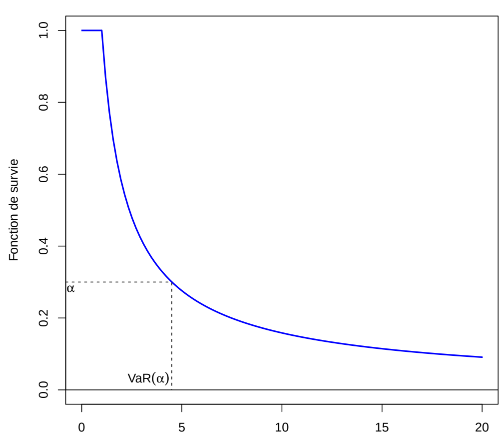
Nonparametric estimation of tail risk measures from heavy-tailed distributions

Jonthan El Methni, Laurent Gardes & Stéphane Girard

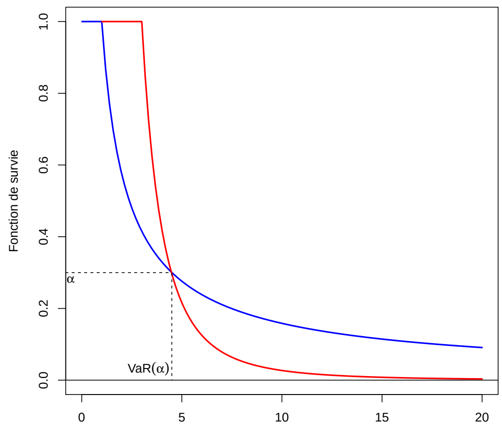
1 Tail risk measures

Let $Y \in \mathbb{R}$ be a real random loss variable. The Value-at-Risk of level $\alpha \in (0,1)$ denoted by $\mathrm{VaR}(\alpha)$ is the α-quantile of the survival function $\bar{F}(x) = P(Y > x)$

$$\mathrm{VaR}(\alpha) := \overline{F}^{-1}(\alpha) = \inf\{t, \overline{F}(t) \leq \alpha\}$$

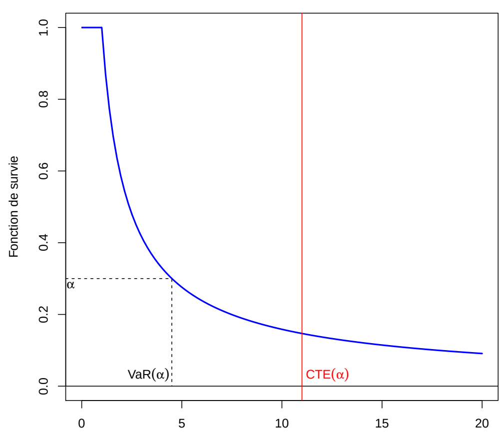


Consider Y_1 and Y_2 two random loss variables with associated survival functions $\overline{F}_1$ and $\overline{F}_2$.



Random variables with light tail probabilities and with heavy tail probabilities may have the same VaR(α). This is one of the main criticisms against the VaR, Embrechts et al. [1997]. The Conditional Tail Expectation of level $\alpha \in (0,1)$ denoted by CTE(α) is defined by

$$\mathrm{CTE}(\alpha) := \mathbb{E}(Y|Y > \mathrm{VaR}(\alpha)).$$



The CTE takes into account the whole information contained in the upper part of the tail distribution.

- Let $Y \in \mathbb{R}$ be a real random loss variable. The Value-at-Risk of level $\alpha \in (0,1)$ is the α-quantile defined by

$$\mathrm{VaR}(\alpha) := \overline{F}^{-1}(\alpha),$$

where $\overline{F}^{-1}$ is the (generalized) inverse of the survival function of Y.

- The Conditional Tail Expectation of level $\alpha \in (0,1)$ is defined by

$$\mathrm{CTE}(\alpha) := \mathbb{E}(Y|Y > \mathrm{VaR}(\alpha)).$$

- The Conditional-Value-at-Risk of level $\alpha \in (0,1)$ introduced by Rockafellar et Uryasev [2000] is defined by

$$\mathrm{CVaR}_\lambda(\alpha) := \lambda \mathrm{VaR}(\alpha) + (1-\lambda)\mathrm{CTE}(\alpha),$$

with $0 \leq \lambda \leq 1$.

