Insurance: Mathematics and Economics 57 (2014) 46–57

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Robust and bias-corrected estimation of the coefficient of tail dependence

Christophe Dutang[a], Yuri Goegebeur[b], Armelle Guillou[c,*]

[a] Laboratoire Manceau de Mathématique, Université du Maine, Avenue Olivier Messiaen, 72000 - Le Mans, France
[b] Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark
[c] Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France

ARTICLE INFO

Article history:
Received December 2013
Received in revised form
January 2014
Accepted 8 May 2014

Keywords:
Bias-correction
Tail dependence
Robustness
Tail quantile process
Insurance risk modelling

ABSTRACT

We introduce a robust and asymptotically unbiased estimator for the coefficient of tail dependence in multivariate extreme value statistics. The estimator is obtained by fitting a second order model to the data by means of the minimum density power divergence criterion. The asymptotic properties of the estimator are investigated. The efficiency of our methodology is illustrated on a small simulation study and by a real dataset from the actuarial context.

1. Introduction

Multivariate extreme value statistics deals with the estimation of the tail of a multivariate distribution function based on a random sample. Of particular interest is the estimation of the extremal dependence between two or more variables. Modeling tail dependence is a crucial issue in actuarial science (see e.g. Joe, 2011), firstly, because of the forthcoming Solvency II regulation framework which will oblige insurers and mutuals to compute 99.5% quantiles. Secondly, tail dependence can be used in the daily work of actuaries, for instance for pricing an excess-of-loss reinsurance treaty (see Cebrián et al., 2003, and the references therein), and for approximating very large quantiles of the distribution of the sums of possibly dependent risks (Barbe et al., 2006). In finance, obvious applications also arise, see e.g. Charpentier and Juri (2006), and Poon et al. (2004). Therefore, accurate modeling of extremal events is needed to better understand the relationship of possibly dependent risks at the tail.

A full characterization of the extremal dependence between variables can be obtained from functions like the spectral distribution function or the Pickands dependence function. We refer to Beirlant et al. (2004), and de Haan and Ferreira (2006), and the references therein, for more details about this approach. Alternatively, similar to classical statistics one can try and summarize the extremal dependency in a number of well chosen coefficients that give a representative picture of the full dependency structure. In this paper we will consider the estimation of such a dependency coefficient, namely the coefficient of tail dependence.

The extremal dependence between the components of a continuous random vector (X, Y) with unit Fréchet margins (note that this can be assumed without loss of generality) can be analyzed with the model of Ledford and Tawn (1997):

$$\mathbb{P}(X > x, Y > y) = x^{-d_1} y^{-d_2} \ell(x, y), \quad x, y > 0,$$

where $d_1, d_2 > 0$ and ℓ is a bivariate slowly varying function, i.e.

$$\frac{\ell(tx, ty)}{\ell(t, t)} \to \zeta(x, y) \quad \text{as } t \to \infty, \text{ for all } x, y > 0,$$

and the function ζ is homogeneous of order zero. The parameter $\eta := (d_1 + d_2)^{-1}$ is called the coefficient of tail dependence. It satisfies $\eta \in (0, 1]$, and larger values of it indicate a stronger extremal dependence. As we can imagine, several attempts have been made to estimate η from data. Since

$$\mathbb{P}(\min(X, Y) > z) = \mathbb{P}(X > z, Y > z) = z^{-1/\eta} \ell(z, z),$$

i.e. the transformed variable $\min(X, Y)$ follows a Pareto-type model with index $1/\eta$, one can estimate η with classical estimators for the extreme value index like the Hill (1975), Pickands (1975)

* Corresponding author.
E-mail addresses: christophe.dutang@univ-lemans.fr (C. Dutang), yuri.goegebeur@imada.sdu.dk (Y. Goegebeur), armelle.guillou@math.unistra.fr (A. Guillou).

http://dx.doi.org/10.1016/j.insmatheco.2014.05.003

or moment estimator (Dekkers et al., 1989). However, this type of estimators typically suffers from bias and also they are not robust with respect to outliers. These issues will be addressed in the present paper.

