

# ESTIMATING THE PROBABILITY OF TWO DEPENDENT CATASTROPHIC EVENTS

LAURENCE LESCOURRET AND CHRISTIAN ROBERT

*University of Cergy Pontoise, France  
Conservatoire National des Arts et Métiers, France*

CREST

Laboratory of Finance and Insurance  
15 Boulevard Gabriel Péri,  
92245 MALAKOFF Cedex, FRANCE  
Tel : (33) 1 41 17 36 05  
Fax : (33) 1 41 17 36 66  
[lescourr@ensae.fr](mailto:lescourr@ensae.fr)  
[christian.robert@cnam.fr](mailto:christian.robert@cnam.fr)

## Abstract

The paper deals with the estimation of the probability that two dependent catastrophic events occur. Because of their nature such events are not often observed. In a two-dimensional space as in a one-dimensional space, the extreme value theory is a powerful tool to do inference in the tail of a distribution outside the range of the observations. This paper considers parametric bivariate extreme value distributions which arise as the limiting distribution of two normalized maxima to estimate the probability of exceedances over high thresholds. The parameter of the extreme value distribution is estimated by using Asymptotic Least Squares (ALS) methods. In case the parameter is a scalar, we derive a new and very simple estimator. Tests for asymptotic independence and overidentifying restrictions are also given.

## Keywords

Probability of catastrophic events, Bivariate extreme value theory, Heavy-tailed distributions, ALS methods.

We thank Charles Lévi for providing us the database

## 1. Introduction

Estimating the probability of catastrophic events from historical data is of much interest to property and casualty insurers. Accurate assessments of the frequency of extreme events are sought in a diversity of applications from environmental impact evaluation to financial risk management.

A first example is the wave height and still-water level which are two important factors for causing floods along a sea coast during storm events. Another example is wind speed data where maximum hourly gusts, maximum hourly mean speeds and the dependence between them are relevant to building safety. Problems involving spatial dependence are multivariate in character: for example, floods may occur at several sites along a coast line, or at various rain-gauges in a national network. In such situation knowledge of the spatial dependence of extremes is essential for risk accumulation assessment. Another example is the study of extremes of stocks prices and stress scenario evaluation.

Understanding relationships among different lines of business is also of much interest for insurers. Belguise [2002] used more than 150,000 individual weather-related domestic insurance claims of a French insurer for a 11-year period (from 1 January 1990 to 31 December 2000) to analyze the dependence structure of storm impact on both motor and household insurance claims. These claims related to damage caused by high winds, tornadoes, lightning, hail, torrential rain and weight of snow, are referred to by insurers as 'storm damage'. There is no uniformly accepted definition of what constitutes storm, and different insurers employ different thresholds of maximum wind gust and temporal scale in order to accept a claim as valid. Most current treaties between reinsurers and insurers define occurrences using a 72 hours scale, which was also chosen by Belguise to obtain 736 storm damages.

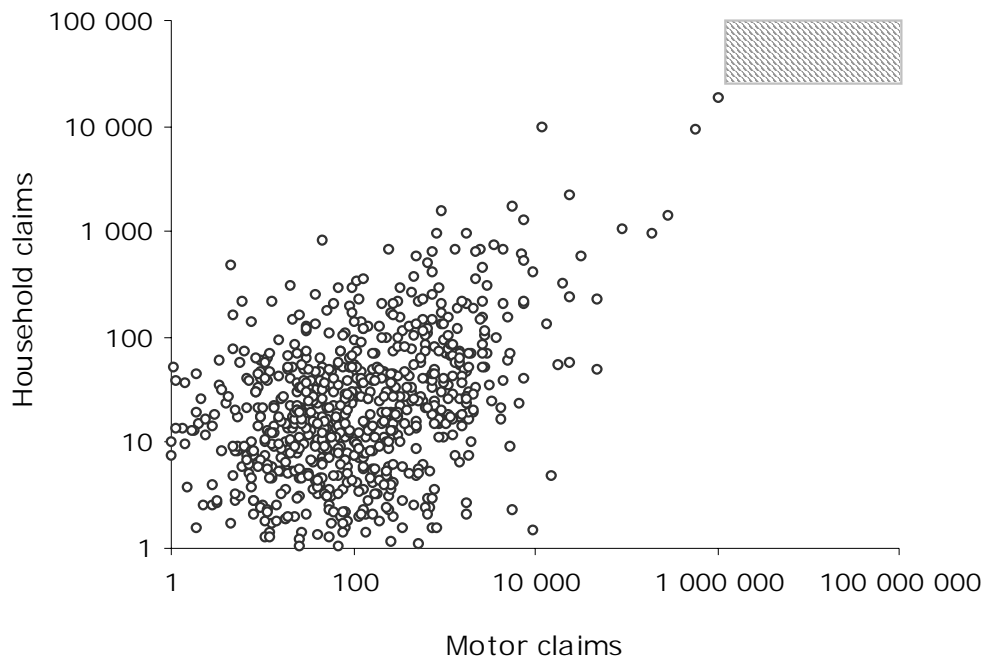


Figure 1: Storm damages (both variables are on a logarithmic scale)

Figure 1 plots Motor claim amounts ( $X$ ) and Household claim amounts ( $Y$ ) of storm events for the period 1990-2000. The data demonstrate an apparent tendency for large claim amounts at both lines. That suggests that temporal and spatial cohesion in the level process induces dependence in the series across lines.

The question we ask in this article is how to estimate the probability  $p$  of the dashed area, i.e. the probability that  $X$  and  $Y$  of the same storm event exceed high thresholds  $p = P(X > t_x \text{ and } Y > t_y)$ . None of the sample points fall into the area, so we can not use the empirical distribution function to estimate  $p$ .

In a two-dimensional space as in a one-dimensional space, if one has to do inference in the tail of a distribution outside the range of the observations, a way to proceed is to use extreme value theory, i.e. to model the tail asymptotically as an extreme-value distribution. Multivariate extreme value theory has recently been an area of much theoretical (Resnick [1987], Galambos [1987], Joe [1997], for example) and practical works (Ledford and Tawn [1996] [1997], Coles and Tawn [1994], de Hann and de Ronde [1998], de Hann and Sinha [1999] among others).

The multivariate extreme value distributions are introduced as limit distributions of normalized componentwise taken maxima. In contrast to the univariate case, no natural parametric family exists for the dependence structure. Thus such a distribution must be modelled in some way, by taking into account restrictions on the dependence structure. The first approach is to model the dependence structure via parametric ad hoc models (see Joe [1993], [1997] or Klüppelberg and May [1998] for explicit classes of models) and to consider appropriate estimating methods: maximum likelihood estimators (Tawn [1990]), bivariate threshold methods (Joe, Smith and Weissman [1992], Coles and Tawn [1991]), or specific estimators for nondifferentiable models (Tawn [1988], Csörgo and Welch [1989]). An alternative approach is to consider non-parametric estimators (Einhmal, de Hann and Huang [1993], de Hann and Resnick [1993], Einhmal, de Hann and Sinha [1997], Einhmal, de Hann and Piterbag [1998], Abdous, Ghoudi and Khoudraji [1998]).

The present paper is concerned with the estimation of parametric bivariate extreme distributions by using Asymptotic Least Squares (ALS) methods. Unlike other authors, we do not assume that the margins are known. The tail parameters have to be estimated as the parameter of the bivariate extreme value distribution. In the case where the parameter is a scalar, we derive a new and very simple estimator. The basic idea behind our approach is a point process representation already widely used for the probabilistic characterization of multivariate extremes (Resnick [1987], de Hann and Resnick [1994] and de Hann and Sinha [1999]).

The bivariate extreme value theory and the point process theory that we need is presented in section 2. At the end of this section we define an estimator  $\hat{p}$  of the probability of catastrophic events. Section 3 tackles on of consistency results. Tests for independence and overidentifying restrictions are also developed. Finally, we present an example of simulated data to illustrate the estimation procedure and the behavior of the new estimator on insurance data. The proofs are gathered in appendices.

## 2. Bivariate extreme value theory

### 2.1. Bivariate extreme value distributions

The bivariate extreme value distributions arise in the following way: let  $\{(X_i, Y_i)\}$  be a

sample of size  $n$  from the bivariate distribution  $F$  (with marginal distributions  $F_X$  and  $F_Y$ ) and suppose there exist sequences of normalizing constants  $a_n > 0$ ,  $c_n > 0$ ,  $b_n$  and  $d_n$  such that

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left( \frac{\max(X_1, \dots, X_n) - b_n}{a_n} \leq x, \frac{\max(Y_1, \dots, Y_n) - d_n}{c_n} \leq y \right) \\ = \lim_{n \rightarrow \infty} F^n(a_n x + b_n, c_n y + d_n) = G(x, y) \end{aligned} \quad (2.1)$$

where  $G$  is a proper nondegenerate bivariate distribution function. It follows that the margins  $G_X$  and  $G_Y$  must be one of the extreme value distributions first identified by Fisher and Tippet [1928]:

$$\begin{aligned} \text{Weibull } (\alpha > 0) \quad \Psi_\alpha(x) &= \exp\{-(-x)^\alpha\} & x \leq 0 \\ \text{Gumbel} \quad \Lambda(x) &= \exp\{-e^{-x}\} & x \in \mathbb{R} \\ \text{Fréchet } (\alpha > 0) \quad \Phi_\alpha(x) &= \exp\{-x^{-\alpha}\} & x \geq 0 \end{aligned}$$

Their results can be considered to have the same status in extreme value theory as the central limit theorem in the study of sums. The class of distributions whose normalized maxima converge in distribution to an extreme value distribution is called the domain of attraction of the extreme value distribution.

