Nonparametric estimation of extreme risks from heavy-tailed distributions

Stéphane GIRARD

joint work with

Jonathan EL METHNI & Laurent GARDES

September 2013
1. Extreme risk measures

2. Estimators and asymptotic results

3. Extrapolation

4. Application
Some risk measures

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

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

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

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

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

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

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

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

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

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

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

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

All the previous risk measures of level $\alpha$ can be rewritten as

$$\text{CTE}(\alpha) = \text{CTM}_1(\alpha),$$
$$\text{CVaR}(\alpha) = \lambda \text{VaR}(\alpha) + (1 - \lambda)\text{CTM}_1(\alpha),$$
$$\text{CTV}(\alpha) = \text{CTM}_2(\alpha) - \text{CTM}_1^2(\alpha).$$

$\implies$ 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 to the case where a covariate $X \in \mathbb{R}^p$ is recorded simultaneously with $Y$.

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

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

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

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

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

where $a > 0$ is such that the moment of order $a$ of $Y$ exists.
Extreme regression risk measures

This yields the following risk measures:

\[
\begin{align*}
\text{RCTE}(\alpha_n|x) &= \text{RCTM}_1(\alpha_n|x), \\
\text{RCVaR}_\lambda(\alpha_n|x) &= \lambda \text{RVaR}(\alpha_n|x) + (1 - \lambda) \text{RCTM}_1(\alpha_n|x), \\
\text{RCTV}_n(\alpha_n|x) &= \text{RCTM}_2(\alpha_n|x) - \text{RCTM}_1^2(\alpha_n|x).
\end{align*}
\]

\[\implies\text{All the risk measures depend on the RVaR and the RCTM}_a.\]

The conditional moment of order \(a \geq 0\) of \(Y\) given \(X = x\) is defined by

\[\varphi_a(y|x) = \mathbb{E}(Y^a \mathbb{I}\{Y > y}\)|X = x),\]

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

\[
\begin{align*}
\text{RVaR}(\alpha_n|x) &= \varphi_0^{-}(\alpha_n|x), \\
\text{RCTM}_a(\alpha_n|x) &= \frac{1}{\alpha_n} \varphi_a(\varphi_0^{-}(\alpha_n|x)|x).
\end{align*}
\]

\[\text{Goal: estimate } \varphi_a(\cdot|x) \text{ and } \varphi_a^{-}(\cdot|x).\]
Inference

Estimator of $\varphi_a(\cdot|x)$:

We propose to use a classical kernel estimator given by

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

- $h_n$ 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 $\varphi_a^{-\cdot}(\cdot|x)$:

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

$$\hat{\varphi}_{a,n}^{-\cdot}(\alpha|x) = \inf \{t, \hat{\varphi}_{a,n}(t|x) < \alpha\}.$$
The conditional survival distribution function of $Y$ given $X = x$ is assumed to be heavy-tailed i.e. for all $\lambda > 0$,

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

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 $a \in [0, 1/\gamma(x))$, $\text{RCTM}_a(.)|x)$ exists, and for all $y > 0$,

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

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

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

(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 $a \in [0, 1/\gamma(x))$. 
Regularity assumptions

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

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

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

$$
\omega_n(y, \xi) = \sup \left\{ \left| \frac{\varphi_a(z|x)}{\varphi_a(z|x^{'})} - 1 \right|, \ z \in [(1 - \xi)y, (1 + \xi)y] \text{ and } d(x, x^{'}) \leq h \right\}.
$$
Main result

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

- \(0 \leq a_1 < a_2 < \cdots < a_J\),
- \(x \in \mathbb{R}^p\) such that \(g(x) > 0\) and \(0 < \gamma(x) < 1/(2a_J)\),
- \(\alpha_n \to 0\) and \(nh^p\alpha_n \to \infty\) as \(n \to \infty\),
- \(\xi > 0\) such that \(\sqrt{nh^p\alpha_n}(h \vee \omega_n(RV\text{VaR}(\alpha_n|x), \xi)) \to 0\),