- The Conditional Tail Variance of level $\alpha \in (0,1)$ introduced by Valdez [2005] is defined by

$$\mathrm{CTV}(\alpha) := \mathbb{E}((Y - \mathrm{CTE}(\alpha))^2 | Y > \mathrm{VaR}(\alpha)).$$

The first goal of this work is to unify the definitions of the previous risk measures. To this end, the Conditional Tail Moment of level $\alpha \in (0,1)$ is introduced:

$$\mathrm{CTM}_a(\alpha) := \mathbb{E}(Y^a | Y > \mathrm{VaR}(\alpha)),$$

where $a \geq 0$ is such that the moment of order a of Y exists. All the previous risk measures of level α can be rewritten as

$$\begin{aligned}
\mathrm{CTE}(\alpha) &= \mathrm{CTM}_1(\alpha), \\
\mathrm{CVaR}(\alpha) &= \lambda \mathrm{VaR}(\alpha) + (1-\lambda)\mathrm{CTM}_1(\alpha), \\
\mathrm{CTV}(\alpha) &= \mathrm{CTM}_2(\alpha) - \mathrm{CTM}_1^2(\alpha).
\end{aligned}$$

All the risk measures depend on the VaR and the CTM_a. Our second aim is to estimate these risk measures in case of extreme losses and in the case where a covariate $X \in \mathbb{R}^p$ is recorded simultaneously with Y.

The fixed level $\alpha \in (0,1)$ is replaced by a sequence $\alpha_n \underset{n\to\infty}{\to} 0$.

Denoting by $\overline{F}(.|x)$ the conditional survival distribution function of Y given $X = x$, the Regression Value-at Risk is defined by:

$$\mathrm{RVaR}(\alpha_n|x) := \overline{F}^{-1}(\alpha_n|x) = \inf\{t, \overline{F}(t|x) \leq \alpha_n\},$$

and the Regression Conditional Tail Moment of order a is defined by:

$$\mathrm{RCTM}_a(\alpha_n|x) := \mathbb{E}(Y^a|Y > \mathrm{RVaR}(\alpha_n|x), X = x),$$

where $a > 0$ is such that the moment of order a of Y exists.This yields the following risk measures:

$$\begin{aligned}
\mathrm{RCTE}(\alpha_n|x) &= \mathrm{RCTM}_1(\alpha_n|x), \\
\mathrm{RCVaR}_\lambda(\alpha_n|x) &= \lambda\mathrm{RVaR}(\alpha_n|x) + (1-\lambda)\mathrm{RCTM}_1(\alpha_n|x), \\
\mathrm{RCTV}_n(\alpha_n|x) &= \mathrm{RCTM}_2(\alpha_n|x) - \mathrm{RCTM}_1^2(\alpha_n|x).
\end{aligned}$$

All the risk measures depend on the RVaR and the RCTM_a. The conditional moment of order $a \geq 0$ of Y given $X = x$ is defined by

$$\varphi_a(y|x) = \mathbb{E}\left(Y^a \mathbb{I}\{Y > y\}|X = x\right),$$

where $\mathbb{I}\{.\}$ is the indicator function. Since $\varphi_0(y|x) = \overline{F}(y|x)$, it follows

$$\begin{aligned}
\mathrm{RVaR}(\alpha_n|x) &= \varphi_0^{-1}(\alpha_n|x), \\
\mathrm{RCTM}_a(\alpha_n|x) &= \frac{1}{\alpha_n}\varphi_a(\varphi_0^{-1}(\alpha_n|x)|x).
\end{aligned}$$

Goal: estimate $\varphi_a(.|x)$ and $\varphi_a^{-1}(.|x)$.

2 Estimators and asymptotic results

Estimator of $\varphi_a(.|x)$ **:** We propose to use a classical kernel estimator given by

$$\widehat{\varphi}_{a,n}(y|x) = \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right) Y_i^a \mathbb{I}\{Y_i > y\} \Bigg/ \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right).$$

- h_n is a sequence called the window-width such that $h_n \to 0$ as $n \to \infty$,
- K is a bounded density on $\mathbb{R}^p$ with support included in the unit ball of $\mathbb{R}^p$.