The domain of attraction of the Weibull extreme value distribution contains short tailed distributions (with finite endpoint) such as the Uniform and Beta distributions. This class is not of interest in insurance applications.

The domain of attraction of the Gumbel extreme value distribution includes the Normal, Exponential, Gamma, Lognormal and Weibull distributions. We say that  $Z$  has a Weibull-like distribution if its d.f. satisfies  $-\ln(1 - F_Z(z)) := z^\tau l_Z(z)$ , for some slowly varying function  $l_Z$ <sup>1</sup>. The Weibull-like distributions have moderately heavy tails and have historically been a popular model for loss severity distributions.

It can be showed that the distribution function  $F_Z$  belongs to the domain of attraction of the Fréchet extreme value distribution if and only if  $1 - F_Z(z) = z^{-\alpha} l_Z(z)$  for some slowly varying function  $l_Z$ , i.e. the tail of the d.f.  $F_Z$  decays like a power function. The class of Pareto-like distributions is quite large and includes the Pareto, Burr, Loggamma, Cauchy and t-distributions as well as various mixture models. These are all so-called heavy tailed distributions.

General texts on the topic of univariate extreme values include Embrechts, Klüppelberg and Mikosch [1997] and Resnick [1987] among others.

In the rest of the paper we will assume that  $F_X$  is a Weibull-like distribution with parameter  $\tau$  and  $F_Y$  is a Pareto-like distribution with parameter  $\alpha$ . The other cases where  $F_X$  and  $F_Y$  are both Weibull-like or Pareto-like (with different parameters) can be easily derived. The first reason why we choose these distributions is that they allowed a large broad of tail behavior and are quite observed on insurance data. The second reason is that they can

---

<sup>1</sup>  $l_Z(z)$  is a slowly varying function if  $\lim_{z \rightarrow \infty} l_Z(tz)/l_Z(z) = 1$  for each  $t > 0$ .

be easily graphically detected with quantile plots (see Beirlant *et al.* [1996]). If the data are generated from a random sample of the reference distribution, the plot should look roughly linear. This statement remains true if the data come from a linear transformation of the distribution. Some differences in the distributional shape may be also deduced from the plot. For example if the reference distribution has a lighter upper-tail, the plot will curve down at the right of the graph.

Let us consider a sample of size  $n$  of a distribution  $F_X : (X_1, \dots, X_n)$ . Let  $X_{(n)} \leq \dots \leq X_{(1)}$  be order statistics. If  $F_X$  is a Weibull-like distribution (resp. a Pareto-like distribution), we expect to see a straight line in the scatter plot  $(\log(-\log(1-i/(n+1))), \log(X_{(n-i)}))$  (resp.  $(-\log(1-i/(n+1)), \log(X_{(n-i)}))$ ). Sometimes quantile plots for conditional distributions given that  $X > a$ , where  $a$  is a threshold, are more relevant to detect the shape of the distribution tails.

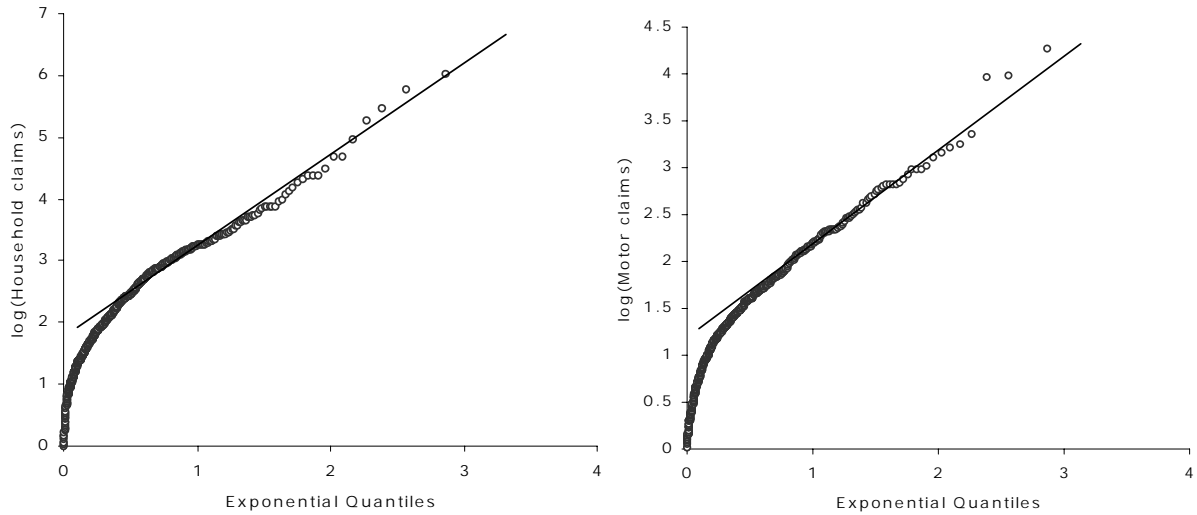


Figure 2: Quantile plots for Motor claim amounts and Household claim amounts

Figure 2 shows two quantile plots for motor claim amounts and household claim amounts. They have been plotted for the whole distribution vs. Pareto quantiles. Empirically both distributions exhibit Pareto tails.

The normalizing constants in (2.1) are determined by (see, for instance, Embrechts, Klüppelberg and Mikosch [1997]):  $a_n = U_X(n)/(\tau \log(n))$ ,  $b_n = U_X(n)$ ,  $c_n = U_Y(n)$  and  $d_n = 0$ , where  $U_X(t) = \inf \{x : F_X(x) \geq 1-t^{-1}\}$  (resp. for  $Y$ ) is the quantile of order  $1-t^{-1}$  and the univariate extreme value distributions are  $G(x, \infty) = \Lambda(x)$  and  $G(\infty, y) = \Phi_\alpha(x)$ .

Since classical characterizations of bivariate extreme value distributions are presented with standard Fréchet distribution for all margins  $\Phi_1(x) = \exp\{-x^{-1}\}$ , we use some transformations on  $x$  and  $y$  and we consider the distribution  $G_*$  given by the condition:

$$\lim_{n \rightarrow \infty} P \left( \frac{\max(X_1, \dots, X_n) - U_X(n)}{U_X(n)/(\tau \log(n))} \leq \log x, \frac{\max(Y_1, \dots, Y_n)}{U_Y(n)} \leq y^\perp \right) = G(\log x, y^\perp) = G_*(x, y).$$

and then  $G_*(x, \infty) = \Phi_1(x)$  and  $G_*(\infty, y) = \Phi_1(y)$ .

As mentioned in introduction, in a multivariate setting, a systematic approach of models seems at first not feasible. Nevertheless there exist various equivalent characterizations among those the one based on  $\chi$  functions. One possible way to write the presentation theorem for  $G_*$  is as follows (see for example Resnick [1987] p.275, Geoffroy [1958], Sibuya [1960]):

**THEOREM 2.1**

$G_*(x, y)$  is a bivariate extreme value distribution with margins  $\Phi_1$  if and only if:

$$G_*(x, y) = \exp \left\{ -\frac{1}{x} - \frac{1}{y} + \frac{1}{y} \chi \left( \frac{y}{x} \right) \right\},$$

for  $x, y > 0$  where  $\chi(t)$ ,  $t \geq 0$ , satisfies:

- (i)  $\chi$  is continuous and concave
- (ii)  $0 \leq \chi(t) \leq \min(1, t)$

If  $\chi(t) = 0$  for any  $t \in \mathbb{R}^+$ , we will say that extremes are asymptotically independent and then  $G_*(x, y) = \Phi_1(x)\Phi_1(y)$ .

Another presentation is based on copulas. A copula is a distribution function with uniform margins. The bivariate extreme value copula is easily derived from Theorem 2.1 :

$$C(u, v) = G_*(\Phi_1^{-1}(u), \Phi_1^{-1}(v)) = \exp \left\{ \ln u + \ln v - \ln v \times \chi \left( \frac{\ln v}{\ln u} \right) \right\}$$

and satisfies the property:  $C(u^\lambda, v^\lambda) = C^\lambda(u, v)$ ,  $\lambda > 0$ .

There is no finite-dimensional parametric family for the dependence function. We must therefore estimate the dependence function by some methods. We restrict here on parametric models, and we assume that the function  $\chi$  can be parameterized by  $\theta \in \mathbb{R}^K$  and it will be denoted  $\chi_\theta$  in the sequel.

*2.2. Point process theory and bivariate extremes*

Our approach follows de Hann and Resnick [1977] and Resnick [1987]: there is a measure  $\nu_*$  concentrating on  $[0, \infty]^2 \setminus \{0\}$  such that:

$$-\log G_*(x, y) =: \nu_* \left( ([0, x] \times [0, y])^c \right)$$

We define a new measure  $\mu^*$  concentrating on  $[0, \infty]^2 \setminus \{0\}$  :

$$\begin{aligned} \mu^* \left( (x, \infty] \times (y, \infty] \right) &:= \nu_* \left( ([0, \infty] \times [0, y])^c \right) + \nu_* \left( ([0, x] \times [0, \infty])^c \right) \\ &- \nu_* \left( ([0, x] \times [0, y])^c \right) = \frac{1}{y} \chi_\theta \left( \frac{y}{x} \right) =: \Phi_\theta(x, y). \end{aligned}$$

Some examples are given below for different sizes  $K$  of the parameter  $\theta$ .

