Kernel estimation of extreme regression risk measures

by

Jonathan EL METHNI

in collaboration with

Stéphane GIRARD & Laurent GARDES

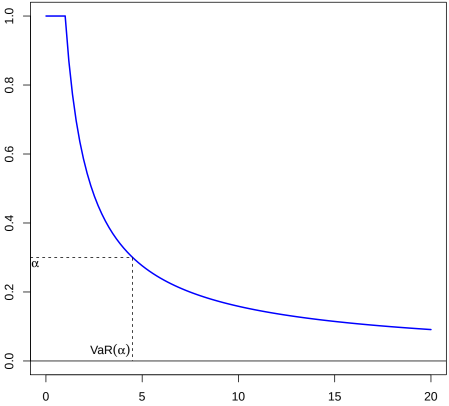
13th International Conference on Operations Research
Havana Cuba, March 2018.

The Value-at-Risk

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

$$\mathrm{VaR}(\alpha) := \overline{F}^{\leftarrow}(\alpha) = \inf\{y, \overline{F}(y) \leq \alpha\},$$

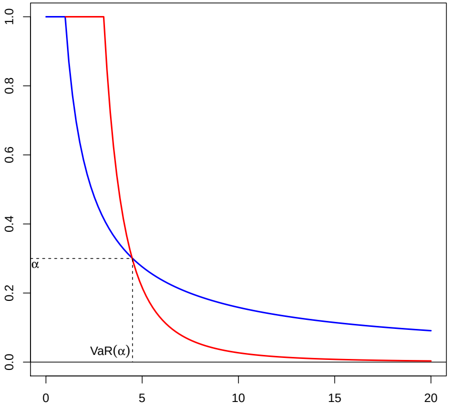
where $\overline{F}^{\leftarrow}$ is the generalized inverse of the survival function $\overline{F}(y) = \mathbb{P}(Y \geq y)$ of Y.



- The $\mathrm{VaR}(\alpha)$ is the quantile of level α of the survival function of the *r.v.* Y.

Drawbacks of the Value-at-Risk

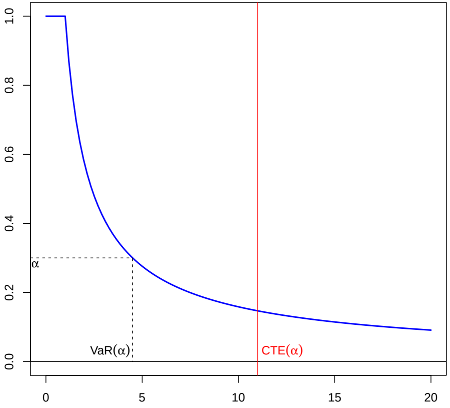
- Let us consider Y_1 and Y_2 two loss *r.v.* with associated survival function $\overline{F}_1$ and $\overline{F}_2$.



$\Longrightarrow$ Random variables with light tail probabilities and with heavy tail probabilities may have the same VaR(α). This is one of the main criticism against VaR as a risk measure (Embrechts *et al.* [1997]).

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

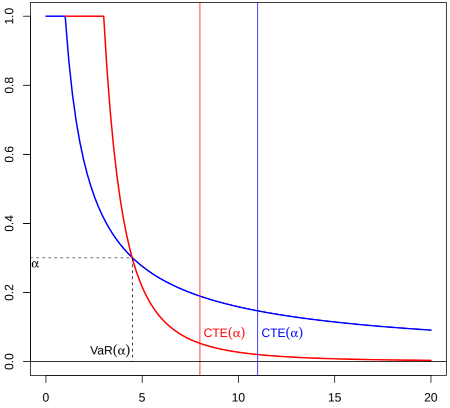
$$\mathrm{CTE}(\alpha) := \mathbb{E}(Y|Y > \mathrm{VaR}(\alpha)).$$



$\Longrightarrow$ The CTE(α) takes into account the whole information contained in the upper part of the tail distribution.