Estimator of $\varphi_a^{-1}(.|x)$ **:** Since $\hat{\varphi}_{a,n}(.|x)$ is a non-increasing function, an estimator of $\varphi_a^{-1}(\alpha|x)$ can be defined for $\alpha \in (0, 1)$ by

$$\hat{\varphi}_{a,n}^{-1}(\alpha|x) = \inf\{t, \hat{\varphi}_{a,n}(t|x) < \alpha\}.$$

Assumptions

(F.1) The conditional survival distribution function of Y given $X = x$ is assumed to be heavy-tailed *i.e.* for all $\lambda > 0$,

$$\lim_{y\to\infty} \frac{\overline{F}(\lambda y|x)}{\overline{F}(y|x)} = \lambda^{-1/\gamma(x)}.$$

In this context, $\gamma(.)$ is a positive function of the covariate x and is referred to as the conditional tail index since it tunes the tail heaviness of the conditional distribution of Y given $X = x$. Condition **(F.1)** also implies that for $a \in [0, 1/\gamma(x))$, $\mathrm{RCTM}_a(.|x)$ exists, and for all $y > 0$,

$$\mathrm{RCTM}_a(1/y|x) = y^{a\gamma(x)}\ell_a(y|x),$$

where for x fixed, $\ell_a(.|x)$ is a slowly-varying function *i.e.* for all $\lambda > 0$,

$$\lim_{y\to\infty} \frac{\ell_a(\lambda y|x)}{\ell_a(y|x)} = 1.$$

(F.2) $\ell_a(.|x)$ is normalized for all $a \in [0, 1/\gamma(x))$.

In such a case, the Karamata representation of the slowly-varying function can be written as

$$\ell_a(y|x) = c_a(x) \exp\left(\int_1^y \frac{\varepsilon_a(u|x)}{u} du\right),$$

where $c_a(.)$ is a positive function and $\varepsilon_a(y|x) \to 0$ as $y \to \infty$.

(F.3) $|\varepsilon_a(.|x)|$ is continuous and ultimately non-increasing for all $a \in [0, 1/\gamma(x))$.

A Lipschitz condition on the probability density function g of X is also required:

(L) There exists a constant $c_g > 0$ such that $|g(x) - g(x')| \le c_g d(x, x')$.

where $d(x, x')$ is the Euclidean distance between x and x'. Finally, for $y > 0$ and $\xi > 0$, the largest oscillation of the conditional moment of order $a \in [0, 1/\gamma(x))$ is defined by

$$\omega(x, y, a, \xi, h_n) = \sup\left\{\left|\frac{\varphi_a(z|x)}{\varphi_a(z|x')} - 1\right|,\ z \in [(1-\xi)y, (1+\xi)y],\ x' \in B(x, h_n)\right\},$$

where $B(x, h_n)$ denotes the ball centred at x with radius h_n.

Theorem 1 *Suppose* **(F.1)**, **(F.2)** *and* **(L)** *hold. Let*

- $0 \le a_1 < a_2 < \cdots < a_J$,
- $x \in \mathbb{R}^p$ *such that* $g(x) > 0$ *and* $0 < \gamma(x) < 1/(2a_J)$,
- $\alpha_n \to 0$ *and* $nh_n^p\alpha_n \to \infty$ *as* $n \to \infty$,
- $\xi > 0$ *such that* $\sqrt{nh_n^p\alpha_n}\,(h_n \vee \max_a \omega(x, \mathrm{RVaR}(\alpha_n|x), a, \xi, h_n)) \to 0$,

Then,

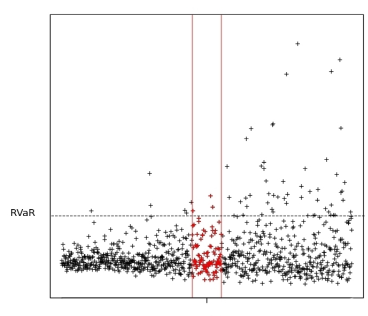
$$\sqrt{nh_n^p\alpha_n}\left\{\left(\frac{\widehat{\mathrm{RCTM}}_{a_j,n}(\alpha_n|x)}{\mathrm{RCTM}_{a_j}(\alpha_n|x)} - 1\right)_{j\in\{1,\ldots,J\}}, \left(\frac{\widehat{\mathrm{RVaR}}_n(\alpha_n|x)}{\mathrm{RVaR}(\alpha_n|x)} - 1\right)\right\}$$

is asymptotically Gaussian, centered, with covariance matrix $\|K\|_2^2\gamma^2(x)\Sigma(x)/g(x)$ *where*

$$\Sigma(x) = \left(\begin{array}{c|c} \frac{a_ia_j(2-(a_i+a_j)\gamma(x))}{(1-(a_i+a_j)\gamma(x))} & \begin{array}{c} a_1 \\ \vdots \\ a_J \end{array} \\ \hline a_1 \cdots a_J & 1 \end{array}\right).$$

Conditions on the sequences α_n and h_n.

