Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory

Elena Di Bernardino & Thomas Laloë & Clémentine Prieur & Véronique Maume-Deschamps

Université de Lyon, Université Lyon 1, ISFA, Laboratoire SAF, 50 avenue Tony Garnier, 69366 Lyon, France, elena.di-bernardino@univ-lyon1.fr

Université de Nice Sophia-Antipolis, Laboratoire J-A Dieudonné, Parc Valrose, 06108 Nice cedex 02, laloe@unice.fr

Université de Lyon, Université Lyon 1, ISFA, Laboratoire SAF, 50 avenue Tony Garnier, 69366 Lyon, France, veronique.maume@univ-lyon1.fr

Université Joseph Fourier, Tour IRMA, MOISE-LJK B.P. 53 38041 Grenoble, France, clementine.prieur@imag.fr

Keywords : Level sets, Plug-in estimation, Hausdorff distance, Conditional Tail Expectation.

1 Introduction and Generalities

The Conditional Tail Expectation (CTE) is a commonly used tool in the univariate risk theory. In this paper we introduce a new bivariate generalization of CTE and provide a consistent procedure to estimate it. To this end we need essentially a generalization of the notion of quantile in dimension higher than one.

As a possible generalization of the univariate quantile function, Belzunce *et al.* [1] defined a multivariate quantile as a set of points which accumulate the same probability for a fixed orthant. They called it level curves or quantile curves. In this paper we will follow this approach in order to propose a new risk measure which generalizes the CTE in the bivariate case. To this end we need to estimate the level sets of a bivariate distribution function F. This leads us into the general field of level sets estimation.

The problem of estimating level sets of an unknown function (for instance of a density function and more recently a regression function) has received attention in the literature. In particular, the problem of estimating general level sets under compactness assumptions has been discussed by Cuevas *et al.* [2].

As we consider the level sets of a bivariate distribution function, the commonly assumed property of compactness for these sets is no more reasonable. This requires special attention in the statement of our problem.

Our general approach will be the following : first, we provide a consistent estimator of the level set

$$L(c) = \{F(x) \geq c\}, \quad \text{for } c \in (0,1).$$

To this end we consider a *plug-in* approach (e.g. see Cuevas *et al.*[2]), that is, $L(c)$ is estimated by

$$L_n(c) = \{F_n(x) \geq c\}, \quad \text{for } c \in (0,1),$$

where F_n is a consistent estimator of F. We obtain two consitency results : one for the Hausdorff distance (Theorem 2.1 in [3]) and one for the volume of the symetrical difference (Theorem 3.1 in [3]) .

We then introduce our new bivariate version of the Conditional Tail Expectation by conditioning the two-dimensional random vector to be in the level set $L(c)$:

$$\mathrm{CTE}_\alpha(\mathrm{X},\mathrm{Y}) = \mathbb{E}[(\mathrm{X},\mathrm{Y})|(\mathrm{X},\mathrm{Y}) \in \mathrm{L}(\alpha)] = \begin{pmatrix} \mathbb{E}[\, X \,|\, (X,Y) \in L(\alpha)\,] \\ \mathbb{E}[\, Y \,|\, (X,Y) \in L(\alpha)\,] \end{pmatrix}.$$

This new risk measure is based on the bivariate Value-at-Risk proposed in Embrechts and Puccetti[4] and deals with the simultaneous joint damages considering the dependence structure of data in a bivariate specific risk's area ($L(c)$). We propose an estimator for this new risk measure using plug-in approach for level sets :

$$\widehat{\mathrm{CTE}}_\alpha(X,Y) = \begin{pmatrix} \dfrac{\sum_{i=1}^n X_i 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}} \\ \dfrac{\sum_{i=1}^n Y_i 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}}{\sum_{i=1}^n 1_{\{(X_i,Y_i)\in L_n(\alpha)\}}} \end{pmatrix}.$$

Again, we provide consistency results and some rates of convergences (Theorem 4.1 and corollary 4.1 in [3]). Finally, we confront our estimator on simulated data sets to state its practical usefullness.

This work has led to a paper accepted in ESAIM PS [3].

Références

[1] F. Belzunce, A. Castaño, A. Olvera-Cervantes, and A. Suárez-Llorens. Quantile curves and dependence structure for bivariate distributions. Computational Statistics & Data Analysis, 51(10) :5112–5129, 2007.

[2] A. Cuevas, W. González-Manteiga, and A. Rodríguez-Casal. Plug-in estimation of general level sets. Australian & New Zealand Journal of Statistics, 48(1) :7–19, 2006.

[3] E. Di Bernardino, T. Laloë, V. Maume-Deschamps, and C. Prieur. Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory. ESAIM PS http://hal.archives-ouvertes.fr/docs/00/62/81/00/PDF/CTEbivar.pdf, 2011. Accepted.

[4] P. Embrechts and G. Puccetti. Bounds for functions of multivariate risks. Journal of Multivariate Analysis, 97(2) :526–547, 2006.