EXAMPLES 2.1 for  $K = 1$ :

- **The Marshall-Olkin model:** for  $\theta \in [0, 1]$

$$\Phi_{\theta}(x, y) = \theta \left( x^{-1} + y^{-1} - \min(x^{-1}, y^{-1}) \right)$$

For  $\theta = 0$  we have independence, for  $\theta = 1$ , we have total dependence. The distribution  $G_*$  is continuous, but a density does not exist for  $\theta > 0$ , because Marshall-Olkin distributions have positive mass at the main diagonal.

- **The Hüsler-Reiss model:** for  $\theta \in ]0, \infty[$

$$\Phi_{\theta}(x, y) = \bar{\Psi} \left( \theta^{-1} + \left( \theta \log(xy^{-1}) \right) / 2 \right) y^{-1} + \bar{\Psi} \left( \theta^{-1} + \left( \theta \log(yx^{-1}) \right) / 2 \right) x^{-1}$$

where  $\bar{\Psi}$  is the survival d.f. of the standard gaussian distribution. For  $\theta \rightarrow 0$  we have independence, for  $\theta \rightarrow \infty$ , we have total dependence.

- **The Logistic model or Gumbel model:** for  $\theta \in ]0, 1]$

$$\Phi_{\theta}(x, y) = x^{-1} + y^{-1} - \left( x^{-\frac{1}{\theta}} + y^{-\frac{1}{\theta}} \right)^{\theta}$$

Independence and total dependence correspond to  $\theta = 1$  and  $\theta \rightarrow 0$  respectively

- **The mixed model:**

$$\Phi_{\theta}(x, y) = \theta (x + y)$$

For  $\theta = 0$  we have independence, but we cannot have complete dependence.

EXAMPLE 2.2 for  $K = 2$ :  $\theta = (\alpha, \beta)$

- **The asymmetric mixed model:**  $\alpha \geq 0$ ,  $\alpha + \beta \leq 1$ ,  $\alpha + 2\beta \leq 1$ ,  $\alpha + 3\beta \geq 0$

$$\Phi_{\theta}(x, y) = [(\alpha + 2\beta)x + (\alpha + \beta)y] / (x + y)^2$$

Independence corresponds to  $\alpha = \beta = 0$ , a corner of the parameter space, but we cannot have complete dependence.

EXAMPLE 2.3 for  $K = 3$ ,  $\theta = (\alpha, \beta, \gamma)$

- **The asymmetric logistic model:**  $\alpha \geq 0$ ,  $\beta \geq 1$ ,  $\gamma \geq 1$

$$\Phi_{\theta}(x, y) = \beta x^{-1} + \alpha y^{-1} - \left( (\alpha x)^{\frac{1}{\gamma}} + (\beta y)^{\frac{1}{\gamma}} \right)^{\gamma} (xy)^{-1}$$

Note that, when  $\alpha = \beta = 1$ , we get the logistic model. Total dependence corresponds to  $\alpha = \beta = 1$  and  $\gamma \rightarrow \infty$ , whereas independence corresponds to  $\alpha = \beta = 0$ .

Given the latter definitions we re-write Condition (2.1) as:

$$\begin{aligned} (n/k)P\left\{\{(X_1 - b_{n/k})/a_{n/k} \geq \log x\} \cup \{(Y_1 - d_{n/k})/c_{n/k} \geq y^{\frac{1}{\alpha}}\}\right\} &\rightarrow \nu^*([0, x] \times [0, y]^c), \\ (n/k)P\left\{\{(X_1 - b_{n/k})/a_{n/k} \geq \log x\} \cap \{(Y_1 - d_{n/k})/c_{n/k} \geq y^{\frac{1}{\alpha}}\}\right\} &\rightarrow \mu^*((x, \infty) \times (y, \infty)), \end{aligned}$$

for  $(x, y) \in [0, \infty]^2 \setminus \{0\}$  and where  $k = k(n)$  is a sequence such that  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$ .

Estimating  $G$  is equivalent to estimating  $\nu^*$  or  $\mu^*$ . It is natural to consider empirical measures as candidates for the estimators. For  $(x, y) \in [0, \infty]^2 \setminus \{0\}$  and  $A \subset [0, \infty]^2 \setminus \{0\}$  define:

$$\mathcal{E}_{(x,y)}(A) = \begin{cases} 1 & \text{if } (x, y) \in A \\ 0 & \text{if } (x, y) \in A^c \end{cases}$$

For  $k \leq n$  the estimators of measures are given by:

$$\begin{aligned} \hat{\nu}_n^*([0, x] \times [0, y]^c) &:= \frac{1}{k} \sum_{i=1}^n \mathcal{E}_{\exp\left(\frac{X_i - \hat{b}_{n/k}}{\hat{a}_{n/k}}\right), \left(\frac{Y_i - \hat{d}_{n/k}}{\hat{c}_{n/k}}\right)^{\hat{\alpha}_k}}([0, x] \times [0, y]^c), \\ \hat{\mu}_n^*((x, \infty) \times (y, \infty)) &:= \frac{1}{k} \sum_{i=1}^n \mathcal{E}_{\exp\left(\frac{X_i - \hat{b}_{n/k}}{\hat{a}_{n/k}}\right), \left(\frac{Y_i - \hat{d}_{n/k}}{\hat{c}_{n/k}}\right)^{\hat{\alpha}_k}}(x, \infty) \times (y, \infty). \end{aligned}$$

where  $\hat{a}_{n/k} = \hat{U}_X(n/k)/\hat{\tau}_k(\log(n/k))$ ,  $\hat{b}_{n/k} = \hat{U}_X(n/k)$ ,  $\hat{c}_{n/k} = \hat{U}_Y(n/k)$ ,  $\hat{d}_{n/k} = 0$  with  $\hat{U}_X(n/k) = X_{(k)}$ ,  $\hat{U}_Y(n/k) = Y_{(k)}$ , and  $\hat{\tau}$  and  $\hat{\alpha}$  are respectively Beirlant et al. [1995] estimator and Hill estimator:

$$\frac{1}{\hat{\tau}_k} = \frac{\log(n/k)}{X_{(k)}} \frac{1}{k} \sum_{i=1}^k (X_{(i)} - X_{(k+1)}), \quad \frac{1}{\hat{\alpha}_k} = \frac{1}{k} \sum_{i=1}^k (\log Y_{(i)} - \log Y_{(k+1)}).$$

A general text on the subject about tail parameters is Beirlant et al. [1996].

### 2.3. An estimator of the probability of dependent catastrophic events

This section proposes an estimator of the probability  $p$ . According to the previous results it seems natural to consider the following threshold  $t_X(n) = a_n \log x + b_n$  and  $t_Y(n) = c_n y^{\frac{1}{\alpha}} + d_n$  since:

$$\frac{n}{k} P\left\{\{X > t_X(n/k)\} \cap \{Y > t_Y(n/k)\}\right\} \rightarrow \Phi_\theta(x, y)$$

where  $k = k(n)$  is a sequence such that  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$ . If we have an estimator of  $\theta$  such that  $\hat{\theta}_{k,n} P \rightarrow \theta$ , a natural estimator of the probability of  $p$  is given by:

$$\hat{p}_n = \frac{k}{n} \Phi_{\hat{\theta}_{k,n}}(x(t_X, k), y(t_Y, k))$$

where:

$$x(t_X, k) = \exp\left(\frac{t_X - X_{(k)}}{\frac{1}{k} \sum_{i=1}^k (X_{(i)} - X_{(k+1)})}\right) = \exp\left(\frac{t_X - \hat{b}_{n/k}}{\hat{a}_{n/k}}\right)$$

$$y(t_Y, k) = \exp\left(\frac{1}{\frac{1}{k} \sum_{i=1}^k (\log Y_{(i)} - \log Y_{(k+1)})} \log\left(\frac{t_Y}{Y_{(k)}}\right)\right) = \left(\frac{t_Y}{\hat{c}_{n/k}}\right)^{\hat{\alpha}}$$

Moreover if  $\hat{\theta}_{k,n}$  is asymptotically normal such that  $\sqrt{k}(\hat{\theta}_{k,n} - \theta)d \rightarrow N(0, V(\theta))$ , then:

$$\sqrt{k}(\Phi_{\hat{\theta}_{k,n}}(x, y) - \Phi_{\theta}(x, y))d \rightarrow N(0, \Sigma(\theta, x, y)),$$

where  $\Sigma(\theta, x, y) = (\nabla_{\theta} \Phi_{\theta}(x, y))' V(\theta) (\nabla_{\theta} \Phi_{\theta}(x, y))$  and  $\nabla_{\theta} \Phi_{\theta}(x, y) = \partial \Phi_{\theta}(x, y) / \partial \theta$  is the gradient of  $\Phi_{\theta}$  with respect to  $\theta$ .

A 95% confidence interval of the probability is given by:

$$\hat{p}_n \in \left[ \frac{k}{n} \Phi_{\hat{\theta}_{k,n}}(x(t_X, k), y(t_Y, k)) \pm 1.96 \frac{\sqrt{k}}{n} \sqrt{\Sigma(\hat{\theta}_{k,n}, x(t_X, k), y(t_Y, k))} \right].$$

## 2. Properties of measure estimators

### 2.1. Consistency and asymptotic normality

We begin by discussing the consistency of the estimators, then we introduce a second order condition to establish the asymptotic normality. The proofs are gathered in Appendix 1.