$nh_n^p\alpha_n \to \infty$: Necessary and sufficient condition for the almost sure presence of at least one point in the region $B(x, h_n) \times [\mathrm{RVaR}(\alpha_n|x), +\infty)$ of $\mathbb{R}^p \times \mathbb{R}$.



$\sqrt{nh_n^p\alpha_n}\,(h_n \vee \max_a \omega(x, \mathrm{RVaR}(\alpha_n|x), a, \xi, h_n)) \to 0$: The biais induced by the smoothing is negligible compared to the standard-deviation.

Corollary 1 *Suppose the assumptions of Theorem 1 hold. Then, if* $0 < \gamma(x) < 1/2$*,*

$$\sqrt{nh_n^p\alpha_n}\left(\frac{\widehat{\mathrm{RCTE}}_n(\alpha_n|x)}{\mathrm{RCTE}(\alpha_n|x)}-1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{2(1-\gamma(x))\gamma^2(x)}{1-2\gamma(x)}\frac{\|K\|_2^2}{g(x)}\right)$$

$$\sqrt{nh_n^p\alpha_n}\left(\frac{\widehat{\mathrm{RCVaR}}_{\lambda,n}(\alpha_n|x)}{\mathrm{RCVaR}_\lambda(\alpha_n|x)}-1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\gamma^2(x)(\lambda^2+2-2\lambda-2\gamma(x))}{1-2\gamma(x)}\frac{\|K\|_2^2}{g(x)}\right)$$

The $\mathrm{RCTV}(\alpha_n|x)$ *estimator involves the computation of a second order moment, it requires the stronger condition* $0 < \gamma(x) < 1/4$*,*

$$\sqrt{nh_n^p\alpha_n}\left(\frac{\widehat{\mathrm{RCTV}}_n(\alpha_n|x)}{\mathrm{RCTV}(\alpha_n|x)}-1\right) \xrightarrow{d} \mathcal{N}\left(0, V_{\gamma(x)}\frac{\|K\|_2^2}{g(x)}\right),$$

where

$$V_{\gamma(x)} = \frac{8(1-\gamma(x))(1-2\gamma(x))(1+2\gamma(x)+3\gamma^2(x))}{(1-3\gamma(x))(1-4\gamma(x))}.$$

3 Illustration on a simulated dataset

A sample $\{(X_i, Y_i), i = 1, \ldots, n\}$ of size $n = 1000$ is generated. The covariate X is uniform on $[0, 1]$. The conditional distribution of $Y|X = x$ is chosen in the Hall class :

$$\overline{F}(y|x) = y^{-1/\gamma(x)} \underbrace{a(1+by^{\rho/\gamma(x)})}_{\ell(y|x)}$$

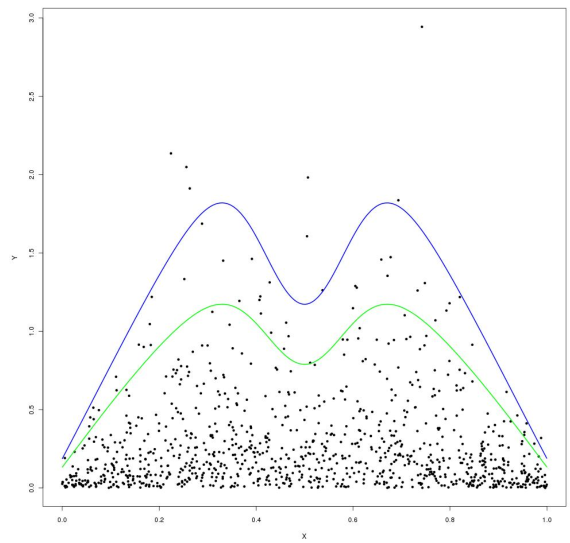
with $a = 1/2$, $b = 1$, $\rho = -1$ and conditional tail index function

$$x \in [0,1] \to \gamma(x) = \frac{1}{2}\left(\frac{1}{10}+\sin(\pi x)\right)\left(\frac{11}{10}-\frac{1}{2}\exp\left(64\left(x-\frac{1}{2}\right)^2\right)\right).$$