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

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

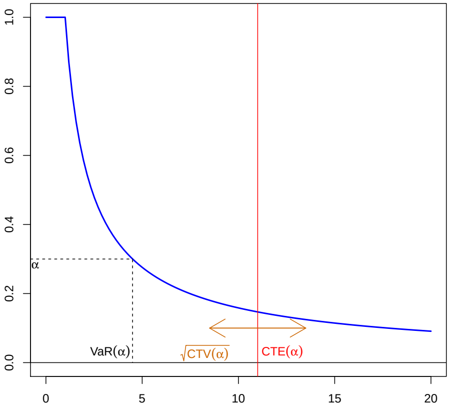


$\Longrightarrow$ The CTE(α) takes into account the whole information contained in the upper part of the tail distribution.

The Conditional Tail Variance

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

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



$\Longrightarrow$ The $\mathrm{CTV}(\alpha)$ measures the conditional variability of Y given that $Y > \mathrm{VaR}(\alpha)$ and indicates how far away the events deviate from $\mathrm{CTE}(\alpha)$.

- The Conditional Tail Skewness of level $\alpha \in (0,1)$ denoted $\mathrm{CTS}(\alpha)$ and introduced by Hong and Elshahat [2010] is defined by

$$\mathrm{CTS}(\alpha) := \frac{\mathbb{E}(Y^3 | Y > \mathrm{VaR}(\alpha))}{(\mathrm{CTV}(\alpha))^{3/2}}$$

  The CTS evaluates the asymmetry of the distribution above the VaR.

$\Longrightarrow$ We can unify the definitions of the previous risk measures using the Conditional Tail Moment introduced by El Methni *et al.* [2014].

Definition

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

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

where $b \geq 0$ is such that the moment of order b of Y exists.

All the previous risk measures of level α can be rewritten as

Risk Measure	Rewritten Risk Measure
$\mathsf{CTE}(\alpha) = \mathbb{E}(Y \mid Y > \mathrm{VaR}(\alpha))$	$\mathrm{CTM}_1(\alpha)$
$\mathsf{CTV}(\alpha) = \mathbb{E}((Y - \mathrm{CTE}(\alpha))^2 \mid Y > \mathrm{VaR}(\alpha))$	$\mathrm{CTM}_2(\alpha) - \mathrm{CTM}_1^2(\alpha)$
$\mathsf{CTS}(\alpha) = \mathbb{E}(Y^3 \mid Y > \mathrm{VaR}(\alpha))/(\mathrm{CTV}(\alpha))^{3/2}$	$\mathrm{CTM}_3(\alpha)/(\mathrm{CTV}(\alpha))^{3/2}$

$\Longrightarrow$ All the risk measures depend on the $\mathrm{CTM}_b(\alpha)$.

$\Longrightarrow$ Our contributions consist in adding two difficulties in the framework of the estimation of risk measures.

1. First we add the presence of a random covariate $X \in \mathbb{R}^p$.

- Y is a positive random variable and $X \in \mathbb{R}^p$ a random vector of regressors recorded simultaneously with Y.
- In what follows, it is assumed that (X, Y) is a continuous random vector.
- The probability density function (p.d.f.) of X is denoted by $g(\cdot)$.
- The conditional p.d.f. of Y given $X = x$ is denoted by $f(\cdot|x)$.

For any $x \in \mathbb{R}^p$ such that $g(x) \neq 0$, the conditional distribution of Y given $X = x$ is characterized by the conditional survival function

$$\overline{F}(\cdot|x) = \mathbb{P}(Y > \cdot|X = x)$$

or, equivalently, by the Regression Value at Risk defined for $\alpha \in (0, 1)$ by

$$\mathrm{RVaR}(\alpha|x) := \overline{F}^{\leftarrow}(\alpha|x) = \inf\{t, \overline{F}(t|x) \leq \alpha\}.$$

The Regression Value at Risk of level α is a generalization to a regression setting of the Value at Risk.

The Regression Conditional Tail Moment of order b is defined by

$$\mathrm{RCTM}_b(\alpha|x) := \mathbb{E}(Y^b|Y > \mathrm{RVaR}(\alpha|x), X = x),$$

where $b \geq 0$ is such that the moment of order b of Y exists.