#### PROPOSITION 3.1

Suppose that (2.1) holds. Let  $k = k(n)$  be a sequence such that:  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$ , then we have:

$$\hat{\nu}_n^*([0, x] \times [0, y])^c P \rightarrow \frac{1}{x} + \frac{1}{y} - \frac{1}{y} \mathcal{X}_{\theta}\left(\frac{y}{x}\right),$$

$$\hat{\mu}_n^*((x, \infty] \times (y, \infty]) P \rightarrow \frac{1}{y} \mathcal{X}_{\theta}\left(\frac{y}{x}\right) = \Phi_{\theta}(x, y).$$

for any  $x$  and  $y$  in  $\mathbb{R}_*^+$ . Moreover if  $k$  is such that  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ ,  $k/\log n \rightarrow \infty$ , as  $n \rightarrow \infty$ , then the estimators converge almost surely.

#### REMARK:

If  $x = 1$  and  $y = 1$ , then:

$$\hat{\mu}_n^*((1, \infty] \times (1, \infty]) = \frac{1}{k} \sum_{i=1}^n \mathcal{E}_{\{X_i > X_{(k)}\} \cap \{Y_i > Y_{(k)}\}} P \rightarrow \mathcal{X}_{\theta}(1).$$

If  $\theta$  is of size 1, it is possible to get an estimator of the links between extremes without estimating tail parameters of margins.

EXAMPLE 3.1: Let us consider the Marshall-Olkin model ( $\theta \in [0,1]$ ). Since  $\chi_\theta(1) = \Phi_\theta(1,1) = \theta$ ,  $\hat{\mu}_n^*((1,\infty] \times (1,\infty])$  is an estimator of  $\theta$ .

Having shown consistency of estimators, we now focus on asymptotic normality. In order to get a rate of converge, the second order refinement will be assumed: there exists a non constant function  $\psi$  from  $\mathbb{R}^2$  to  $\mathbb{R}$  and a positive function  $c$  such that:

$$\lim_{n \rightarrow \infty} \frac{n(1 - F(a_n x + b_n, c_n y + d_n)) - G(x, y)}{c(U_X(n), U_Y(n))} = \psi(x, y) < \infty,$$

locally uniformly for  $(x, y) \in (-\infty, \infty] \times (0, \infty]$ , where  $c(U_X(\cdot), U_Y(\cdot))$  is regularly varying,  $c(U_X(t), U_Y(t)) \rightarrow 0$  as  $t \rightarrow \infty$ . For more details, see de Hann and Stadtmüller [1994].

### PROPOSITION 3.2

Suppose that (2.1) holds and  $k = k(n)$  is such that  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$  and  $\lim_{n \rightarrow \infty} \sqrt{k} c(U_X(n/k), U_Y(n/k)) = 0$  then:

$$\begin{aligned} & \sqrt{k} \left( \hat{\mu}_n^*((x, \infty] \times (y, \infty]) - \frac{1}{y} \chi_\theta \left( \frac{y}{x} \right) \right) d \rightarrow V(x, y) =: W(x, \infty) + W(\infty, y) - W(x, y) \\ & + \frac{\partial \Phi_\theta(x, y)}{\partial x} x (W(1, \infty) + \tau \log x \Gamma_\tau) + \frac{\partial \Phi_\theta(x, y)}{\partial y} y (W(\infty, 1) + \alpha \log y \Gamma_\alpha). \end{aligned}$$

where  $W$  is a zero mean Gaussian random field ( $W(\{x, y\}), \{x, y\} \in [0, \infty]^2 \setminus \{0\}$ ) with covariance function:

$$\text{cov} \left( W(\{x, y\}), W(\{x', y'\}) \right) = \nu_* \left( ([0, x] \times [0, y])^c \cap ([0, x'] \times [0, y']^c) \right),$$

and

$$\Gamma_\tau = \int_0^\infty W(e^{tu}, \infty) du - \frac{1}{\tau} W(1, \infty), \quad \Gamma_\alpha = \int_1^\infty W(\infty, v^\alpha) \frac{dv}{v} - \frac{1}{\alpha} W(\infty, 1).$$

### 3.2. An estimator of $\theta$ when $K = 1$

In the case where  $\theta$  is a scalar and  $\Phi_\theta(1,1)$  is invertible with respect to  $\theta$ , an estimator of  $\theta$ ,  $\hat{\theta}$ , is easily derived and the asymptotic normality of  $\hat{\mu}_n^*((1,\infty] \times (1,\infty]) = \hat{\Phi}_\theta(1,1)$  gives the asymptotic normality of  $\hat{\theta}$ . A simple expression of the asymptotic variance can be also derived:

$$\sqrt{k} \left( \hat{\mu}_n^*((1,\infty] \times (1,\infty]) - \Phi_\theta(1,1) \right) d \rightarrow N(0, \Omega_\theta),$$

with:

$$\Omega_\theta = \left( \frac{\partial \Phi_\theta(1,1)}{\partial x} \right)^2 + \left( \frac{\partial \Phi_\theta(1,1)}{\partial y} \right)^2 + \Phi_\theta(1,1) \left( 2 \left( 1 + \frac{\partial \Phi_\theta(1,1)}{\partial x} \right) \left( 1 + \frac{\partial \Phi_\theta(1,1)}{\partial y} \right) - 1 \right).$$

Since  $\Phi_\theta$  is homogeneous of order  $-1$ :

$$\frac{\partial \Phi_\theta(x, y)}{\partial x} x + \frac{\partial \Phi_\theta(x, y)}{\partial y} y = -\Phi_\theta(x, y).$$

Distributions with a scalar parameter are symmetric in  $x$  and  $y$ , which implies that:

$$\frac{\partial \Phi_\theta(1, 1)}{\partial x} + \frac{\partial \Phi_\theta(1, 1)}{\partial y} = 2 \frac{\partial \Phi_\theta(1, 1)}{\partial x} = 2 \frac{\partial \Phi_\theta(1, 1)}{\partial y} = -\Phi_\theta(1, 1),$$

and:

$$\Omega_\theta = \frac{1}{2} \Phi_\theta(1, 1) (\Phi_\theta(1, 1) - 1) (\Phi_\theta(1, 1) - 2).$$

EXAMPLE 3.2: For the Logistic model ( $\theta \in ]0, 1[$ )

$$\Phi_\theta(1, 1) = 2(1 - 2^{\theta-1}) \quad \text{and} \quad \Omega_\theta = (1 - 2^{\theta-1})(2^\theta - 1)2^\theta.$$

Then we have:

$$\sqrt{k_n} (\hat{\theta} - \theta) d \rightarrow N \left( 0, \frac{(1 - 2^{\theta-1})(2^\theta - 1)2^\theta}{(2^\theta \log 2)^2} \right),$$

where the estimator of  $\theta$  is given by  $\hat{\theta} = \log(2 - \hat{\Phi}_\theta(1, 1)) / \log 2$ .

EXAMPLE 3.3: For the Hüsler and Reiss model ( $\theta \in ]0, \infty[$ )

$$\Phi_\theta(1, 1) = 2 \left( 1 - \Psi(\theta^{-1}) \right) \quad \text{and} \quad \Omega_\theta = 2 \left( 1 - \Psi(\theta^{-1}) \right) (2\Psi(\theta^{-1}) - 1) \Psi(\theta^{-1}).$$

Then we have:

$$\sqrt{k} (\hat{\theta} - \theta) d \rightarrow N \left( 0, \frac{2\theta^4 \left( 1 - \Psi(\theta^{-1}) \right) (2\Psi(\theta^{-1}) - 1) \Psi(\theta^{-1})}{4(\Psi')^2(\theta^{-1})} \right),$$

where the estimator of  $\theta$  is given by  $\hat{\theta} = \left( \Psi^{-1} \left( 1 - \hat{\Phi}_\theta(1, 1)/2 \right) \right)^{-1}$ .

### 3.3. A test for independence

The componentwise maxima are asymptotically independent for many distributions (for instance bivariate normal distributions with correlation coefficient strictly smaller than one). The independence play an important role in many applications and it is possible to define moderate dependence in the class of asymptotically independent distribution (see Ledford and Tawn [1997]).

In this section we consider a test for asymptotically independence. If the test is not rejected then Proposition 3.2 is not valid and the asymptotic Least Squares method discussed below is not necessary.

#### PROPOSITION 3.3.

Suppose that (2.1) holds and  $k = k(n)$  is such that  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$  and  $\lim_{n \rightarrow \infty} \sqrt{k} c(U_X(n/k), U_Y(n/k)) = 0$  then:

$$\sqrt{k} \left( \hat{\nu}_n^*((x, \infty] \times (y, \infty]) - \left( \frac{1}{x} + \frac{1}{y} - \Phi_\theta(x, y) \right) \right) d \rightarrow W(x, y) - \left( \frac{1}{x^2} + \frac{\partial \Phi_\theta(x, y)}{\partial x} \right) x (W(1, \infty) + \tau \log x \Gamma_\tau) - \left( \frac{1}{y^2} + \frac{\partial \Phi_\theta(x, y)}{\partial y} \right) y (W(\infty, 1) + \alpha \log y \Gamma_\alpha).$$

The null hypothesis of independence is:

$$H_0 : \nu^*((x, \infty] \times (y, \infty]) = \frac{1}{x} + \frac{1}{y}.$$

Under  $H_0$  the asymptotic law is given by:

$$\sqrt{k} \left( \hat{\nu}_n^*((x, \infty] \times (y, \infty]) - \frac{1}{x} + \frac{1}{y} \right) d \rightarrow W(x, y) - \frac{1}{x} (W(1, \infty) + \tau \log x \Gamma_\tau) - \frac{1}{y} (W(\infty, 1) + \alpha \log y \Gamma_\alpha).$$

#### COROLLARY 3.4.

Suppose that (2.1) holds and  $k = k(n)$  is such that  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$  and  $\lim_{n \rightarrow \infty} \sqrt{k} c(U_X(n/k), U_Y(n/k)) = 0$ . Let  $x < 1$  and  $y < 1$ . Under the null hypothesis  $H_0$ , the statistic:

$$\xi_n^I = k \frac{\left( \hat{\nu}_n^*((x, \infty] \times (y, \infty]) - \frac{1}{x} + \frac{1}{y} \right)^2}{\left( \frac{1}{x^2} (1 + (\log x)^2 - x) + \frac{1}{y^2} (1 + (\log y)^2 - y) \right)}$$

is asymptotically distributed as a  $\chi^2(1)$  distribution. The test with critical region  $\{\xi_n^I \geq \chi_{1-\alpha}^2(1)\}$  has asymptotic level  $\alpha$  and is consistent.