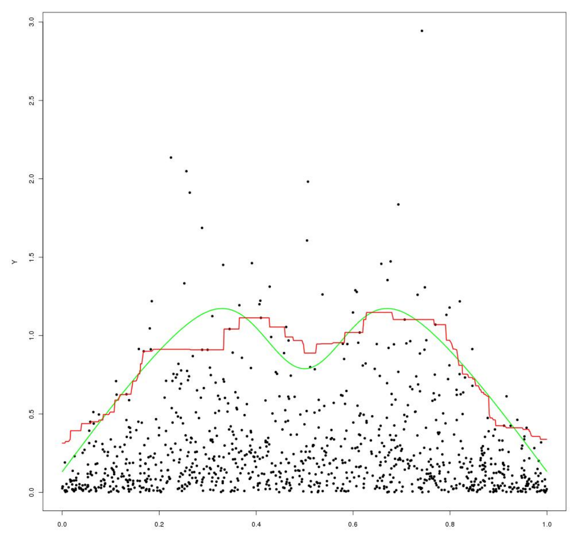
We have chosen a bi-quadratic kernel

$$K(u) \propto (1-u^2)^2\mathbb{I}_{\{|u|\leq 1\}}$$

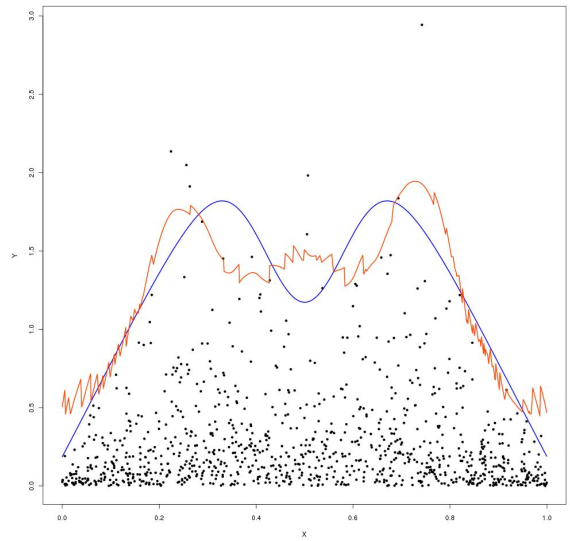
with smoothing parameter $h_n = 0.1$. Our goal if to estimate $\mathrm{RCTE}(\alpha_n|x)$ and $\mathrm{RVaR}(\alpha_n|x)$ with $\alpha_n = 0.05$.



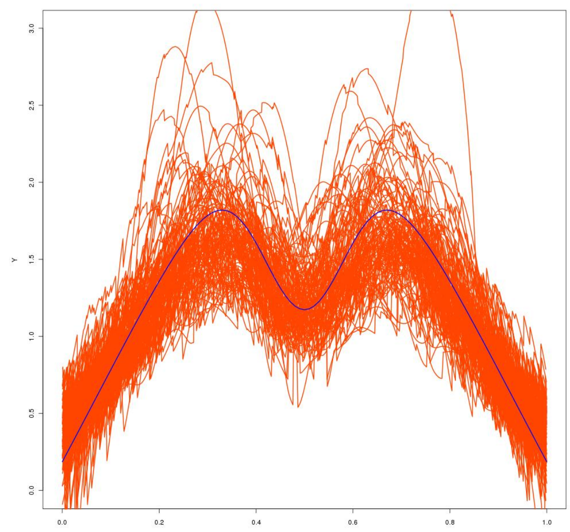
Theoretical RCTE($\alpha_n|x$) and RVaR($\alpha_n|x$) with a logarithmic scale.