2. Second we are interested in the estimation of risk measures in the case of extreme losses.

$\Longrightarrow$ To this end, we replace the fixed order $\alpha \in (0,1)$ by a sequence $\alpha_n \to 0$ as the sample size $n \to \infty$.

$$\begin{aligned} \mathrm{RVaR}(\alpha_n|x) &:= \overline{F}^{\leftarrow}(\alpha_n|x) \\ \mathrm{RCTM}_b(\alpha_n|x) &:= \mathbb{E}(Y^b|Y > \mathrm{RVaR}(\alpha_n|x), X = x) \end{aligned}$$

$\Longrightarrow$ All the risk measures depend on the $\mathrm{RCTM}_b(\alpha|x)$.

$$\begin{aligned} \mathrm{RCTE}(\alpha_n|x) &= \mathrm{RCTM}_1(\alpha_n|x), \\ \mathrm{RCTV}(\alpha_n|x) &= \mathrm{RCTM}_2(\alpha_n|x) - \mathrm{RCTM}_1^2(\alpha_n|x), \\ \mathrm{RCTS}(\alpha_n|x) &= \mathrm{RCTM}_3(\alpha_n|x)/(\mathrm{RCTV}(\alpha_n|x))^{3/2}. \end{aligned}$$

Starting from n independent copies $(X_1, Y_1), \ldots, (X_n, Y_n)$ of the random vector (X, Y), we address here the estimation of the Regression Conditional Tail Moment of level α_n and order $b \geq 0$ given by

$$\mathrm{RCTM}_b(\alpha_n|x) := \frac{1}{\alpha_n}\mathbb{E}\left(Y^b\mathbb{I}\{Y > \mathrm{RVaR}(\alpha_n|x)\}|X = x\right),$$

where b is such that the moment of order b of Y exits and $\mathbb{I}\{\cdot\}$ is the indicator function.

$\Longrightarrow$ We want to estimate all the above mentioned risk measures.

To do it, we need the asymptotic joint distribution of

$$\left\{\left(\widehat{\mathrm{RCTM}}_{b_j,n}(\alpha_n|x),\ j = 1, \ldots, J\right)\right\},$$

with $0 \leq b_1 < \ldots < b_J$ and where J is an integer.

Estimator of the RVaR

The estimator of the Regression Value at Risk of level α_n considered is given by

$$\widehat{\mathrm{RVaR}}_n(\alpha_n|x) = \inf\{t,\ \hat{\bar{F}}_n(t|x) \leq \alpha_n\}$$

with

$$\hat{\bar{F}}_n(y|x) = \frac{\sum_{i=1}^n \mathcal{K}_{k_n}(x - X_i)\mathbb{I}\{Y_i > y\}}{\sum_{i=1}^n \mathcal{K}_{k_n}(x - X_i)}.$$

- The bandwidth (k_n) is a non random sequence converging to 0 as $n \to \infty$.
- It controls the smoothness of the kernel estimator.
- For $z > 0$, we have also introduced the notation $\mathcal{K}_z(\cdot) = z^{-p}\mathcal{K}(\cdot/z)$ where $\mathcal{K}(\cdot)$ is a density on $\mathbb{R}^p$.
- The estimation of the $\mathrm{RVaR}(\alpha_n|x)$ has been addressed for instance by Daouia *et al.* [2013].

The estimator of the Regression Conditional Tail Moment of level α_n and order b is given by

$$\widehat{\mathrm{RCTM}}_{b,n}(\alpha_n|x) = \frac{1}{\alpha_n} \frac{\sum_{i=1}^n \mathcal{K}_{h_n}(x - X_i) Y_i^b \mathbb{I}\{Y_i > \widehat{\mathrm{RVaR}}_n(\alpha_n|x)\}}{\sum_{i=1}^n \mathcal{K}_{h_n}(x - X_i)}$$

where

$$\widehat{\mathrm{RVaR}}_n(\alpha_n|x) = \inf\{t,\ \hat{\bar{F}}_n(t|x) \leq \alpha_n\}$$

with

$$\hat{\bar{F}}_n(y|x) = \frac{\sum_{i=1}^n \mathcal{K}_{k_n}(x - X_i)\mathbb{I}\{Y_i > y\}}{\sum_{i=1}^n \mathcal{K}_{k_n}(x - X_i)}.$$

- The bandwidths (h_n) and (k_n) are non random sequences converging to 0 as $n \to \infty$.
- They control the smoothness of the kernel estimators. In what follows, the dependence on n for these two sequences is omitted.
- For the sake of simplicity we have chosen the same kernel $\mathcal{K}(\cdot)$.