In order to obtain a bias-corrected estimator we will, as usual in extreme value statistics, invoke a second order condition. In particular we will work under the following condition from Draisma et al. (2004), which can be seen as an extension of the above discussed Ledford and Tawn condition.

Condition $\mathcal{SO}$: Let (X, Y) be a random vector with joint distribution function F and continuous marginal distribution functions F_X and F_Y such that

$$\lim_{t\downarrow 0} q_1(t)^{-1}\left(\frac{\mathbb{P}(1-F_X(X)<tx, 1-F_Y(Y)<ty)}{\mathbb{P}(1-F_X(X)<t, 1-F_Y(Y)<t)} - c(x,y)\right) =: c_1(x,y) \quad (1)$$

exists for all $x \geq 0, y \geq 0$ with $x + y > 0$, a function q_1 tending to zero as $t \downarrow 0$, and c_1 a function neither constant nor a multiple of c. Moreover, we assume that the convergence is uniform on $\{(x, y) \in x^2 + y^2 = 1\}$.

Essentially, this condition is a second order multivariate regular variation condition on the function $R(x, y) := \mathbb{P}\big(1 - F_X(X) < x, 1 - F_Y(Y) < y\big)$. It can be shown that $R(t, t)$ is regularly varying at zero with index $1/\eta$, $|q_1|$ is regularly varying at zero with index $\tau \geq 0$, and that the function c is homogeneous of order $1/\eta$, that is $c(tx, ty) = t^{1/\eta}c(x, y)$. Also, $c_1(x, x) = x^{1/\eta}(x^\tau - 1)/\tau$.

The robust and asymptotically unbiased estimator for η will be derived from a second order model obtained from condition ($\mathcal{SO}$), which will be fitted to the data by the minimum density power divergence (MDPD) criterion. The specific second order model will be introduced in the next section. The density power divergence criterion was originally introduced by Basu et al. (1998) for the purpose of developing a robust estimation method. In particular, the density power divergence between density functions f and h is given by

$$\Delta_\alpha(f,h) := \begin{cases} \displaystyle\int_{\mathbb{R}}\left[h^{1+\alpha}(z) - \left(1+\frac{1}{\alpha}\right)h^\alpha(z)f(z) + \frac{1}{\alpha}f^{1+\alpha}(z)\right]dz, & \alpha > 0,\\ \displaystyle\int_{\mathbb{R}}\log\frac{f(z)}{h(z)}f(z)dz, & \alpha = 0.\end{cases}$$

Note that for $\alpha = 0$ one recovers the Kullback–Leibler divergence, whereas setting $\alpha = 1$ leads to the L_2 divergence. Assume that the density function h depends on a parameter vector θ, and let f be the true density function of the random variable under consideration. The idea is then to estimate θ by minimizing an empirical version of Δ_α based on a random sample $Z_1, \dots, Z_n$ from f: if $\alpha > 0$ one considers

$$\widehat{\Delta}_\alpha(\theta) := \int_{\mathbb{R}} h^{1+\alpha}(z)dz - \left(1+\frac{1}{\alpha}\right)\frac{1}{n}\sum_{i=1}^n h^\alpha(Z_i),$$

whereas for $\alpha = 0$

$$\widehat{\Delta}_0(\theta) := -\frac{1}{n}\sum_{i=1}^n \log h(Z_i).$$

For $\alpha = 0$, one fits the model h to the data using the maximum likelihood method. The parameter α controls the trade-off between efficiency and robustness of the MDPD estimator: the estimator becomes more efficient but less robust against outliers as α gets closer to zero, whereas for increasing α the robustness increases and the efficiency decreases.

In Beirlant et al. (2011), an asymptotically unbiased estimator for η was proposed, based on fitting the extended Pareto distribution with the method of maximum likelihood to properly transformed random variables. Goegebeur and Guillou (2013) obtained asymptotic unbiasedness by taking a properly weighted sum of two biased estimators for η. However, these methods are not robust with respect to outliers.