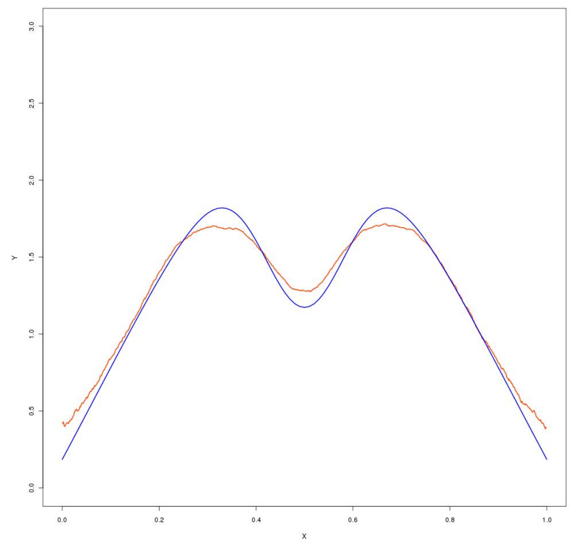
Theoretical and estimated RVaR($\alpha_n|x$) with a logarithmic scale.



Theoretical and estimated RCTE($\alpha_n|x$) with a logarithmic scale.



Theoretical and estimated RCTE($\alpha_n|x$) with a logarithmic scale.



Theoretical and estimated $\mathrm{RCTE}(\alpha_n|x)$ with a logarithmic scale.

4 Extrapolation: A Weissman type estimator

- In Theorem 1, the condition $nh_n^p\alpha_n \to \infty$ provides a lower bound on the level of the risk measure to estimate.
- This restriction is a consequence of the use of a kernel estimator which cannot extrapolate beyond the maximum observation in the ball $B(x, h_n)$.
- In consequence, α_n must be an order of an extreme quantile within the sample.

Let us consider $(\alpha_n)_{n\geq 1}$ and $(\beta_n)_{n\geq 1}$ two positive sequences such that $\alpha_n \to 0$, $\beta_n \to 0$ and $0 < \beta_n < \alpha_n$. A kernel adaptation of Weissman's estimator [1978] is given by

$$\widehat{\mathrm{RCTM}}^{W}_{a,n}(\beta_n|x) = \widehat{\mathrm{RCTM}}_{a,n}(\alpha_n|x)\underbrace{\left(\frac{\alpha_n}{\beta_n}\right)^{a\hat{\gamma}_n(x)}}_{\text{extrapolation}}$$

Theorem 2 *Suppose the assumptions of Theorem 1 hold together with* **(F.3)**. *Let $\hat{\gamma}_n(x)$ be an estimator of the conditional tail index such that*

$$\sqrt{nh_n^p\alpha_n}(\hat{\gamma}_n(x) - \gamma(x)) \xrightarrow{d} \mathcal{N}\left(0, v^2(x)\right),$$

with $v(x) > 0$. If, moreover $(\beta_n)_{n\geq 1}$ is a positive sequence such that $\beta_n \to 0$ and $\beta_n/\alpha_n \to 0$ as $n \to \infty$, then

$$\frac{\sqrt{nh_n^p\alpha_n}}{\log(\alpha_n/\beta_n)}\left(\frac{\widehat{\mathrm{RCTM}}^W_{a,n}(\beta_n|x)}{\mathrm{RCTM}_a(\beta_n|x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, (av(x))^2\right).$$

The condition $\beta_n/\alpha_n \to 0$ allows us to extrapolate and choose a level β_n arbitrarily small.

Estimation of the conditional tail index

- Without covariate : Hill [1975]. Let $(k_n)_{n\geq 1}$ be a sequence of integers such that $k_n \in \{1 \ldots n\}$. The Hill estimator is given by

$$\hat{\gamma}_{n,\alpha_n} = \frac{1}{k_n - 1}\sum_{i=1}^{k_n-1} \log Z_{n-i+1,n} - \log Z_{n-k_n+1,n},$$

  where $Z_{1,n} \leq \cdots \leq Z_{n,n}$ are the order statistics associated with i.i.d. random variables $Z_1, \ldots, Z_n$ and $\alpha_n = k_n/n$.

- With a covariate : A kernel version of the Hill estimator is given by

$$\hat{\gamma}_{n,\alpha_n}(x) = \sum_{j=1}^{J}(\log \widehat{\mathrm{RVaR}}_n(\tau_j\alpha_n|x) - \log \widehat{\mathrm{RVaR}}_n(\tau_1\alpha_n|x)) \Bigg/ \sum_{j=1}^{J}\log(\tau_1/\tau_j),$$

  where $J \geq 1$ and $(\tau_j)_{j\geq 1}$ is a decreasing sequence of weights.

