \quad \text{and} \quad \nabla L_{\hat{\mathbf{\Lambda}}}(\mathbf{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} \nabla L_{\mathbf{\Lambda}}(\mathbf{\Theta}).$$

Moreover, given identical starting values $\mathbf{\Theta}^0$, H_0, $\sigma \in (0, 1/2)$, $\rho \in (\sigma, 1)$ *and* $\epsilon \geq 0$, *for any step* k, *it holds that*

$$\widehat{\mathbf{\Theta}}^k \xrightarrow[n\to\infty]{\mathbb{P}} \mathbf{\Theta}^k.$$

We now proceed by using an iterated two-step procedure.

Firstly we provided above an adequate estimate of the true loss function (and its gradient);

Secondly proceed with the optimization procedure.

Theorem (Limit behaviour in n)

Let $\hat{\boldsymbol{\Lambda}} = (\hat{\theta}, \hat{c}_2, \dots, \hat{c}_d, \hat{\lambda}^{ik}_{\mathrm{Beta}})$ *as in Table above. Then*

$$L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta}) \quad \text{and} \quad \nabla L_{\hat{\boldsymbol{\Lambda}}}(\boldsymbol{\Theta}) \xrightarrow[n\to\infty]{\mathbb{P}} \nabla L_{\boldsymbol{\Lambda}}(\boldsymbol{\Theta}).$$

Moreover, given identical starting values $\boldsymbol{\Theta}^0$, H_0, $\sigma \in (0, 1/2)$, $\rho \in (\sigma, 1)$ *and* $\epsilon \geq 0$, *for any step* k, *it holds that*

$$\widehat{\boldsymbol{\Theta}}^k \xrightarrow[n\to\infty]{\mathbb{P}} \boldsymbol{\Theta}^k.$$

We now proceed by using an iterated two-step procedure.

Firstly we provided above an adequate estimate of the true loss function (and its gradient);

Secondly proceed with the optimization procedure.

Two-steps estimation procedure for MEEs

(Step 1) Taking the limit $n \to \infty$. Establish the consistency of $\hat{\Lambda}$ and subsequently $L_{\hat{\Lambda}}$ and $\nabla L_{\hat{\Lambda}}$. Then also the step-wise convergence of the BFGS algorithm.

(Step 2) Taking the limit $k \to \infty$. Optimize the consistently approximated problem from Step 1 using the BFGS algorithm.

Corollary (Non-exchangeable iterated limit in n and k)

Under the assumption that the BFGS algorithm solves for the global minimum, it holds that

$$\lim_{k\to\infty} \left(\operatorname{plim}_{n\to\infty} \widehat{\Theta}^k\right) = \Theta^*.$$

Two-steps estimation procedure for MEEs

(Step 1) Taking the limit $n \to \infty$. Establish the consistency of $\hat{\Lambda}$ and subsequently $L_{\hat{\Lambda}}$ and $\nabla L_{\hat{\Lambda}}$. Then also the step-wise convergence of the BFGS algorithm.

(Step 2) Taking the limit $k \to \infty$. Optimize the consistently approximated problem from Step 1 using the BFGS algorithm.

Corollary (Non-exchangeable iterated limit in n and k)

Under the assumption that the BFGS algorithm solves for the global minimum, it holds that

$$\lim_{k\to\infty} (\operatorname{plim}_{n\to\infty} \widehat{\Theta}^k) = \Theta^*.$$

Two-steps estimation procedure for MEEs

(Step 1) Taking the limit $n \to \infty$. Establish the consistency of $\hat{\Lambda}$ and subsequently $L_{\hat{\Lambda}}$ and $\nabla L_{\hat{\Lambda}}$. Then also the step-wise convergence of the BFGS algorithm.

(Step 2) Taking the limit $k \to \infty$. Optimize the consistently approximated problem from Step 1 using the BFGS algorithm.

Corollary (Non-exchangeable iterated limit in n and k)

Under the assumption that the BFGS algorithm solves for the global minimum, it holds that