The plan of the paper is as follows. In Section 2, we will introduce a second order Pareto-type model, which is derived from a submodel of condition ($\mathcal{SO}$), and discuss the robust estimation method. In Section 3, the asymptotic properties of our estimator are established. In particular, we will establish a uniform consistency result for the tail quantile process and use this to obtain the limiting distribution of the robust estimator for η. The estimation method is illustrated with a small simulation study in Section 4, and a real dataset concerning workers' compensation in Section 5. Section 6 contains some concluding remarks. The proofs of all results are deferred to the Appendix.

2. Model and assumptions

Let (X, Y) be a bivariate random vector with continuous marginal distributions satisfying

$$\mathbb{P}(1-F_X(X)<x, 1-F_Y(Y)<y) = x^{d_1}y^{d_2}g(x,y) \times \left(1+\frac{1}{\eta}\delta(x,y)\right), \quad x\geq 0, y\geq 0, \quad (2)$$

where d_1, d_2 are positive constants, $\eta := (d_1 + d_2)^{-1} \in (0, 1)$ is the tail dependence coefficient, g is a continuous function that is homogeneous of order 0 and δ is a function of constant sign in the neighborhood of zero, with $|\delta|$ being a bivariate regularly varying function, that is, there exists a function ξ such that

$$\lim_{t\downarrow 0}\frac{|\delta|(tx,ty)}{|\delta|(t,t)} = \xi(x,y),$$

for all $x, y \geq 0$. We assume additionally that ξ is continuous, homogeneous of order $\tau > 0$, and that the convergence is uniform on $\{(x, y) \in [0, \infty)^2|x^2 + y^2 = 1\}$. Note that we exclude the case $\eta = 1$, as was also done in Beirlant and Vandewalle (2002), Beirlant et al. (2011), and Goegebeur and Guillou (2013). For the sequel, it is instructive to keep the following elementary property in mind.

Lemma 1. *Model (2) satisfies assumption ($\mathcal{SO}$).*

Many commonly used joint distribution functions satisfy model (2). Note that this model is in fact a condition on the copula function C. Indeed, one easily verifies that

$$\mathbb{P}(1-F_X(X)<x, 1-F_Y(Y)<y) = x+y-1+C(1-x,1-y).$$

Example 1. The Farlie Gumbel Morgenstern (FGM) distribution. The Farlie Gumbel Morgenstern copula function is given by

$$C(x,y) = xy\left[1+\beta(1-x)(1-y)\right], \quad (x,y)\in[0,1]^2,$$

with $\beta \in [-1, 1]$. Straightforward calculations lead to

$$\mathbb{P}(1-F_X(X)<x, 1-F_Y(Y)<y) = xy[1+\beta-\beta(x+y)+\beta xy].$$

In the case where $\beta \in (-1, 1]$, we get that $d_1 = d_2 = 1$, $\eta = 1/2$, $g(x, y) = 1 + \beta$, $\delta(x, y) = -\eta\beta(x + y - xy)/(1 + \beta)$, $\xi(x, y) = (x + y)/2$ and $\tau = 1$. In terms of condition ($\mathcal{SO}$) this gives then $c(x, y) = xy$, $c_1(x, y) = xy(x+y-2)/2$ and $q_1(t) \sim -2\beta t/(1+\beta)$. In the case $\beta = -1$, we have $d_1 = d_2 = 3/2$, $\eta = 1/3$, $g(x, y) = (x + y)/\sqrt{xy}$, $\delta(x, y) = -xy/(3(x + y))$, $\xi(x, y) = 2xy/(x + y)$ and $\tau = 1$. Condition ($\mathcal{SO}$) is also satisfied with $c(x, y) = xy(x + y)/2$, $c_1(x, y) = xy(2xy - x - y)/2$ and $q_1(t) \sim -t/2$.