To obtain the asymptotic property of the Regression Conditional Tail Moment estimator, an assumption on the right tail behavior of the conditional distribution of Y given $X = x$ is required. In the sequel, we assume that,

(F) The function $\mathrm{RVaR}(\cdot|x)$ is differentiable and

$$\lim_{\alpha \to 0} \frac{\mathrm{RVaR}'(t\alpha|x)}{\mathrm{RVaR}'(\alpha|x)} = t^{-(\gamma(x)+1)},$$

locally uniformly in $t \in (0, \infty)$.

$\Longrightarrow$ In other words :

$-\mathrm{RVaR}'(\cdot|x)$ is said to be regularly varying at 0 with index $-(\gamma(x) + 1)$

The condition **(F)** entails that the conditional distribution of Y given $X = x$ is in the maximum domain of attraction of the extreme value distribution with extreme value index $\gamma(x)$.

The unknown function $\gamma(x)$ is referred as the conditional extreme-value index.

It controls the behaviour of the tail of the survival function and by consequence the behaviour of the extreme values.

$\Longrightarrow$ if $\gamma(x) < 0$, $F(.|x)$ belongs to the domain of attraction of Weibull. It contains distributions with finite right tail, *i.e.* short-tailed.

$\Longrightarrow$ if $\gamma(x) = 0$, $F(.|x)$ belongs to the domain of attraction of Gumbel. It contains distributions with survival function exponentially decreasing, *i.e.* light-tailed.

$\Longrightarrow$ if $\gamma(x) > 0$, $F(.|x)$ belongs to the domain of attraction of Fréchet. It contains distributions with survival function polynomially decreasing, *i.e.* heavy-tailed.

The case $\gamma(x) > 0$ has already been investigated by El Methni *et al.* [2014].

Assumptions

First, a Lipschitz condition on the probability density function g of X is required. For all $(x, x') \in \mathbb{R}^p \times \mathbb{R}^p$, denoting by $d(x, x')$ the distance between x and x', we suppose that

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

The next assumption is devoted to the kernel function $\mathcal{K}(\cdot)$.

(K) $\mathcal{K}(\cdot)$ is a bounded density on $\mathbb{R}^p$, with support S included in the unit ball of $\mathbb{R}^p$.

Before stating our main result, some further notations are required.

For $\xi > 0$, the largest oscillation at point $(x, y) \in \mathbb{R}^p \times \mathbb{R}_*^+$ associated with the Regression Conditional Tail Moment of order $b \in [0, 1/\gamma_+(x))$ is given by

$$\omega\left(x, y, b, \xi, h\right) = \sup\left\{\left|\frac{\varphi_b(z|x)}{\varphi_b(z|x')} - 1\right| \text{ with } \left|\frac{z}{y} - 1\right| \leq \xi \text{ and } x' \in B(x, h)\right\},$$

where $\varphi_b(\cdot|x) := \overline{F}(\cdot|x) RCTM_b(\overline{F}(\cdot|x)|x)$ and $B(x, h)$ denotes the ball centred at x with radius h.

Asymptotic normality of $\widehat{\mathrm{RVaR}}_n(\alpha_n|x)$

Theorem 1

Suppose **(F)**, **(L)** and **(K)** hold. For $x \in \mathbb{R}^p$ such that $g(x) > 0$, let $\alpha_n \to 0$ such that

$$nk^p\alpha_n \to \infty \quad \text{as} \quad n \to \infty$$

If there exists $\xi > 0$ such that

$$nk^p\alpha_n \left(k \vee \omega(x, \mathrm{RVaR}(\alpha_n|x), 0, \xi, k)\right)^2 \to 0,$$

then

$$(nk^p\alpha_n^{-1})^{1/2} f(\mathrm{RVaR}(\alpha_n|x)|x) \left(\widehat{\mathrm{RVaR}}_n(\alpha_n|x) - \mathrm{RVaR}(\alpha_n|x)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|\mathcal{K}\|_2^2}{g(x)}\right).$$

$\Longrightarrow$ We thus find back the result established in Daouia *et al.* [2013] under weaker assumptions.