Then,

\[
\sqrt{nh^p\alpha_n} \left\{ \left( \frac{\hat{\text{RCTM}}_{a_j,n}(\alpha_n|x)}{\text{RCTM}_{a_j}(\alpha_n|x)} - 1 \right) \right\}_{j \in \{1, \ldots, J\}}, \left( \frac{\hat{\text{RVaR}}_n(\alpha_n|x)}{\text{RVaR}(\alpha_n|x)} - 1 \right)
\]

is asymptotically Gaussian, centered, with covariance matrix

\[
\|K\|^2_2 \gamma^2(x) \Sigma(x)/g(x)
\]

where

\[
\Sigma(x) = \begin{pmatrix}
\frac{a_1 a_j (2 - (a_i + a_j) \gamma(x))}{1 - (a_i + a_j) \gamma(x)} & a_1 \\
\vdots & \ddots \\
\frac{a_j}{1 - (a_i + a_j) \gamma(x)} & \ddots & a_J \\
\frac{a_1 \cdots a_j}{a_1 \cdots a_J} & \cdots & 1
\end{pmatrix}.
\]
Conditions on the sequences $\alpha_n$ and $h_n$

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

\[ \sqrt{nh_n^p \alpha_n} (h \lor \omega_n(\text{RVaR}(\alpha_n|x), \xi)) \to 0 : \text{The bias induced by the smoothing is negligible compared to the standard-deviation.} \]
Suppose the assumptions of Theorem 1 hold. Then, if $0 < \gamma(x) < 1/2$,

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

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

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

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

where

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

- In Theorem 1, the condition $nh^p \alpha_n \to \infty$ 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, $\alpha_n$ must be an order of an extreme quantile within the sample.

**Definition**

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

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

**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

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

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

\[
\sqrt{nh_n^p \alpha_n} \left( \frac{\text{RCTM}_{a,n}(\beta_n|x)}{\text{RCTM}_a(\beta_n|x)} - 1 \right) \xrightarrow{d} \mathcal{N}(0, (av(x))^2).
\]

The condition \( \beta_n/\alpha_n \to 0 \) allows us to extrapolate and choose a level \( \beta_n \) arbitrarily small.
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 \(Z_1, \ldots, Z_n\).

With a covariate:

A kernel version of the Hill estimator is given by

\[
\hat{\gamma}_{n, \alpha_n}(x) = \sum_{j=1}^{J} \left( \log \widehat{RVaR}_n(\tau_j \alpha_n | x) - \log \widehat{RVaR}_n(\tau_1 \alpha_n | x) \right) \left/ \sum_{j=1}^{J} \log(\tau_1 / \tau_j) \right.,
\]

where \(J \geq 1\) and \((\tau_j)_{j \geq 1}\) is a decreasing sequence of weights.
The asymptotic normality of $\hat{\gamma}_{n,\alpha_n}(x)$ and
\[
\hat{RVaR}_n^W(\beta_n|x) = \hat{RVaR}_n(\alpha_n|x) \left( \frac{\alpha_n}{\beta_n} \right)^{\hat{\gamma}_n(x)}.
\]

has been established by Daouia et al. [2011].

As a consequence, replacing $\hat{RVaR}_n$ by $\hat{RVaR}_n^W$ and $\hat{RCTM}_{a,n}$ by $\hat{RCTM}_{a,n}^W$ provides (asymptotically Gaussian) estimators for all the risk measures considered in this talk, and for arbitrarily small levels.

In particular, since $\hat{RCTE}(\alpha_n|x) = \hat{RCTM}_1(\alpha_n|x)$, we obtain
\[
\hat{RCTE}_n^W(\beta_n|x) = \hat{RCTE}_n(\alpha_n|x) \left( \frac{\alpha_n}{\beta_n} \right)^{\hat{\gamma}_n(x)}.
\]
Daily rainfalls in the Cévennes-Vivarais region

The Cévennes-Vivarais region

523 Stations / 1958–2000 / in mm

Estimation of risk measures associated to return periods of 100 years
A cross validation procedure to choose $h_n$ and $\alpha_n$: Step 1

- Double loop on $\mathcal{H} = \{h_i; i = 1, \ldots, M\}$ and on $\mathcal{A} = \{\alpha_j; j = 1, \ldots, R\}$.
- Loop on all raingauge stations $\{x_t; t = 1, \ldots, N\}$.
A cross validation procedure to choose $h_n$ and $\alpha_n$ : Step 1

1. Double loop on $\mathcal{H} = \{ h_i; i = 1, \ldots, M \}$ and on $\mathcal{A} = \{ \alpha_j; j = 1, \ldots, R \}$.
2. Loop on all raingauge stations $\{ x_t; t = 1, \ldots, N \}$.