#### 4. Asymptotic Least Squares methods (ALS)

The basic principle of ALS methods is to choose the parameter estimate so that the distance between the corresponding sample conditions and the theoretical conditions are equal to 0. We assume that extremes are asymptotically dependent and  $\theta$  is of size  $K$ . Suppose that the true value  $\theta_0$  is characterized by the  $H$  orthogonality conditions:

$$\mu((x_i, \infty] \times (y_i, \infty]) - \frac{1}{y_i} \chi_{\theta_0} \left( \frac{y_i}{x_i} \right) = 0, i = 1, \dots, H.$$

Its sample analogue is the sample mean evaluated at some hypothetical value  $\theta$ :

$$\hat{\mu}_n^*((x_i, \infty] \times (y_i, \infty]) - \frac{1}{y_i} \chi_\theta \left( \frac{y_i}{x_i} \right), i = 1, \dots, H. \quad (4.1)$$

The following notation will be used in the sequel:  $\mathbf{x}$  denotes a generic vector of size  $H$  and

$$\hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_\theta \left( \frac{\mathbf{y}}{\mathbf{x}} \right)$$

is the vector of size  $H$  with components given by (4.1).

If the number of orthogonality conditions is equal to the number of parameters (i.e.  $K = H$ ), then applying the ALS principle amounts to choosing  $\theta_0$  such that:

$$\hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) = 0 \quad (4.2)$$

But, if there are more orthogonality conditions than parameters (i.e.  $H > K$ ), the system may not have a solution. The basic idea of ALS methods is to choose  $\theta_0$  so that (4.2) is as close to 0 as possible.

One way to proceed is to define:  $\theta_n(S_n) = \arg \min_{\theta} J(\theta, S_n)$  where:

$$J(\theta, S_n) = \left[ \hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_n \left[ \hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]$$

That is, the ALS estimator  $\theta_n(S_n)$  is the value of  $\theta$  that minimizes the objective function  $J(\theta, S_n)$ , where  $S_n$  is a  $(H, H)$  symmetric positive definite matrix sometimes, called the weighting matrix.

Under technical assumptions defined in Appendix 2, the ALS estimator is consistent. The proofs are also gathered in Appendix 2.

**PROPOSITION 4.1.:**

*The ALS estimator associated with  $S_n$  exists asymptotically and converges to  $\theta_0$ .*

The asymptotically multivariate normality of  $\hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty])$  is easily derived from a sample modification of Proposition 3.2. If the sequence  $k$  satisfies  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ ,  $\lim_{n \rightarrow \infty} \sqrt{k} c(U_X(n/k), U_Y(n/k)) = 0$ , as  $n \rightarrow \infty$ , then:

$$\sqrt{k} \left( \hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right) d \rightarrow N(0, \Omega_{\theta_0}(\mathbf{x}, \mathbf{y})),$$

where:

$$\Omega_{\theta_0}(\mathbf{x}, \mathbf{y})_{i,j} = \text{cov}(V(x_i, y_i), V(x_j, y_j)).$$

The asymptotic normality of the ALS estimator is then obtained by adding some extra technical assumptions (see Appendix 2).

**PROPOSITION 4.2.:**

*The ALS estimator associated with  $S_n$  is asymptotically normally distributed such that:*

$$\sqrt{k}(\theta_n(S_n) - \theta_0) d \rightarrow N(0, \Sigma(S_0, \mathbf{x}, \mathbf{y})),$$

*and:*

$$\Sigma(S_0, \mathbf{x}, \mathbf{y}) = \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1} \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \Omega_{\theta_0}(\mathbf{x}, \mathbf{y}) S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right] \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1}.$$

where  $\Phi_{\theta_0}(x, y)$  is the vector  $(\chi_{\theta_0}(y_i/x_i)/y_i)_{i=1, \dots, H}$ .

Now we would like to know whether there exists an optimal choice for the sequence matrix  $S_n$ , or equivalently whether there exists an optimal choice for  $S_0$ , since the asymptotic variance covariance matrix depends only on the non-random matrix  $S_0$ , if

**PROPOSITION 4.3.:**

*There exists a best estimator which is obtained when the limit matrix is  $S_0^* = \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y})$ .*

*Its asymptotic variance covariance matrix is equal to:*

$$\Sigma(S_0^*, \mathbf{x}, \mathbf{y}) = \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y}) \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1}.$$

This description of  $S_0^*$  is somewhat circular because  $\theta_0$  is an estimate that we want to obtain by the ALS estimation. The practical procedure, called the two-step efficient ALS procedure, is as follows:

- Step 1:  $S_n^1 = Id_H$  and  $\theta_n^1(S_n^1)$  is computed
- Step 2:  $S_n^2 = \Omega_{\theta_n^1(S_n^1)}^{-1}(\mathbf{x}, \mathbf{y})$  and  $\theta_n^2(S_n^2)$  is computed

The iterative GMM estimator repeats this procedure until  $\theta_n^i$  converges, but in theory the two previous steps are just needed.

Now we would like to test the validity of the model with family  $\chi_{\theta}$  by using overidentifying restrictions. We know that, for the true value  $\theta_0$  and for a sequence matrix  $S_n^* n \rightarrow \infty \rightarrow \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y})$

$$kJ\left(\theta, \Omega_{\theta_n^1(S_n^1)}^{-1}(\mathbf{x}, \mathbf{y})\right) d \rightarrow \chi^2(K)$$

If  $\theta_0$  is replaced by  $\theta_n(S_n^*)$ , then the number of freedom degrees changes from  $H$  to  $H - K$  in the overidentified case. Then it is possible to test whether every restrictions of the model are satisfied.

Let the null hypothesis be:

$$H_0 : \exists \theta_0 \text{ such that } \mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) = 0,$$

**PROPOSITION 4.4.:**

*Let  $\theta_n^* = \theta_n(S_n^*)$  be the best ALS estimator (i.e. such that  $S_n^* n \rightarrow \infty \rightarrow \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y})$ ). Under the null hypothesis  $H_0$ , the statistic:*

$$\xi_n^{over} = k \left[ \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \frac{1}{\mathbf{y}} \chi_{\theta_n^*} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_n^* \left[ \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \frac{1}{\mathbf{y}} \chi_{\theta_n^*} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]$$

is asymptotically distributed as a  $\chi^2(H-K)$ . The test with critical region  $\{\xi_n \geq \chi_{1-\alpha}^2(H-K)\}$  has an asymptotic level  $\alpha$  and is consistent.

REMARK 4.1.: Estimators defined in Section 3.2 (see examples 3.2 and 3.3) are (ALS) estimators built with the orthogonality condition

$$\hat{\mu}_n^*([1; \infty] \times [1; \infty]) - \Phi_\theta(1,1) = 0.$$

## 5. Applications

### 5.1. Simulations

Three samples of one thousand bivariate data were generated for the Hüsler and Reiss model ( $\theta = 0.5$ ,  $\theta = 1$ ,  $\theta = 2$ ) by using the statistical software Xtremes. Several estimators of  $\hat{\theta}$  was computed for different values of  $k$ . Let us recall that the Hüsler and Reiss model is defined by:

$$\Phi_\theta(x, y) = \bar{\Psi} \left( \theta^{-1} + \left( \theta \log(xy^{-1}) \right) / 2 \right) y^{-1} + \bar{\Psi} \left( \theta^{-1} + \left( \theta \log(yx^{-1}) \right) / 2 \right) x^{-1},$$

with  $\theta \in ]0, \infty]$  and the estimator of  $\theta$  is given by:

$$\hat{\theta} = \left( \Psi^{-1} \left( 1 - \hat{\Phi}_\theta(1,1) / 2 \right) \right)^{-1}, \text{ with } \hat{\Phi}_\theta(1,1) = \frac{1}{k} \sum_{i=1}^n \mathcal{E}_{\{X_i > X_{(k)}\} \cap \{Y_i > Y_{(k)}\}}.$$