Theorem 2

Suppose **(F)**, **(L)** and **(K)** hold. For $x \in \mathbb{R}^p$ such that $g(x) > 0$:

- Let $0 \leq b_1 \leq \ldots \leq b_J < 1/(2\gamma_+(x))$,
- $\underline{\ell} = h \wedge k$ and $\overline{\ell} = h \vee k$.
- Let $\alpha_n \to 0$ such that $n\underline{\ell}^p \alpha_n \to \infty$ as $n \to \infty$.
- If there exists $\xi > 0$ such that

$$n\overline{\ell}^p \alpha_n \left(\overline{\ell} \vee \max_b \omega(x, \mathrm{RVaR}(\alpha_n|x), b, \xi, \overline{\ell}) \right)^2 \to 0,$$

then, if

$$h/k \to 0 \quad \text{or} \quad k/h \to 0$$

the random vector

$$(n\underline{\ell}^p \alpha_n)^{1/2} \left\{ \left(\frac{\widehat{\mathrm{RCTM}}_{b_j,n}(\alpha_n|x)}{\mathrm{RCTM}_{b_j}(\alpha_n|x)} - 1 \right) \right\}_{j \in \{1,\ldots,J\}}$$

is asymptotically Gaussian, centred, with a $J \times J$ covariance matrix.

Covariance matrix two cases

In what follows, $(\cdot)_+$ (resp. $(\cdot)_-$) denotes the positive (resp. negative) part function.

1. If $k/h \to 0$ then the covariance matrix is given by

$$\frac{\|\mathcal{K}\|_2^2 \Sigma^{(1)}(x)}{g(x)}$$

where for $(i,j) \in \{1,\ldots,J\}^2$,

$$\Sigma^{(1)}_{i,j}(x) = (1 - b_i\gamma_+(x))(1 - b_j\gamma_+(x)).$$

2. If $h/k \to 0$ then the covariance matrix is given by

$$\frac{\|\mathcal{K}\|_2^2 \Sigma^{(2)}(x)}{g(x)}$$

where for $(i,j) \in \{1,\ldots,J\}^2$,

$$\Sigma^{(2)}_{i,j}(x) = \frac{(1 - b_i\gamma_+(x))(1 - b_j\gamma_+(x))}{1 - (b_i + b_j)\gamma_+(x)} = \frac{\Sigma^{(1)}_{i,j}(x)}{1 - (b_i + b_j)\gamma_+(x)}$$

Covariance matrix two cases

Recall that

$$\Sigma^{(1)}_{i,j}(x) = (1 - b_i\gamma_+(x))(1 - b_j\gamma_+(x)) \quad \text{and} \quad \Sigma^{(2)}_{i,j}(x) = \frac{\Sigma^{(1)}_{i,j}(x)}{1 - (b_i + b_j)\gamma_+(x)}$$

- Note that if $\gamma(x) \leq 0$, asymptotic covariance matrices do not depend on $\{b_1, \ldots, b_J\}$ and thus the estimators share the same rate of convergence.

- Conversely, when $\gamma(x) > 0$, asymptotic variances are increasing functions of the RCTM order.

- Moreover, in this case, note that for all $i \in \{1, \ldots, J\}$

$$\Sigma^{(2)}_{i,i}(x) > \Sigma^{(1)}_{i,i}(x)$$

$\Longrightarrow$ Taking $k/h \to 0$ leads to more efficient estimators than $h/k \to 0$.

Proposition

Under **(F)**, the Regression Conditional Tail Moment or order b is asymptotically proportional to the Regression Value at Risk to the power b.

Proposition

Under **(F)**, *for all* $b \in [0, 1/\gamma_+(x))$,

$$\lim_{\alpha \to 0} \frac{\mathrm{RCTM}_b(\alpha|x)}{[\mathrm{RVaR}(\alpha|x)]^b} = \frac{1}{1 - b\gamma_+(x)},$$

and $\mathrm{RCTM}_b(\cdot|x)$ *is regularly varying with index* $-b\gamma_+(x)$.

In particular, the Proposition is an extension to a regression setting of the result established in Hua and Joe [2011] for the Conditional Tail Expectation ($b = 1$) in the framework of heavy-tailded distributions ($\gamma = \gamma(x) > 0$).

Right endpoint

Let us note $y^*(x) = \mathrm{RVaR}(0|x) = \bar{F}^{\leftarrow}(0|x) \in (0, \infty]$ the endpoint of Y given $X = x$

Two cases :

1. If the endpoint $y^*(x)$ is infinite :

$$y^*(x) = \infty \quad \text{then} \quad \gamma(x) \geq 0$$

$\Longrightarrow$ We can make risk measure estimation.

