HAL
open science

Estimation de mesures de risque extrêmes

Jonathan El Methni, Laurent Gardes, Stéphane Girard

▶ **To cite this version:**

Jonathan El Methni, Laurent Gardes, Stéphane Girard. Estimation de mesures de risque extrêmes. 45èmes Journées de Statistique, May 2013, Toulouse, France. pp.électronique. hal-00829625

HAL Id: hal-00829625

https://hal.science/hal-00829625v1

Submitted on 3 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation de Mesures de Risque Extrêmes

Jonathan El Methni [1], Laurent Gardes [2] & Stéphane Girard [1]

[1] *Equipe Mistis, INRIA Rhône-Alpes & Laboratoire Jean Kuntzmann*
655, avenue de l'Europe, Montbonnot, 38334 Saint-Ismier cedex, France.

[2] *Université de Strasbourg & CNRS, IRMA, UMR 7501,*
7, rue René Descartes, 67084 Strasbourg cedex, France.

jonathan.el-methni@inria.fr, gardes@unistra.fr & stephane.girard@inria.fr

Résumé. Des mesures de risques classiques sont la Value-at-Risk, la Conditional Tail Expectation, la Conditional Value-at-Risk et la Conditional Tail Variance. En termes statistique, la Value-at-Risk est le quantile de niveau de confiance $\alpha \in]0,1[$ de la distribution des pertes. On s'intéresse aux propriétés de ces mesures de risque dans le cas de pertes extrêmes (où α n'est plus fixé mais tend vers 0) qu'on supposera modélisées par des lois à queues lourdes. On considérera aussi ces mesures de risque avec la présence d'une covariable. On ajoute ainsi deux difficultés dans l'estimation de mesures de risque. Par conséquent, le but principal de cette communication est de proposer des estimateurs de toutes les mesures de risque énoncées ci-dessus pour des pertes extrêmes dans le cas de lois à queues lourdes en présence d'une covariable. On établira les propriétés asymptotiques de nos estimateurs et on illustrera leurs comportements sur des données simulées et sur un jeu de données pluviométriques.

Mots-clés. Mesure de risques, quantiles conditionnels, lois à queues lourdes, normalité asymptotique, estimateur à noyau, statistique des valeurs extrêmes.

Abstract. Classical risk measures are the Value-at-Risk, the Conditional Tail Expectation, the Conditional Value-at-Risk and the Conditional Tail Variance. In statistical terms, the Value-at-Risk is the upper α-quantile of the loss distribution where $\alpha \in (0,1)$ is the confidence level. Here, we focus on the properties of these risk measures for cases of extreme loss (where $\alpha \to 0$ is no longer fixed). To assign probabilities to extreme loss we assume that we are in the case of heavy-tailed losses. We also consider these risk measures in the presence of a covariate. Thus we add two difficulties in estimating risk measures. Consequently, the main goal of this communication is to propose estimators of all risk measures cited before in the case of heavy-tailed distributions for extreme loss and to include a covariate in the estimation. We establish the asymptotic properties of our estimators and we illustrate their behavior on simulated data and on a real data set of pluviometrical measurements.

Keywords. Risk measures, conditional quantiles, heavy-tailed distributions, asymptotic normality, kernel estimator, extreme-value statistics.

1 Introduction

La maîtrise des risques est un sujet de préoccupation aussi bien en hydrologie, météorologie qu'en finance et en actuariat. On appelle mesure de risque une fonction associant à une variable aléatoire de perte Y une valeur positive ou nulle quantifiant le risque. La plus utilisée d'entre elles est la Value-at-Risk au niveau de confiance $\alpha \in]0,1[$ notée $VaR(\alpha)$ [4]. Elle représente le quantile d'ordre α de la fonction de survie $\overline{F}$ de la variable aléatoire Y et elle est définie par

$$VaR(\alpha) := q(\alpha) = \overline{F}^{\leftarrow}(\alpha) = \inf\{z : \overline{F}(z) \leq \alpha\}.$$

Cette mesure de risque ne fournit qu'une information ponctuelle et donc sous-estime l'impact du sinistre. Dans le but de trouver une alternative à la Value-at-Risk, d'autres mesures de risque ont été proposées afin de prendre en compte les incertitudes sur les événements extrêmes. La Conditional Tail Expectation [1] au niveau de confiance $\alpha \in]0,1[$ notée $CTE(\alpha)$ est définie par