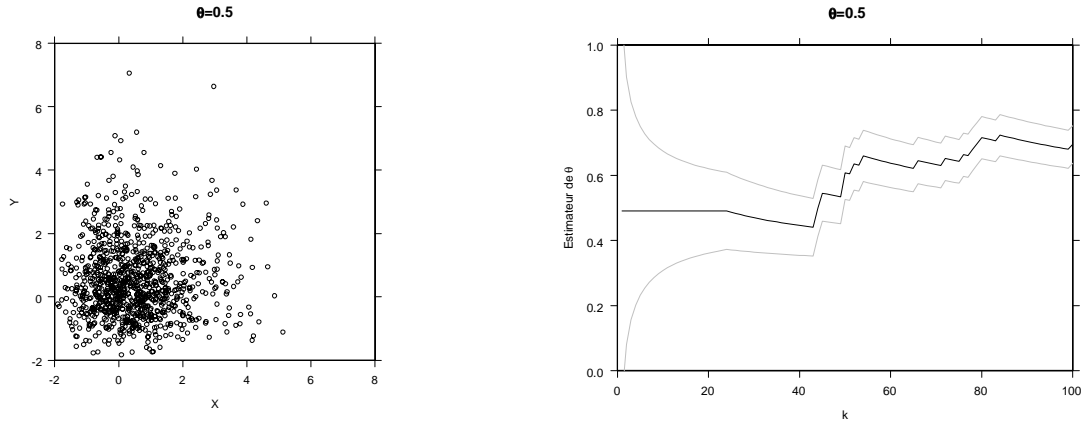


Figure 4.1: Simulated data and estimators of  $\hat{\theta}$  ( $\theta = 0.5$ )  
(Spearman correlation coefficient:  $\rho = 0.026$ )

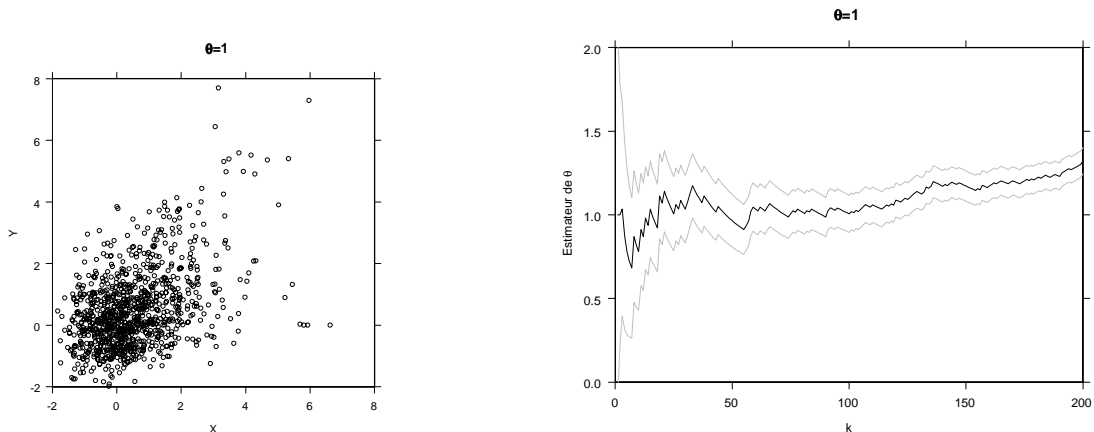


Figure 4.2: Simulated data and estimators of  $\hat{\theta}$  ( $\theta = 1$ )  
 (Spearman correlation coefficient:  $\rho = 0.438$ )

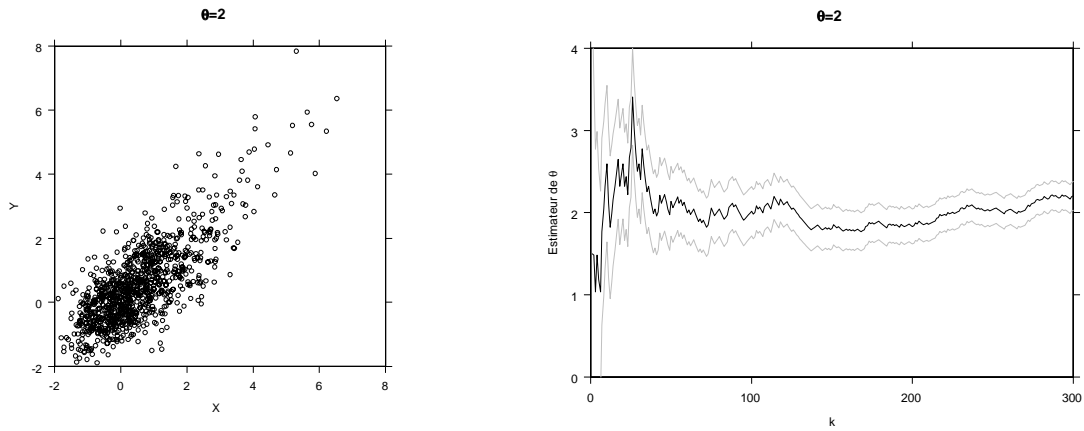


Figure 4.3: Simulated data and estimators of  $\hat{\theta}$  ( $\theta = 2$ )  
 (Spearman correlation coefficient:  $\rho = 0.773$ )

Let us remember that for  $\theta \rightarrow 0$  we have independence, for  $\theta \rightarrow \infty$ , we have total dependence. The estimator is quite good on small samples.

### 1.2. Insurance application: storm damages

In this sub-section the method developed in Section 2 and 3 are applied to the problem outlined in introduction. The data we will consider consist of 736 pairs of storm damages: motor claim amounts and household claim amounts (see Figure 1). The two outliers correspond to the windstorms Lothar and Martin passed over France in December of 1999 causing the largest insured catastrophe loss in Europe's history.

We first estimate tail parameters. Quantile plots have shown that both distributions exhibit Pareto tails. In Figure 3(a) Hill estimators are plotted with respect to the  $k$  upper order statistics.

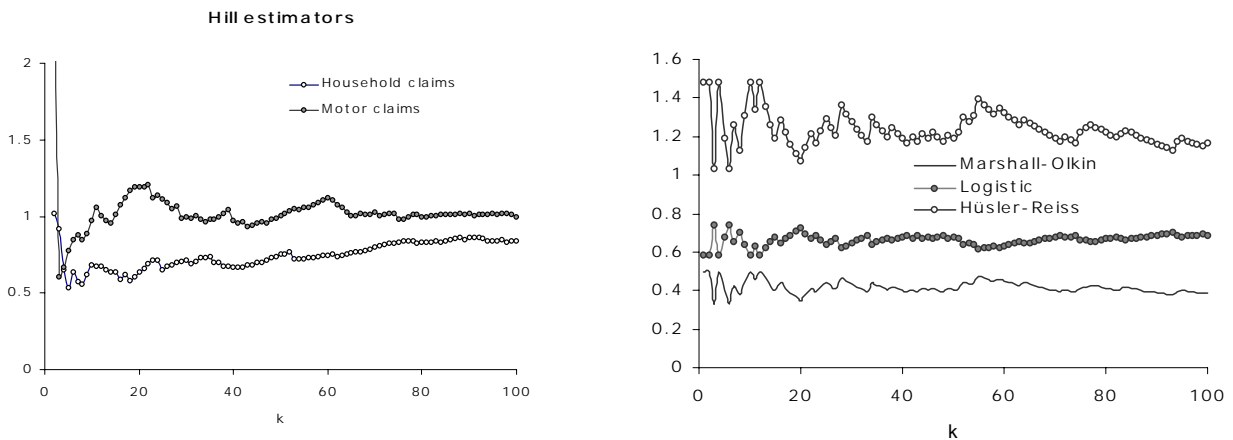


Figure 3: (a) Hill estimators for Household claim amounts and Motor claim amounts  
 (b)  $\hat{\theta}$  for the Marshall-Olkin model, the logistic model and the Hüsler-Reiss model

The tail parameter estimates  $\hat{\tau}_{X,k}$  and  $\hat{\tau}_{Y,k}$  should be inferred from a stable region in the graph. But the choice of  $k$  which is a critical parameter depends also on the dependence structure. Figure 3(b) gives the estimators  $\hat{\Phi}_\theta(1,1)$  (Marshall and Olkin curve) based on the pairs  $(X_{(k)}, Y_{(k)})$ . As a completion of these graphs the previous simulations can also be used to decide on the optimal  $k$ . Finally we choose  $k = 60$ , which gives  $X_{(60)} = 2,054$ ,  $Y_{(60)} = 205$ ,  $\hat{\tau}_{X,60} = 0.754$ ,  $\hat{\tau}_{Y,60} = 1.104$ . The graph on Figure 4 shows the claim amounts after normalization transformations:  $\left( \left( X_i / X_{(60)} \right)^{\hat{\tau}_{X,60}}, \left( Y_i / Y_{(60)} \right)^{\hat{\tau}_{Y,60}} \right)$  where  $X_i \geq 0.1 \times X_{(60)}$  and  $Y_i \geq 0.1 \times Y_{(60)}$ .