$\Longrightarrow$ An application in pluviometry has already been done in El Methni *et al.* [2014].

Right endpoint

2. If the endpoint $y^*(x))$ is finite, the risk measures do not have sense :

$$y^*(x) < \infty \quad \text{then} \quad \gamma(x) \leq 0$$

As a consequence of the Proposition

$$\mathrm{RCTM}_b(\alpha|x) = [\mathrm{RVaR}(\alpha|x)]^b(1+o(1)) \to [y^*(x)]^b \quad \text{as} \quad \alpha \to 0.$$

For all $b > 0$, a natural estimator of the right endpoint (or frontier) is thus given by

$$\hat{y}^*_{b,n}(x) := \left[\widehat{\mathrm{RCTM}}_{b,n}(\alpha_n|x)\right]^{1/b}$$

where α_n is a sequence converging to 0 as $n \to \infty$.

$\Longrightarrow$ We can use our Proposition to make frontier estimation.

$\Longrightarrow$ We propose an application in nuclear reactor reliability.

Simulation and Procedure

- The performance of the frontier estimator $\hat{y}^*_{b,n}(x)$ is illustrated on simulated data.
- $\hat{y}^*_{b,n}(x)$ depends on two hyper-parameters h and α :
  - The choice of the bandwidth h, which controls the degree of smoothing, is a recurrent problem in non-parametric statistics.
  - Besides, the choice of α is crucial, it is equivalent to the choice of the number of upper order statistics in the non-conditional extreme-value theory.
- We propose a data driven procedure to select h and α.
- The performance of the data-driven selection of the hyper-parameters is compared to an oracle one. Our procedure yields reasonable results.

- We have compared 10 estimators $\hat{y}^*_{1,n}, \ldots, \hat{y}^*_{10,n}$ deduced from $\hat{y}^*_{b,n}(x)$ with $\widehat{\mathrm{RVaR}}$ and three estimators $\hat{y}_n^{(*,gj)}$, $\hat{y}_n^{(*,mc)}$ and $\hat{y}_n^{(*,mv)}$ from Girard and Jacob [2008] and Girard *et al.* [2013]

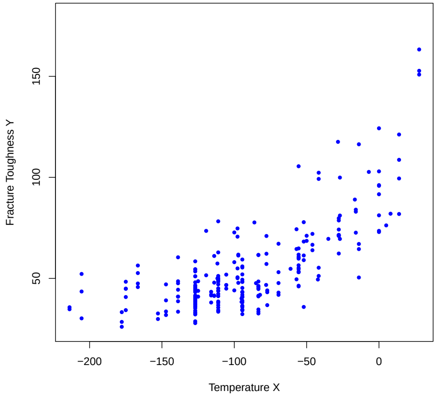
- It appears that $\hat{y}^*_{1,n} = \widehat{\mathrm{CTE}}$ does not yield very good results but $\hat{y}^*_{2,n}, \ldots, \hat{y}^*_{10,n}$ all perform better than $\widehat{\mathrm{RVaR}}$, $\hat{y}_n^{(*,gj)}$, $\hat{y}_n^{(*,mc)}$ and $\hat{y}_n^{(*,mv)}$ in all situations.

- Among them, $\hat{y}^*_{7,n}$ yields the best results but the behavior of $\hat{y}^*_{4,n}$, $\hat{y}^*_{5,n}$ and $\hat{y}^*_{6,n}$ are very close.

- As a conclusion it appears on this numerical study that $\hat{y}^*_{b,n}$ combined with the data-driven hyper-parameters selection are efficient frontier estimators for $b \geq 2$.

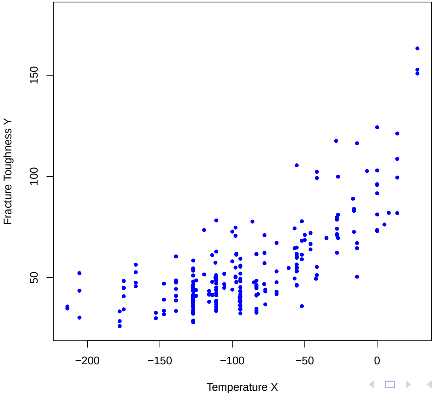
- Their performance seems to be stable with $b \geq 2$ but an automatic selection of b could be of interest.