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

Elle donne des informations sur la distribution de Y au delà de la $VaR(\alpha)$ et donc contrairement à la $VaR(\alpha)$, sur l'épaisseur de la queue de distribution. Pour un niveau de confiance $\alpha \in]0,1[$, elle représente la moyenne des $(1-\alpha)100\%$ sinistres les plus élevés. La Conditional Value-at-Risk notée CVaR [6] au niveau de confiance $\alpha \in]0,1[$ est une moyenne pondérée de la $VaR(\alpha)$ et de la $CTE(\alpha)$ définie par

$$CVaR_\lambda(\alpha) := \lambda VaR(\alpha) + (1-\lambda)CTE(\alpha),$$

où $0 \leq \lambda \leq 1$. La Conditional Tail Variance notée CTV [8] et définie par

$$CTV(\alpha) := \mathbb{E}\left((Y - CTE(\alpha))^2 \,|Y > q(\alpha)\right),$$

mesure la variabilité de Y sachant $Y > q(\alpha)$. Les mesures de risque introduites précédemment sont toutes basées sur le moment d'ordre $a \geq 0$ de Y sachant $Y > y > 0$ défini par

$$\varphi_a(y) = \mathbb{E}\left(Y^a \mathbb{I}\{Y > y\}\right).$$

2 Mesures de risque extrêmes avec covariable

L'apport nouveau de ce travail consiste en l'ajout de deux difficultés supplémentaires dans le cadre de l'estimation des mesures de risque citées précédemment.

1. On ajoute la présence d'une covariable $X \in \mathbb{R}^p$.

2. On s'intéresse à l'estimation de mesures de risques pour des pertes extrêmes et pour des lois à queues lourdes. Pour cela on remplace α par une suite $\alpha_n \to 0$ quand la taille de l'échantillon $n \to \infty$.

La loi conditionnelle de Y sachant $X = x$ est $F(y|x) = \mathbb{P}(Y < y|X = x)$. On définit le moment conditionel d'ordre $a \geq 0$ de Y sachant $Y > y > 0$ par

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

En particulier si $a = 0$, on a $\varphi_0(y|x) = \overline{F}(y|x)$. Pour tout $x \in \mathbb{R}^p$ la fonction

$$VaR(\alpha_n|x) := q(\alpha_n|x) = \varphi_0^{\leftarrow}(\alpha_n|x),$$

est appelée quantile conditionnel. On dit d'un quantile conditionnel qu'il est extrême lorsque $\alpha_n \to 0$ quand $n \to \infty$. Ainsi les mesures de risque en présence d'une covariable pour des quantiles conditionnels extrêmes s'écrivent

$$\begin{aligned} CTE(\alpha_n|x) &:= \frac{1}{\alpha_n}\varphi_1(\varphi_0^{\leftarrow}(\alpha_n|x)|x), \\ CVaR_\lambda(\alpha_n|x) &:= \lambda VaR(\alpha_n|x) + (1-\lambda)CTE(\alpha_n|x), \\ CTV(\alpha_n|x) &:= \frac{1}{\alpha_n}\varphi_2(\varphi_0^{\leftarrow}(\alpha_n|x)|x) - CTE^2(\alpha_n|x). \end{aligned}$$

3 Définition des estimateurs et leurs lois asymptotiques

Soient $\{(X_i, Y_i), i = 1, \cdots, n\}$ des copies indépendantes du couple aléatoire $(X, Y) \in \mathbb{R}^p \times \mathbb{R}$ où Y est une variable d'intérêt associée à une covariable X. Afin d'estimer $\varphi_a(.|x)$, on se propose d'utiliser un estimateur à noyau classique introduit par Parzen-Rosenblatt ([5,7]). Il est défini pour $(x, y) \in \mathbb{R}^p \times \mathbb{R}$ par

$$\hat{\varphi}_{a,n}(y|x) = \sum_{i=1}^{n} K_h(x - X_i) Y_i^a \mathbb{I}\{Y_i > y\} \Big/ \sum_{i=1}^{n} K_h(x - X_i),$$

où $\mathbb{I}\{.\}$ est la fonction indicatrice et $h = h_n$ est une suite non aléatoire telle que $h \to 0$ quand $n \to \infty$ appelée paramètre de lissage. On a aussi introduit $K_h(t) = K(t/h)/h^p$ où la fonction $K(.)$ appelée noyau est une densité de probabilité sur $\mathbb{R}^p$. Puisque $\hat{\varphi}_{a,n}(.|x)$ est une fonction décroissante, on peut définir un estimateur de $\varphi_a^{\leftarrow}(\alpha|x)$ pour $\alpha \in]0, 1[$ par

