Estimation de l'espérance conditionnelle des pertes extrêmes dans le cas de lois à queues lourdes en présence d'une covariable

Jonathan El Methni, Laurent Gardes, Stephane Girard

Estimation de l'espérance conditionnelle des pertes extrêmes dans le cas de lois à queues lourdes en présence d'une covariable.

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. L'espérance conditionnelle des pertes extrêmes (Conditional Tail Expectation, CTE) est une importante mesure de risque fréquemment utilisée en actuariat et en finance. Elle représente la perte attendue au delà d'un certain quantile. Dans la littérature, il existe plusieurs méthodes d'estimation de la CTE basées sur l'estimation du quantile par exemple dans le cas de lois à queues lourdes. L'objectif de cette communication est double. On propose un estimateur de la CTE dans le cas de lois à queues lourdes en présence d'une covariable et cela pour des quantiles conditionnels extrêmes.

Mots-clés. Conditional tail expectation, quantiles conditionnels, lois à queue lourde, normalité asymptotique, estimateur à noyau, statistique des valeurs extrêmes.

Abstract. The Conditional Tail Expectation (CTE) is an important mesure of right-tail risks, which are frequently encountered in the insurance and financial investment. It describes in risk analysis the expected amount of risk above a given quantile. The asymptotic normality of estimators of the CTE based on the quantile estimation has already been established in the literature for instance in the case of heavy-tailed distributions. The main goal of this communication, is to propose an estimator of the CTE in the case of heavy-tailed distributions with a covariable for extreme conditional quantiles.

Keywords. Conditional tail expectation, conditional quantiles, heavy-tailed distributions, asymptotic normality, kernel estimator, extreme-value statistics.

1 Introduction

Une des plus importantes mesures de risque en actuariat et en finance est l'espérance conditionnelle des pertes (Conditional Tail Expectation, CTE) [1], elle représente la perte attendue au delà d'un quantile d'ordre α. Pour un niveau de confiance $\alpha \in]0,1[$, elle représente la moyenne des $(1-\alpha)100\%$ sinistres les plus élevés. Considérons une variable aléatoire Y de perte de fonction de survie $\bar{F}$, l'espérance conditionnelle des pertes au niveau de probabilité $\alpha \in]0,1[$ notée $CTE(\alpha)$ est définie par

$$CTE(\alpha) = \mathbb{E}(Y|Y > q(\alpha))$$

où $q(\alpha)$ est le quantile d'ordre α défini par

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

Cette mesure de risque donne des informations sur la distribution de Y au delà du quantile d'ordre α et donc sur l'épaisseur de la queue de distribution. Dans la littérature, il existe plusieurs méthodes d'estimation de la CTE basées sur l'estimation du quantile $q(\alpha)$ citons par exemple [1]. Necir et al. [5] ont proposé un estimateur de la CTE dans le cas de lois à queue lourde où le moment d'ordre 2 n'existe pas.

Pour tout $x \in \mathbb{R}^p$ la fonction $q(\alpha|x) = \bar{F}^{\leftarrow}(\alpha|x) = \inf\{z : \bar{F}(z|x) \leq \alpha\}$ est appelée quantile conditionnel. On dit d'un quantile conditionnel qu'il est extrême lorsque l'on a $q(\alpha_n|x)$ avec une suite réelle $\alpha_n \to 0$ quand la taille de l'échantillon $n \to \infty$.

L'apport nouveau de ce travail consiste en l'ajout de deux difficultés supplémentaires dans le cadre de l'estimation de l'espérance conditionnelle des pertes pour des lois à queues lourdes. En premier lieu on ajoute la présence d'une covariable $X \in \mathbb{R}^p$ et en second lieu on estime l'espérance conditionnelle des pertes avec $\alpha_n \to 0$ quand $n \to \infty$ c'est-à-dire pour des quantiles conditionnels extrêmes. Cela s'écrit

$$CTE(\alpha_n|X = x) = \mathbb{E}(Y|Y > q(\alpha_n|x), X = x), \tag{1}$$

où 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 $y > 0$,

$$\bar{F}(y|x) = 1 - F(y|x) = y^{-1/\gamma(x)}\ell(y|x), \tag{2}$$

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"[4] 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. On peut dans ce cas écrire à partir de (1)

$$CTE(\alpha_n | X = x) = q(\alpha_n|x) + \frac{1}{\alpha_n} \int_{q(\alpha_n|x)}^{\infty} \bar{F}(y|x) dy, \tag{3}$$

avec $0 < \gamma(x) < 1$.