- The dataset comes from the US Electric Power Research Institute and consists of $n = 254$ toughness results obtained from non-irradiated representative steels.
- The variable of interest Y is the fracture toughness and the unidimensional covariate X is the temperature measured in degrees Fahrenheit.
- As the temperature decreases, the steel fissures more easily.



Frontier estimation

- In a worst case scenario, it is important to know the upper limit of fracture toughness of each material as a function of the temperature, that is $y^*(x)$.
- An accurate knowledge of the change in fracture toughness of the reactor pressure vessel materials as a function of the temperature is of prime importance in a nuclear power plant's lifetime programme.

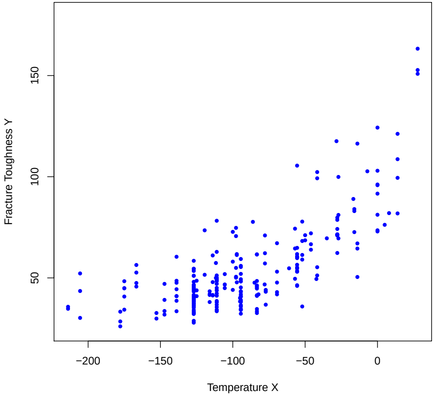


Frontier estimation

- The hyper-parameters associated with $\hat{y}^*_{7,n}$ are chosen in the sets

$$\mathcal{H} = \{17, 18, \ldots, 120\} \quad \text{and} \quad \mathcal{A} = \{0.01, 0.011, \ldots, 0.1\}$$

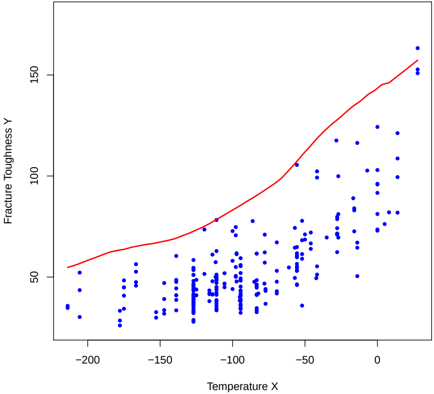
- The selection yields $(h_{data}, \alpha_{data}) = (98, 0.085)$



- The hyper-parameters associated with $\hat{y}^*_{7,n}$ are chosen in the sets

$$\mathcal{H} = \{17, 18, \dots, 120\} \quad \text{and} \quad \mathcal{A} = \{0.01, 0.011, \dots, 0.1\}$$

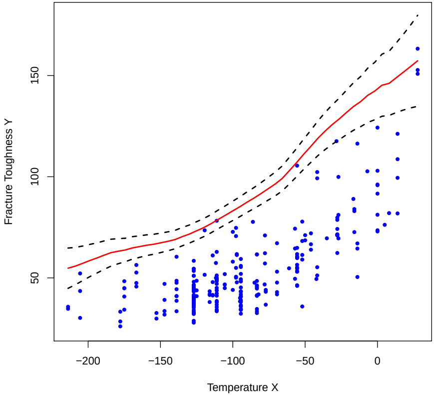
- The selection yields $(h_{data}, \alpha_{data}) = (98, 0.085)$



- The hyper-parameters associated with $\hat{y}^*_{7,n}$ are chosen in the sets

$$\mathcal{H} = \{17, 18, \dots, 120\} \quad \text{and} \quad \mathcal{A} = \{0.01, 0.011, \dots, 0.1\}$$