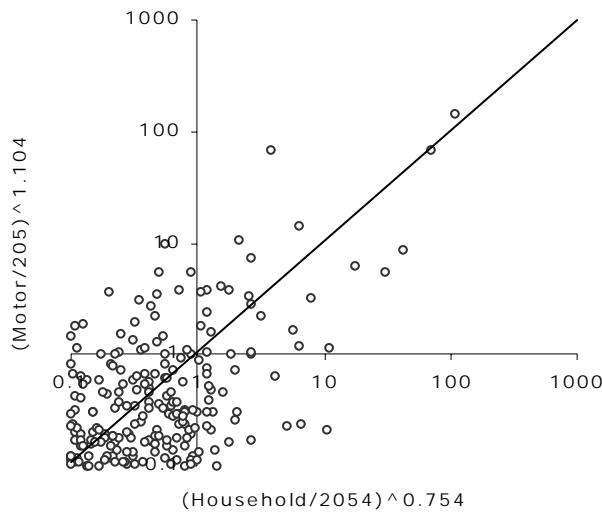


Figure 4: Claim amounts after transformations (both variables are on a logarithmic scale)

The independence test gives  $\xi_n^I = 9.53$  for  $x = 0.5$  and  $y = 0.5$  which clearly rejects the null hypothesis of asymptotic independence. Now we have to decide which model to use. The biplot on Figure 4 is quite symmetric with respect to the main diagonal and thus a model with a symmetric function i.e.  $\Phi_\theta(x, y) = \Phi_\theta(y, x)$  seems a good choice. We will consider the Mixed model, the Logistic model and the Hüsler-Reiss model which have all a parameter of size 1. Figure 3(b) gives estimates of the parameters

$k = 60$	Mixed model	Logistic Model	Hüsler-Reiss model
$\hat{\theta}_k$	0.9	0.632	1.323
$\sigma_k(\theta)$	0.113	0.0526	0.165
$\xi_{n,k}^{over}$	0.21	0.16	0.14

Table 1: Comparison of bivariate extreme value distribution

Table 1 gives for each model the parameter estimate, its standard error and the statistic of a test for three overidentifying restrictions with the points (1;1), (0.5;10) and (10;0.5). In all the cases the test is not rejected (level 5%). To discriminate between them, the statistic  $\xi_{n,k}^{over}$  is a reasonable tool of comparison leading to the choice of the Hüsler-Reiss model.

Finally the probability that the claim amounts due to a storm (for the French insurer) are bigger than twice the largest losses from windstorm Lothar is 0.001439 (with a 95% confidence interval [0.0001086;0.0001791]), which corresponds to a storm every 104 years (resp. [84;138]). This return period confirms figures given by Swiss Re experts (Bresch, Bisping and Lemcke [2000]) and asks the recurrent question of a price structure for the storm risk at an inadequate level.

## REFERENCES

- [1] ABDOUS, B., GHOUDI, K. and KHOUDRAJI, A. (1998) Nonparametric estimation of the limit dependence function of multivariate extremes. *Extremes*, **2**: 243-265.
- [2] BEIRLANT, J., M., TEUGELS, J.L., and VYNCKIER, P. (1996) *Practical Analysis of Extreme Values*. Leuven University Press.
- [3] BEIRLANT, J., BRONIATOWSKI, M., TEUGELS, J.L., and VYNCKIER, P. (1995) The mean residual life function at great age: Applications to tails estimation. *J. Stat. Plann. Inf.*, **45**: 21-48.
- [4] COLES, S.G. and TAWN, J.A. (1991) Modelling extreme multivariate events. *J. R. Statist. Soc. B*, **53**: 377-392.
- [5] COLES, S.G. and TAWN, J.A. (1991) Statistical methods for multivariate extremes: an application to structural design. *Appl. Statist.*, **43**: 1-48.
- [6] CSÖRGO, S. and WELSH, A.H. (1989) Testing for exponential and Marshall-Olkin distributions. *J. Statist. Plan. Inference*, **23**: 287-300.
- [7] DEHEUVELS, P., HAEUSLER, E. de and MASSON, M.D. (1988) Almost sure convergence of the Hill estimator. *Math. Camb. Phil. Soc.*, **104**: 371-381.
- [8] EINMAHL, J.H.J., HANN, L. de and HUANG, X. (1993) Estimating a multidimensional extreme value distribution. *J. Mult. Analysis*, **47**: 35-47.
- [9] EINMAHL, J.H.J., HANN, L. de and PITERBARG, V.I. (1998) Nonparametric estimation of the spectral measure of an extreme value distribution. *Annals of Statistics*, **29**: 1401-1423.
- [10] EINMAHL, J.H.J., HANN, L. de and SINHA, A.K. (1997) Estimating the spectral measure of an extreme value distribution. *Stoch. Proc. Models*, **70**: 143-171.
- [11] EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- [12] FISHER, R.A. and TIPPET, L.H.C. (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Phil. Soc.*, **24**: 180-190.
- [13] GALAMBOS, J. (1987) *The Asymptotic Theory of Extreme Order Statistic*. 2nd ed. Melbourne: Krieger.
- [14] GEOFFROY, J. (1958) Contribution à la théorie des valeurs extrêmes. *Publication de l'ISUP*, **7**: 37-121.
- [15] HAAN, L. de and RESNICK, S. (1977) Limit theory for multivariate samples extremes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **40**: 317-337.
- [16] HAAN, L. de and RESNICK, S. (1993) Estimating the limiting distribution of multivariate extremes. *Comm. Statist. Stoch. Models*, **9**: 275-309.

- [17] HAAN, L. de and RESNICK, S. (1996) On asymptotic normality of the Hill estimator. *Stochastic Models*, **14**: 849-867
- [18] HAAN, L. de and RONDE, J. (1998) Sea and Wind: multivariate extremes at work. *Extremes*, **1**: 7-45.
- [19] HAAN, L. de and SINHA, A.K. (1999) Estimating the probability of a rare event. *Annals of Statistics*, **27**: 739-759.
- [20] HAAN, L. de and STADTMÜLLER, U. (1996) Generalized regular variation of second order. *J. Australian Math. Soc. A*, **61**: 381-395.
- [21] JOE, H. (1993). Parametric Families of multivariate distributions with given margin. *J. Mult. Analysis*, **46**: 262-282.
- [22] JOE, H. (1997) *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability 73, Chapman & Hall.
- [23] JOE, H., SMITH, R.L. and WEISSMAN, I. (1992) Bivariate threshold methods for extremes. *J. R. Statist. Soc. B*, **59**: 171-183.
- [24] KLÜPPELBERG, C., and MAY, A. (1998) The dependence function for bivariate extreme value distributions. *Working paper*, Muenchen University.
- [25] LEDFORD, A.W. and TAWN, J.A. (1996) Statistics for near independence in multivariate extreme values. *Biometrika*, **83**: 169-187.
- [26] LEDFORD, A.W. and TAWN, J.A. (1997). Modelling Dependence within joint tail regions. *J. R. Statist. Soc. B*, **59**: 475-499.
- [27] RESNICK, S. (1987) *Extreme Values, Regular Variation, and Point Processes*. Springer, New-York.
- [28] RESNICK, S. and STARICA, C. (1999) Smoothing the moment estimator of the extreme value parameter. *Extremes*, **1**: 263-293.
- [29] RESNICK, S. (1986) Point processes, regular variation and weak convergence. *Adv. Appl. Probab.*, **18**: 66-138.
- [30] SIBUYA, M. (1960) Bivariate extreme statistics. I, *Annals of the Institute of Statistical Mathematics, Tokyo*, **19**: 195–210.
- [31] TAWN, J.A. (1988) Bivariate extreme value theory: Models and estimation. *Biometrika*, **75**: 397–415.
- [32] TAWN, J.A. (1988) Modelling multivariate extreme value distributions. *Biometrika*, **77**: 245–253.

## Appendix 1

We first introduce some sets and definitions. Let  $E := [0, \infty]^2 \setminus \{0\}$  be the one point uncompactification of  $[0, \infty] \times [0, \infty]$  so that the compact sets of  $E$  are of the form  $U^c$ , where  $U$  is an open set of  $[0, \infty) \times [0, \infty)$  and  $0 \in U$ . Let  $C^+(E)$  be the space of continuous, non-negative functions on  $E$  with compact support. The vague topology on  $M_+(E)$ , the space of positive radon measures on  $E$ , can be generated by a countable family of semi-norms:

$$\{p_f : M_+(E) \rightarrow \mathbb{R}^+ : p_f(\mu) = \mu(f), |f| \leq 1, f \in C^+(E)\}$$

turning  $M_+(E)$  into a complete, separable, metric space. Convergence of  $\mu_n \in M_+(E)$  to  $\mu_0 \in M_+(E)$  in the vague topology is then denoted by  $\mu_n \rightarrow_v \mu_0$ .

**PROOF OF PROPOSITION 3.1.:**

To prove the probability consistency in Proposition 3.1, we first define the estimators:

$$\begin{aligned}\tilde{\nu}_n(([0, x] \times [0, y])^c) &:= \frac{1}{k} \sum_{i=1}^n \mathcal{E} \frac{X_i - U_X(n/k)}{U_X(n/k) / \log(n/k)} \frac{Y_i}{U_Y(n/k)} ([0, x] \times [0, y])^c, \\ \hat{\nu}_n(([0, x] \times [0, y])^c) &:= \frac{1}{k} \sum_{i=1}^n \mathcal{E} \frac{X_i - \hat{U}_X(n/k)}{\hat{U}_X(n/k) / \log(n/k)} \frac{Y_i}{\hat{U}_Y(n/k)} ([0, x] \times [0, y])^c,\end{aligned}$$

and establish some intermediary Propositions.

**PROPOSITION A.1.**

Suppose that (2.1) holds. Let  $k(n)$  be a sequence such that:  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$ , then we have for each  $x > 0$  and  $y > 0$ :

$$\tilde{\nu}_n(([0, x] \times [0, y])^c) \rightarrow_p \nu(([0, x] \times [0, y])^c). \quad (\text{A.1})$$

**PROOF OF PROPOSITION A.1.:**

(2.1) is equivalent to vague convergence of measures on  $E$ :

$$\frac{n}{k} P \left( \left\{ \frac{X_i - U_X(n/k)}{U_X(n/k) / \log(n/k)} > x \right\} \cup \left\{ \frac{Y_i}{U_Y(n/k)} > y \right\} \right) \rightarrow_v \nu(([0, x] \times [0, y])^c).$$

In the vague topology on  $M_+(E)$ , we have:

$$\tilde{\nu}_n(([0, x] \times [0, y])^c) \rightarrow_p \nu(([0, x] \times [0, y])^c). \quad (\text{A.2})$$

by Proposition 5.3 of Resnick [1986].