2 Définition de l'estimateur et sa loi asymptotique

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. On se propose d'estimer (3) avec $\alpha_n \to 0$ lorsque $n \to \infty$ par

$$\widehat{CTE}(\alpha_n | X = x) = \hat{q}_n(\alpha_n|x) + \frac{1}{\alpha_n} \int_{\hat{q}_n(\alpha_n|x)}^{\infty} \hat{\bar{F}}_n(y|x) dy. \tag{4}$$

où on estime $q(\alpha_n|x)$ par l'inverse généralisé de l'estimateur de la fonction de survie conditionnelle défini par

$$\hat{q}_n(\alpha_n|x) = \hat{\bar{F}}_n^{\leftarrow}(\alpha_n|x) = \inf\{t, \hat{\bar{F}}_n(t|x) \leq \alpha_n\}.$$

Pour estimer la fonction de survie conditionnelle de Y sachant $X = x$, on se propose d'utiliser un estimateur à noyau introduit par Collomb [2]. Il est défini pour $(x, y) \in \mathbb{R}^p \times \mathbb{R}$ par

$$\hat{\bar{F}}_n(y|x) = \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right) \mathbb{I}\{Y_i > y\} \Bigg/ \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right)$$

où la fonction $K(.)$ appelée noyau est une densité de probabilité sur $\mathbb{R}^p$ et h_n est une suite non aléatoire telle que $h_n = h \to 0$ quand $n \to \infty$ appelée paramètre de lissage.
Afin d'établir la loi asymptotique de $\widehat{CTE}(\alpha_n | X = x)$ on introduit

$$\widetilde{CTE}(\alpha_n | X = x) = q(\alpha_n|x) + \frac{1}{\alpha_n} \int_{q(\alpha_n|x)}^{\infty} \hat{\bar{F}}_n(y|x) dy. \tag{5}$$

En utilisant (4) et (5) on obtient

$$\widetilde{CTE}(\alpha_n | X = x) - CTE(\alpha_n | X = x) = \frac{1}{\alpha_n} \int_{q(\alpha_n|x)}^{\infty} \left(\hat{\bar{F}}_n(y|x) - \bar{F}(y|x)\right) dy.$$

On établit la normalité asymptotique de $\widetilde{CTE}(\alpha_n | X = x)$ à l'aide du théorème suivant

Théorème 1 *Soit $\bar{F}$ vérifiant (2), sous certaines hypothèses sur K, $\ell(.|x)$ et $\gamma(x)$, on introduit*

- $\tau_1 > \tau_2 > \cdots > \tau_J > 0$ *où J est un entier naturel non nul,*
- $\alpha_{n,j} = \tau_j \alpha_n (1 + o(1))$ *pour* $j = 1, \ldots, J$,
- $\alpha_n \to 0$ *tel que* $nh^p \alpha_n \to \infty$ *et* $nh^p \alpha_n q(\alpha_n|x)^2 \left(h \vee \delta_{n,x}(h)\right)^2 \to 0$ *quand* $n \to \infty$, *avec*
$$\delta_{n,x}(h) = \sup_{d(x,x') \leq h} \left| \frac{\int_{q(\alpha_n|x)}^{\infty} \bar{F}(z|x) dz}{\int_{q(\alpha_n|x)}^{\infty} \bar{F}(z|x') dz} - 1 \right|.$$

Ainsi, pour tout $x \in \mathbb{R}^p$ tel que la fonction densité $g(x)$ de X soit positive et $0 < \gamma(x) < 1/2$, le vecteur aléatoire

$$\left\{ \sqrt{\frac{nh^p}{\alpha_n q(\alpha_n|x)^2}} \left(\widetilde{CTE}(\alpha_{n,j}|X = x) - CTE(\alpha_{n,j}|X = x) \right) \right\}_{j=1,\ldots,J}$$

est asymptotiquement Gaussien, centré, avec pour matrice de variance-covariance $\frac{\|K\|_2^2}{g(x)} C(x)$ où $C_{j,j'}(x) = ((1 - 2\gamma)(1 - \gamma))^2 \tau_{j \wedge j'}^{1-2\gamma(x)} / \gamma^4$ pour $(j, j') \in \{1, \ldots, J\}^2$.

Dans un travail futur, il reste à remplacer $q(\alpha_n|x)$ par $\hat{q}_n(\alpha_n|x)$ et utiliser le théorème 2 de [3] pour en déduire la loi asymptotique de $\widehat{CTE}(\alpha_n|X = x)$.

Bibliographie

[1] 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.

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

[3] Daouia, A., Gardes, L., Girard, S. et Lekina, A. (2010), Kernel estimators of extreme level curves, *Test*, 20, 311–333.

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

[5] 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.