- The selection yields $(h_{data}, \alpha_{data}) = (98, 0.085)$

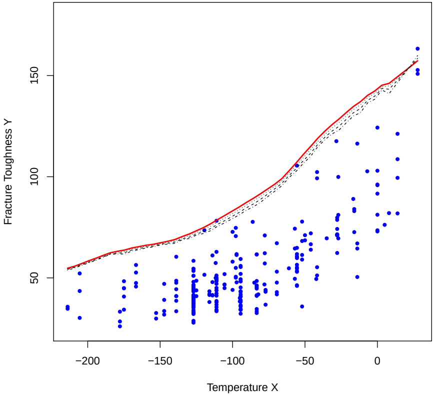


Frontier estimation

- The hyper-parameters associated with $\hat{y}^*_{7,n}$ are chosen in the sets

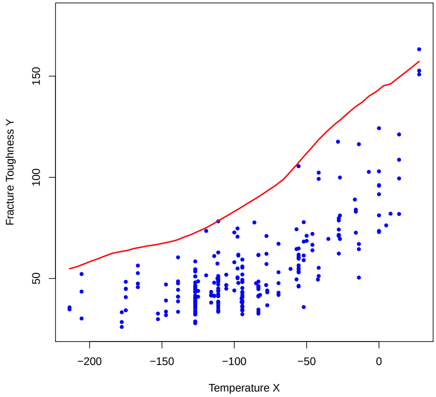
$$\mathcal{H} = \{17, 18, \dots, 120\} \quad \text{and} \quad \mathcal{A} = \{0.01, 0.011, \dots, 0.1\}$$

- The selection yields $(h_{data}, \alpha_{data}) = (98, 0.085)$

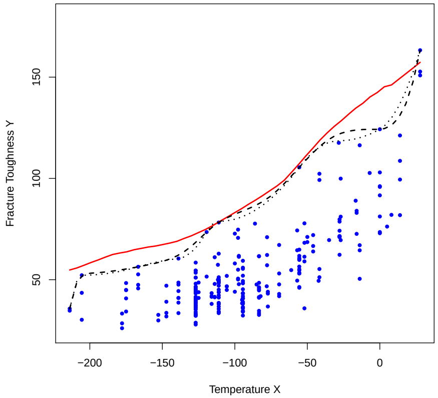


Frontier estimation

- We compare $\hat{y}^*_{7,n}$ to the spline-based estimators CS-B and QS-B recently introduced in Daouia *et al.* [2016] for monotone boundaries.
- The BIC criterion is used to determine the complexity of the spline approximation.



- We compare $\hat{y}^*_{7,n}$ to the spline-based estimators CS-B and QS-B recently introduced in Daouia *et al.* [2016] for monotone boundaries.
- The BIC criterion is used to determine the complexity of the spline approximation.



- CS-B and QS-B simply interpolate the boundary points whereas $\hat{y}^*_{7,n}$ estimates a heavier tail and thus a higher value for the limit of fracture toughness.
- Moreover, unlike us, they make different hypothesis on the form of the curve.

Conclusions

Commentaries

+ New tool for the prevention of risk and frontier estimation.
+ Theoretical properties similar to the univariate case (extreme or not) and with or without a covariate.
+ Our results are similar to those obtained by Daouia *et al.* [2013] and El Methni *et al.* [2014]. We have filled in the gap between these two works.
+ Capable to estimate risk measures based on conditional moments of the *r.v.* of losses given that the losses are greater than RVaR(α) for short, light and heavy-tailed distributions.
+ Tuning parameter selection procedure to choose (h, α).

Illustration on real data

$\Longrightarrow$ Application in pluviometry.

$\Longrightarrow$ Application in nuclear reactors reliability.

Long-term perspectives

- Curse of dimensionality.

Further readings

This presentation is based on the research article

El Methni, J., Gardes, L. and Girard, S. Kernel estimation of extreme regression risk measures, to appear in Electronic Journal of Statistics, 2018.

On my personal web page you can find a link to the Preprint version.

Further readings

This presentation is based on the research article

El Methni, J., Gardes, L. and Girard, S. Kernel estimation of extreme regression risk measures, to appear in Electronic Journal of Statistics, 2018.

On my personal web page you can find a link to the Preprint version.

Thank you for your attention

Principals references

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

Daouia, A., Noh, H. & Park, B.U. (2016). Data envelope fitting with constrainded polynomial spline, *Journal of the Royal Statistical Society, B*, **78(1)**, 3–36.

El Methni, J., Gardes, L. & Girard, S. (2014). Nonparametric estimation of extreme risk measures from conditional heavy-tailed distributions, *Scandinavian Journal of Statistics*, **41(4)**, 988–1012.

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

Girard, S., and Jacob, P. (2008). Frontier estimation via kernel regression on high power transformed data. *Journal of Multivariate Analysis*, **99**, 403–420.

Girard, S., Guillou, A. and Stupfler, G. (2013) Frontier estimation with kernel regression on high order moments, *Journal of Multivariate Analysis*, **116**, 172–189.

Hong, J. and Elshahat, A. (2010). Conditional tail variance and conditional tail skewness. *Journal of Financial and Economic Practice*, **10**, 147–156.

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