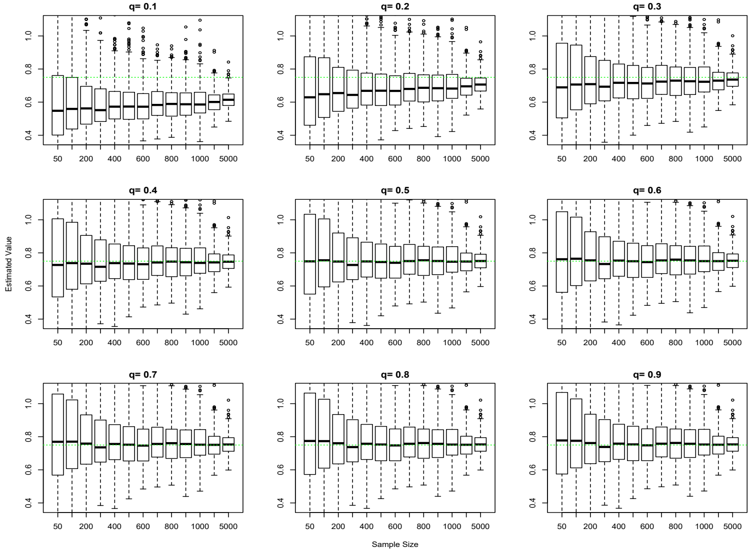
$$\lim_{k\to\infty} (\operatorname*{plim}_{n\to\infty} \widehat{\boldsymbol{\Theta}}^k) = \boldsymbol{\Theta}^*.$$

A numerical analysis

We consider a **3-dimensional** random vector with

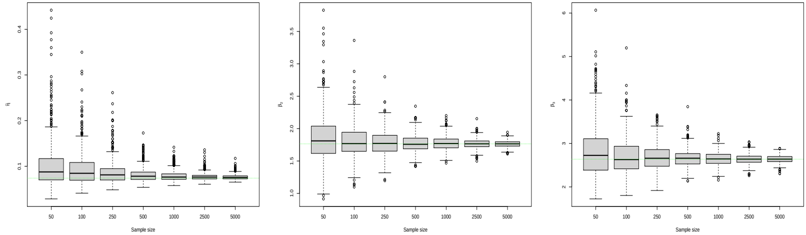
- $\hookrightarrow$ Pareto type I margins $X_i \sim \text{P}\{3.5, 1.25(1+i)\}$, $i \in \{1,2,3\}$;
- $\hookrightarrow$ Various sample sizes n;
- $\hookrightarrow$ Intermediate integer sequences $\ell_\theta = \ell_i = n^{0.75}$;
- $\hookrightarrow$ Dependence structures: independency, comonotonicity and non-trivial tail dependence structure (survival Clayton copula).

Performance of the integral of estimated UTD function

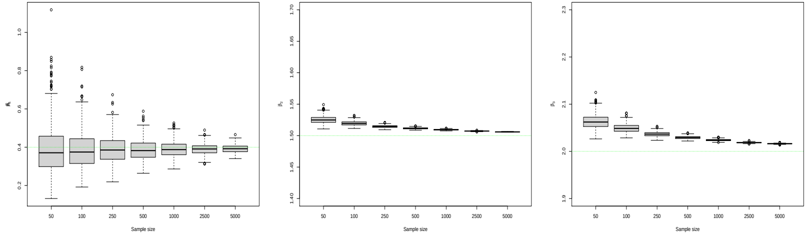


Behaviour of $\int_{\frac{\beta_2}{\beta_3}}^{\infty} \hat{\lambda}^{23}_{\text{Beta}}\left(\frac{\hat{c}_2}{\hat{c}_3} t^{-\hat{\theta}}, 1\right) dt$ for various sample sizes and subsequences $\ell_\lambda = n^q$, $q \in \{0.1, 0.2, \ldots, 0.9\}$ with several sample sizes n. The true value under **comonotonic** Pareto margins is displayed in green horizontal line.

Boxplots for the estimated solution vector

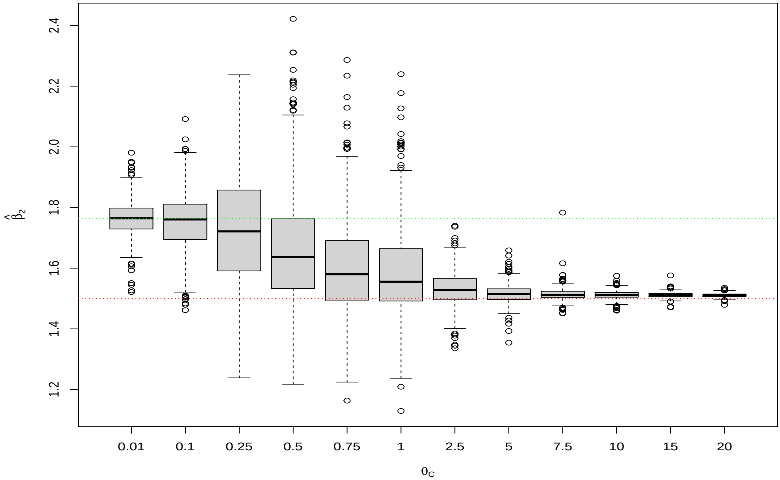


(a) Results for independent margins



(b) Results for comonotonic margins

Boxplots for the estimated solution vector for various sample sizes n with $\ell_\lambda = n^{0.50}$. Results are for $\widehat{\eta}$ (left), $\widehat{\beta}_2$ (center), $\widehat{\beta}_3$ (right). True values for **independent** $\Theta^\perp$ and **comonotonic** Θ^+ dependence structure in dashed green lines.