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

On a donc un estimateur pour toutes les mesures de risque mentionnées avant

$$\begin{aligned}
\widehat{VaR}_n(\alpha_n|x) &:= \hat{\varphi}^{\leftarrow}_{0,n}(\alpha_n|x),\\
\widehat{CTE}_n(\alpha_n|x) &:= \frac{1}{\alpha_n}\hat{\varphi}_{1,n}(\hat{\varphi}^{\leftarrow}_{0,n}(\alpha_n|x)|x),\\
\widehat{CVaR}_{\lambda,n}(\alpha_n|x) &:= \lambda\widehat{VaR}_n(\alpha_n|x) + (1-\lambda)\widehat{CTE}_n(\alpha_n|x),\\
\widehat{CTV}_n(\alpha_n|x) &:= \frac{1}{\alpha_n}\hat{\varphi}_{2,n}(\hat{\varphi}^{\leftarrow}_{0,n}(\alpha_n|x)|x) - \widehat{CTE}_n^2(\alpha_n|x).
\end{aligned}$$

Afin d'établir la normalité asymptotique de nos estimateurs on suppose que la fonction de survie conditionnelle de Y sachant $X = x$ est à variations régulières d'indice $-1/\gamma(x)$ à l'infini. Cela signifie que pour tout $y > 0$,

$$\overline{F}(y|x) = 1 - F(y|x) = y^{-1/\gamma(x)}\ell(y|x),$$

avec $\gamma(.)$ une fonction inconnue et positive de la covariable x que l'on appelle "indice de queue conditionnel "ou "indice des valeurs extrêmes conditionnel"[3] et pour tout x fixé, $\ell(.|x)$ une fonction à variations lentes à l'infini c'est-à-dire pour tout $\lambda > 0$,

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

Autrement dit cela revient à supposer que la loi conditionnelle de Y sachant $X = x$ est à queue lourde. Rappelons que l'estimation de la VaR des pertes extrêmes en présence d'une covariable pour des lois à queues lourdes a déjà été étudiée [2].

On donne des conditions suffisantes pour avoir les convergences en loi de nos estimateurs. En particulier on suppose que, pour $x \in \mathbb{R}^p$ la fonction densité $g(x)$ de X est positive et que certaines hypothèses sur K et $\ell(.|x)$ sont vérifiées. On introduit une suite $(\alpha_n)_{n\geq 1}$ telle que $\alpha_n \to 0$ et $nh^p\alpha_n \to \infty$ quand $n \to \infty$, on a alors

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

et

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

pour $0 < \gamma(x) < 1/2$. On a de plus

$$\sqrt{nh^p\alpha_n}\left(\frac{\widehat{CTV}_n(\alpha_n|x)}{CTV(\alpha_n|x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{8(1-\gamma(x))(1-2\gamma(x))(1+2\gamma(x)+3\gamma^2(x))}{(1-3\gamma(x))(1-4\gamma(x))}\frac{\|K\|_2^2}{g(x)}\right),$$

pour $0 < \gamma(x) < 1/4$.

On remarque que la condition $nh^p\alpha_n \to \infty$ est restrictive sur le choix de α_n, on ne peut donc pas extrapoler au delà de l'échantillon. Afin de s'affranchir de cette condition et de pouvoir estimer des mesures de risque pour un niveau de confiance aussi petit que l'on souhaite on introduit alors une méthode d'extrapolation inspirée de l'estimateur de Weissman [9].

4 Conclusion et applications

Après avoir illustré l'efficacité de nos estimateurs sur des simulations, on appliquera nos travaux à un jeu de données pluviométriques fourni par le Laboratoire d'étude des Transferts en Hydrologie et Environnement de Grenoble. On a mesuré les hauteurs de pluies journalières (en mm) entre les années 1958 et 2000 sur 524 stations situées dans la région des Cévennes-Vivarais. Dans notre contexte, les variables d'intérêt sont les précipitations journalières et les covariables sont les coordonnées géographiques des stations. En conclusion de cette étude, on pourra estimer des mesures de risque extrêmes en des sites où l'on ne dispose pas de mesure.

Bibliographie

[1] P. Artzner, F. Delbaen, J.M. Eber and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[2] A. Daouia, L. Gardes, S. Girard and A. Lekina. Kernel estimators of extreme level curves, *TEST*, 20:311–333, 2011.

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

[4] P. Jorion. *Value at risk: the new benchmark for managing financial risk*, McGraw-Hill New York, 2007.

[5] E. Parzen. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 1065:1076–33, 1962.

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

[7] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 832–837, 1956.

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

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