The asymptotic normality of $\hat{\gamma}_{n,\alpha_n}(x)$ has been established by Daouia et al. [2011] . As a consequence, replacing $\widehat{\mathrm{RVaR}}_n$ by $\widehat{\mathrm{RVaR}}_n^W$ and $\widehat{\mathrm{RCTM}}_{a,n}$ by $\widehat{\mathrm{RCTM}}_{a,n}^W$ provides (asymptotically Gaussian) estimators for all the risk measures considered here, and for arbitrarily small levels. In particular, since $\mathrm{RCTE}(\alpha_n|x) = \mathrm{RCTM}_1(\alpha_n|x)$, we obtain

$$\widehat{\mathrm{RCTE}}_n^W(\beta_n|x) = \widehat{\mathrm{RCTE}}_n(\alpha_n|x)\left(\frac{\alpha_n}{\beta_n}\right)^{\hat{\gamma}_n(x)}.$$

References

[1] Artzner, P., Delbaen, F., Eber, J. M. et Heath, D. (1999). Coherent measures of risk, *Mathematical finance*, **9(3)**, 203–228.

[2] Brazaukas, V., Jones, B. L., Puri, M. L. et Zitikis, R. (2008). Estimating conditional tail expectation with actuarial applications in view, *Journal of Statistical Planning and Inference*, **128**, 3590–3604.

[3] Carreau, J., Ceresetti, D., Ursu, E., Anquetin, E., Creutin, J.D., Gardes, L., Girard,S. and Molinié, G. (2012). Evaluation of classical spatial-analysis schemes of extreme rainfall, *Natural Hazards and Earth System Sciences*, **12**, 3229–3240.

[4] Collomb, G. (1980). Estimation non paramétrique de probabilités conditionnelles, *Comptes Rendus de l'Académie des Sciences*, **291**, 427–430.

[5] Daouia, A., Gardes, L., Girard, S. and Lekina, A. (2011). Kernel estimators of extreme level curves, *Test*, **20**, 311–333.

[6] Daouia, A., Gardes, L. and Girard, S. (2013). On kernel smoothing for extremal quantile regression, *Bernoulli*, **19**, 2557–2589.

[7] Deme, E., Girard, S. and Guillou, A. (2015). Reduced-bias estimators of the Conditional Tail Expectation for heavy-tailed distributions, In M. Hallin et al, editors, *Mathematical Statistics and Limit Theorems*, pages 105–123, Springer.

[8] Deme, E., Girard, S. and Guillou, A. (2013). Reduced-bias estimator of the Proportional Hazard Premium for heavy-tailed distributions, *Insurance: Mathematics and Economics*, **52**, 550–559.

[9] Embrechts, P., Kluppelberg, C. and Mikosh, T. (1997). Modelling Extremal Events, *Springer*.

[10] El Methni, J., Gardes, L. and Girard, S. (2014). Nonparametric estimation of extreme risks from conditional heavy-tailed distributions, *Scandinavian Journal of Statistics*, 41, 988–1012.

[11] Gardes, L. and Girard, S. (2008). A moving window approach for nonparametric estimation of the conditional tail index, *Journal of Multivariate Analysis*, **99**, 2368–2388.

[12] Gardes, L. and Girard, S. (2010). Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels, *Extremes*, **13**, 177–204.

[13] Girard, S. and Stupfler, G. (2015). Extreme geometric quantiles in a multivariate regular variation framework, *Extremes*, **18**, 629–663.

[14] Girard, S. and Stupfler, G. (2016). Intriguing properties of extreme geometric quantiles", REVSTAT - Statistical Journal, to appear.

[15] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, **3**, 1163–1174.

[16] Rockafellar, R.T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, **2**, 21–42.

[17] Valdez, E.A. (2005). Tail conditional variance for elliptically contoured distributions, *Belgian Actuarial Bulletin*, **5**, 26–36.

[18] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association*, **73**, 812–815.

[19] Necir, A., Rassoul, A. et Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses, *Journal of Probability and Statistics*, **ID 596839**, 17 pages.