Results for the estimate of $\hat{\beta}_2$ under the **Pareto–Clayton model** with varying dependence parameter θ_C with $\ell_\lambda = n^{0.50}$ with $n = 5000$ and $\theta_C \in \{0.01, 0.1, 0.25, 0.5, 0.75, 1, 2.5, 5, 7.5, 10, 15, 20\}$.

Dotted lines provide true values for asymptotic independence (green) and comonotonicity (red) with $\beta_2^{\perp} = 1.764$ and $\beta_2^{+} = 1.5$, respectively.

We presented some results from :

N. Beck, E. Di Bernardino and M. Mailhot, *Semi-parametric Estimation of Multivariate Extreme Expectiles*, Journal of Multivariate Analysis, **2021**, Vol. 184, https://doi.org/10.1016/j.jmva.2021.104758 .

Possible improvements :

- ↪ Clearly for extreme multivariate expectiles it is required that $\Theta > 0$ componentwise and include $\Theta^{\perp}$ and Θ^{+} as lower and upper bounds ⇒ box-constrained BFGS algorithm (or BFGS-B).
- ↪ Incorporate limited memory storage of the inverse hessian H_k (beneficial when the dimension of the problem is large) ⇒ limited-memory box-constrained BFGS algorithm (L-BFGS-B).

Future works :

- ↪ To consider the functional conditional multidimensional L^1-expectile extension and to estimate extreme $\mathbf{e}_{\alpha}(\mathbf{X}, z)$ by using the extrapolation technique when $\alpha \to 1$.

Thank you very much for your attention!

We presented some results from :

N. Beck, E. Di Bernardino and M. Mailhot, *Semi-parametric Estimation of Multivariate Extreme Expectiles*, Journal of Multivariate Analysis, **2021**, Vol. 184, https://doi.org/10.1016/j.jmva.2021.104758 .

Possible improvements :

↪ Clearly for extreme multivariate expectiles it is required that $\mathbf{\Theta} > 0$ componentwise and include $\mathbf{\Theta}^{\perp}$ and $\mathbf{\Theta}^{+}$ as lower and upper bounds ⇒ box-constrained BFGS algorithm (or BFGS-B).

↪ Incorporate limited memory storage of the inverse hessian H_k (beneficial when the dimension of the problem is large) ⇒ limited-memory box-constrained BFGS algorithm (L-BFGS-B).

Future works :

↪ To consider the functional conditional multidimensional L^1-expectile extension and to estimate extreme $\mathbf{e}_{\alpha}(\boldsymbol{X}, z)$ by using the extrapolation technique when $\alpha \to 1$.

Thank you very much for your attention!

We presented some results from :

N. Beck, E. Di Bernardino and M. Mailhot, *Semi-parametric Estimation of Multivariate Extreme Expectiles*, Journal of Multivariate Analysis, **2021**, Vol. 184, https://doi.org/10.1016/j.jmva.2021.104758 .

Possible improvements :

↪ Clearly for extreme multivariate expectiles it is required that $\boldsymbol{\Theta} > 0$ componentwise and include $\boldsymbol{\Theta}^{\perp}$ and $\boldsymbol{\Theta}^{+}$ as lower and upper bounds ⇒ box-constrained BFGS algorithm (or BFGS-B).

↪ Incorporate limited memory storage of the inverse hessian H_k (beneficial when the dimension of the problem is large) ⇒ limited-memory box-constrained BFGS algorithm (L-BFGS-B).

Future works :

↪ To consider the functional conditional multidimensional L^1-expectile extension and to estimate extreme $e_\alpha(\boldsymbol{X}, z)$ by using the extrapolation technique when $\alpha \to 1$.

Thank you very much for your attention!

We presented some results from :

N. Beck, E. Di Bernardino and M. Mailhot, *Semi-parametric Estimation of Multivariate Extreme Expectiles*, Journal of Multivariate Analysis, **2021**, Vol. 184, https://doi.org/10.1016/j.jmva.2021.104758 .

Possible improvements :

- $\hookrightarrow$ Clearly for extreme multivariate expectiles it is required that $\boldsymbol{\Theta} > 0$ componentwise and include $\boldsymbol{\Theta}^{\perp}$ and $\boldsymbol{\Theta}^{+}$ as lower and upper bounds $\Rightarrow$ box-constrained BFGS algorithm (or BFGS-B).
- $\hookrightarrow$ Incorporate limited memory storage of the inverse hessian H_k (beneficial when the dimension of the problem is large) $\Rightarrow$ limited-memory box-constrained BFGS algorithm (L-BFGS-B).