Consider one raingauge station $x_r$.
A cross validation procedure to choose $h_n$ and $\alpha_n$: Step 1

- Double loop on $\mathcal{H} = \{h_i; i = 1, \ldots, M\}$ and on $\mathcal{A} = \{\alpha_j; j = 1, \ldots, R\}$.
- Loop on all raingauge stations $\{x_t; t = 1, \ldots, N\}$.

Remove all other raingauge stations.
A cross validation procedure to choose $h_n$ and $\alpha_n$ : **Step 1**

- Double loop on $H = \{h_i; i = 1, \ldots, M\}$ and on $A = \{\alpha_j; j = 1, \ldots, R\}$.
- Loop on all raingauge stations $\{x_t; t = 1, \ldots, N\}$.

Remove all other raingauge stations

- Estimate $\gamma > 0$ using the classical Hill estimator.
- It only depends on $\alpha_j$. 
A cross validation procedure to choose $h_n$ and $\alpha_n$: Step 1

- Double loop on $\mathcal{H} = \{h_i; \ i = 1, \ldots, M\}$ and on $\mathcal{A} = \{\alpha_j; \ j = 1, \ldots, R\}$.
- Loop on all raingauge stations $\{x_t; \ t = 1, \ldots, N\}$.

Remove all other raingauge stations

- Estimate $\gamma > 0$ using the classical Hill estimator.
- It only depends on $\alpha_j$.

$\Rightarrow$ We obtain $\hat{\gamma}_{n,t,\alpha_j}$
A cross validation procedure to choose $h_n$ and $\alpha_n$ : Step 2

Remove the station $x_t$
A cross validation procedure to choose $h_n$ and $\alpha_n$ : Step 2

Work in $B(x_t, h_i) \setminus \{x_t\}$
A cross validation procedure to choose $h_n$ and $\alpha_n$: Step 2

Work in $B(x_t, h_i) \setminus \{x_t\}$

- Estimate $\gamma(x) > 0$ using the kernel version of the Hill estimator.
- It depends on $\alpha_j$ and on $h_i$. 
A cross validation procedure to choose $h_n$ and $\alpha_n$ : Step 2

Work in $B(x_t, h_i) \setminus \{x_t\}$

- Estimate $\gamma(x) > 0$ using the kernel version of the Hill estimator.
- It depends on $\alpha_j$ and on $h_i$.

$\Rightarrow$ We obtain $\hat{\gamma}_{n,h_i,\alpha_j}(x_t)$
A cross validation procedure to choose $h_n$ and $\alpha_n$ : Step 2

- Estimate $\gamma(x) > 0$ using the kernel version of the Hill estimator.
- It depends on $\alpha_j$ and on $h_i$.

$\implies$ We obtain $\hat{\gamma}_{n,h_i,\alpha_j}(x_t)$

$$\left( h_{\text{emp}}, \alpha_{\text{emp}} \right) = \arg\min_{(h_i, \alpha_j) \in \mathcal{H} \times \mathcal{A}} \text{median}\{ (\hat{\gamma}_{n,t,\alpha_j} - \hat{\gamma}_{n,h_i,\alpha_j}(x_t))^2, \; t \in \{1, \ldots, N\} \}. $$
Computation of $\widehat{RVaR}_n^W$ and $\widehat{RCTE}_n^W$

- Two dimensional covariate $X = (\text{latitude}, \text{longitude})$.
- Bi-quadratic kernel: $K(x) \propto (1 - \|x\|^2)^2 \mathbb{I}_{\{\|x\| \leq 1\}}$.
- Harmonic sequence of weights: $(\tau_j)_{j \in \{1, \ldots, 9\}} = 1/j$.
- Results of the procedure $(h_{\text{emp}}, \alpha_{\text{emp}}) = (24, 1/(3 \times 365.25))$. 

523 Stations

Regular grid: $200 \times 200$
Estimated risk measures for a return period of 3 years

\[ \hat{RVaR}_n(1/(3 \times 365.25)|x) \]

\[ \hat{RCTE}_n(1/(3 \times 365.25)|x) \]
Estimated conditional tail index
RVaR\textsubscript{\text{W}} \left( \frac{1}{(100 \times 365.25)} \right) : 100-year return level
\( \text{RCTE}_n^W \left( \frac{1}{(100 \times 365.25)} | x \right) \) above the 100-year return level