Future works :

- $\hookrightarrow$ To consider the functional conditional multidimensional L^1-expectile extension and to estimate extreme $\boldsymbol{e}_{\alpha}(\boldsymbol{X}, z)$ by using the extrapolation technique when $\alpha \to 1$.

Thank you very much for your attention!

We presented some results from :

N. Beck, E. Di Bernardino and M. Mailhot, *Semi-parametric Estimation of Multivariate Extreme Expectiles*, Journal of Multivariate Analysis, **2021**, Vol. 184, https://doi.org/10.1016/j.jmva.2021.104758 .

Possible improvements :

- $\hookrightarrow$ Clearly for extreme multivariate expectiles it is required that $\boldsymbol{\Theta} > 0$ componentwise and include $\boldsymbol{\Theta}^{\perp}$ and $\boldsymbol{\Theta}^{+}$ as lower and upper bounds $\Rightarrow$ box-constrained BFGS algorithm (or BFGS-B).
- $\hookrightarrow$ Incorporate limited memory storage of the inverse hessian H_k (beneficial when the dimension of the problem is large) $\Rightarrow$ limited-memory box-constrained BFGS algorithm (L-BFGS-B).

Future works :

- $\hookrightarrow$ To consider the functional conditional multidimensional L^1-expectile extension and to estimate extreme $\boldsymbol{e}_\alpha(\boldsymbol{X}, z)$ by using the extrapolation technique when $\alpha \to 1$.

Thank you very much for your attention!

Definition (MRV definition)

Let $\boldsymbol{X}$ be a random vector on $\mathbb{R}^d$, the following definitions are equivalent:

The vector $\boldsymbol{X}$ has regularly varying tail of index θ.

There exists for all $x > 0$ a finite measure μ on the unit sphere $\mathbb{S}^{d-1}$, a normalizing function $b : (0, \infty) \mapsto (0, \infty)$ such that

$$\lim_{t \to +\infty} \mathbb{P}\left\{ \|\boldsymbol{X}\| > xb(t), \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in \cdot \right\} = x^{-\theta}\mu(\cdot).$$

The measure μ depends on the chosen norm, it is called the *spectral measure* of $\boldsymbol{X}$.

Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton descent algorithm

(Step 0) Put counter $k := 0$ and choose initial values $\mathbf{\Theta}^0 \in \mathbb{R}^d$, $H_0 \in \mathbb{R}^{d\times d}$ initial approximation to the inverse of the Hessian matrix of $L_{\mathbf{\Lambda}}$, $\sigma \in (0, 1/2)$, $\rho \in (\sigma, 1)$, and $\epsilon \geq 0$.

(Step 1) Let $L_{\mathbf{\Lambda}}$ as in Definition 2. If $\|\nabla L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k\right)\| \leq \epsilon$: STOP.

(Step 2) Calculate the direction $d^k = -H_k \nabla L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k\right)$.

(Step 3) Determine the step size $t_k > 0$ such that

$$L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k + t_k \boldsymbol{d}^k\right) \leq L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k\right) + \sigma t_k \nabla L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k\right),$$

$$\nabla L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k + t_k \boldsymbol{d}^k\right)^\top \boldsymbol{d}^k \geq \rho \nabla L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k\right)^\top \boldsymbol{d}^k.$$

(Step 4) Let $\boldsymbol{\rho}_k = 1/\boldsymbol{y}_k^\top \boldsymbol{s}_k$. Update the following:

- $\mathbf{\Theta}^{k+1} := \mathbf{\Theta}^k + t_k d^k$
- $\boldsymbol{y}^k := \nabla L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^{k+1}\right) - \nabla L_{\mathbf{\Lambda}}\left(\mathbf{\Theta}^k\right)$,
- $\boldsymbol{s}^k := \mathbf{\Theta}^{k+1} - \mathbf{\Theta}^k$
- $H_{k+1} := \left(\mathbb{I} - \boldsymbol{\rho}_k \boldsymbol{s}_k \boldsymbol{y}_k^\top\right) H_k \left(\mathbb{I} - \boldsymbol{\rho}_k \boldsymbol{s}_k \boldsymbol{y}_k^\top\right) + \boldsymbol{\rho}_k \boldsymbol{s}_k \boldsymbol{s}_k^\top$.

(Step 5) Set $k \leftarrow k + 1$ and go to **(Step 1)**.

back to main slides