Now we replace the unknown quantities by sample versions.

**PROPOSITION A.2.**

Suppose that (2.1) holds. Let  $k$  be a sequence such that:  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ , as  $n \rightarrow \infty$ , then we have for each  $x > 0$  and  $y > 0$ :

$$\hat{\nu}_n(([0, x] \times [0, y])^c) \rightarrow_p \nu(([0, x] \times [0, y])^c).$$

**PROOF OF PROPOSITION A.2.:**

We replace the unknown quantities  $U_X(n/k)$  and  $U_Y(n/k)$  by sample versions and use a simple continuity argument. Let us remark that:

$$\hat{\nu}_n(([0, x] \times [0, y])^c) = \tilde{\nu}_n \left( \left( \left[ 0, x \frac{\hat{U}_X(n/k)}{U_X(n/k)} + \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k) / (\tau \log(n/k))} \right] \times \left[ 0, y \frac{\hat{U}_Y(n/k)}{U_Y(n/k)} \right] \right)^c \right).$$

Define a continuous map  $T : M_+(E) \times (\mathbb{R}^+)^3 \rightarrow M_+(E)$ :

$$T(\nu, a, b, c) := \nu \left( ([0, xa + (a-1)/c] \times [0, yb])^c \right).$$

From Proposition (A.1) and classical arguments:

$$\left( \tilde{\nu}_n, \frac{\hat{U}_X(n/k)}{U_X(n/k)}, \frac{\hat{U}_Y(n/k)}{U_Y(n/k)}, \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right) \rightarrow_P (\nu, 1, 1, 0)$$

in  $M_+(E) \times (\mathbb{R}^+)^3$  Apply the map  $T$  and the continuous mapping theorem to obtain the result.

The proof of the probability consistency of both measures in Proposition 3.1 is similar to the previous proof by considering also that:

$$\frac{1}{\hat{\tau}_k} P \rightarrow \frac{1}{\tau}, \frac{1}{\hat{\alpha}_k} P \rightarrow \frac{1}{\alpha} \text{ and}$$

$$\hat{\mu}_n^*((x, \infty] \times (y, \infty]) = \hat{\nu}_n^*(([0, x] \times [0, \infty])^c) + \hat{\nu}_n^*(([0, \infty] \times [0, y])^c) - \hat{\nu}_n^*(([0, x] \times [0, y])^c).$$

The almost sure consistency of  $\tilde{\nu}_n$  is obtained with the same arguments as de Hann and Resnick [1993]. Moreover Beirlant et al. [1995] and Deheuvels et al. [1988] show with similar assumptions that:

$$\frac{1}{\hat{\tau}_k} \rightarrow_{p.s.} \frac{1}{\tau}, \text{ and } \frac{1}{\hat{\alpha}_k} \rightarrow_{p.s.} \frac{1}{\alpha}.$$

The classical convergence:

$$\left( \frac{\hat{U}_X(n/k)}{U_X(n/k)}, \frac{\hat{U}_Y(n/k)}{U_Y(n/k)} \right) \rightarrow_{p.s.} (1, 1)$$

and the mapping theorem give the final result.

**PROOF OF PROPOSITION 3.2.:**

To prove Proposition 3.2, we begin by discussing the limiting properties of:

$$\begin{aligned} \tilde{\nu}_n^*(([0, x] \times [0, y])^c) &:= \frac{1}{k} \sum_{i=1}^n \mathcal{E} \exp\left(\tau \frac{X_i - U_X(n/k)}{U_X(n/k)/\log(n/k)}\right) \left(\frac{Y_i}{U_Y(n/k)}\right)^\alpha ([0, x] \times [0, y])^c, \\ \tilde{\mu}_n^*((x, \infty] \times (y, \infty]) &:= \frac{1}{k} \sum_{i=1}^n \mathcal{E} \exp\left(\tau \frac{X_i - U_X(n/k)}{U_X(n/k)/\log(n/k)}\right) \left(\frac{Y_i}{U_X(n/k)}\right)^\alpha (x, \infty] \times (y, \infty]. \end{aligned}$$

**PROPOSITION A.3.**

Suppose (2.1) holds and for  $\{x, y\} \in E$  define:

$$W_n(\{x, y\}) = \sqrt{k} \left( \tilde{\nu}_n^*(([0, x] \times [0, y])^c) - E \tilde{\nu}_n^*(([0, x] \times [0, y])^c) \right),$$

There exists a zero mean Gaussian random field  $(W(\{x, y\}), \{x, y\} \in E)$  with covariance function:

$$\text{cov}\left(W(\{x, y\}), W(\{x', y'\})\right) = \nu_* \left( ([0, x] \times [0, y])^c \cap ([0, x'] \times [0, y']^c) \right),$$

such that in  $D((0, \infty] \times (0, \infty])$ :

$$W_n \Rightarrow W.$$

Moreover there exists a version of  $W$  which is sample path continuous.

**PROOF OF PROPOSITION A.3:**

The proof is standard and details can be read in de Hann and Resnick [1994], Section 3.

Moreover we assume that  $\chi_\theta$  (or equivalently  $\Phi_\theta$ ) has continuous first order partial derivatives.

**PROPOSITION A.4:**

Suppose that (2.1) holds and that  $k = k(n)$  is such that  $\lim_{n \rightarrow \infty} \sqrt{k} c(U_X(n/k), U_Y(n/k)) = 0$ , then in  $D((0, \infty] \times (0, \infty]) \times \mathbb{R}^4$ :

$$\begin{aligned} & \sqrt{k} \left( \begin{array}{c} \left( \tilde{\mu}_n^*((x, \infty] \times (y, \infty]) - \frac{1}{y} \chi_\theta \left( \frac{y}{x} \right) \right), \frac{1}{\hat{\tau}} - \frac{1}{\tau}, \\ \frac{1}{\hat{\alpha}} - \frac{1}{\alpha}, \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))}, \frac{\hat{U}_Y(n/k) - U_Y(n/k)}{U_Y(n/k)/\alpha} \end{array} \right) \\ & \Rightarrow (W(x, \infty) + W(\infty, y) - W(x, y), \Gamma_\tau, \Gamma_\alpha, W(1, \infty), W(\infty, 1)), \end{aligned}$$

where:

$$\Gamma_\tau = \int_0^\infty W(e^{\tau u}, \infty) du - \frac{1}{\tau} W(1, \infty) \quad \Gamma_\alpha = \int_1^\infty W(\infty, v^\alpha) \frac{dv}{v} - \frac{1}{\alpha} W(\infty, 1).$$

**PROOF OF PROPOSITION A.4.:**

The proof is presented in a series of steps.

**STEP 1:**

As previously,

$$\tilde{\mu}_n^*((x, \infty] \times (y, \infty]) = \tilde{\nu}_n^*(([0, x] \times [0, \infty])^c) + \tilde{\nu}_n^*(([0, \infty] \times [0, y])^c) - \tilde{\nu}_n^*(([0, x] \times [0, y])^c),$$

This implies:

$$\begin{aligned} & \sqrt{k} \left( \tilde{\mu}_n^*((x, \infty] \times (y, \infty]) - \Phi_\theta(x, y) \right) \\ & = \sqrt{k} \left( \tilde{\nu}_n^*(([0, x] \times [0, \infty])^c) - E \tilde{\nu}_n^*(([0, x] \times [0, \infty])^c) \right) \\ & + \sqrt{k} \left( \frac{n}{k} (1 - F_X(a_{n/k} \log x + b_{n/k})) - \nu^*(([0, x] \times [0, \infty])^c) \right) \\ & + \sqrt{k} \left( \tilde{\nu}_n^*(([0, \infty] \times [0, y])^c) - E \tilde{\nu}_n^*(([0, \infty] \times [0, y])^c) \right) \\ & + \sqrt{k} \left( \frac{n}{k} (1 - F_Y(c_{n/k} y^{1/\alpha} + d_{n/k})) - \nu^*(([0, \infty] \times [0, y])^c) \right) \\ & - \sqrt{k} \left( \tilde{\nu}_n^*(([0, x] \times [0, y])^c) - E \tilde{\nu}_n^*(([0, x] \times [0, y])^c) \right) \\ & - \sqrt{k} \left( \frac{n}{k} (1 - F(a_{n/k} \log x + b_{n/k}, c_{n/k} y^{1/\alpha} + d_{n/k})) - \nu^*(([0, x] \times [0, y])^c) \right). \end{aligned}$$

Proposition 3.1 and Assumption (3.2) imply:

$$\sqrt{k} \left( \tilde{\mu}_n^*((x, \infty] \times (y, \infty]) - \Phi_\theta(x, y) \right) \Rightarrow W(x, \infty) + W(\infty, y) - W(x, y).$$

STEP 2:

We use Proposition 2.1 in Resnick and Starica [1999] to get:

$$\sqrt{k} \left( \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right) = \sqrt{k} \left( \frac{X_{(k)} - b_{n/k}}{a_{n/k}} \right) \rightarrow_d W(1, \infty),$$

and:

$$\frac{\hat{U}_Y(n/k) - U_Y(n/k)}{U_Y(n/k)/\alpha} \rightarrow_d W(\infty, 1).$$

STEP 3:

Proposition 2.2 in Resnick et Starica [1999] implies that:

$$\sqrt{k} \left( \frac{\tau \hat{U}_X(n/k)}{\hat{\tau}_k U_X(n/k)} - 1 \right) d \rightarrow \tau \int_0^\infty W(e^{\tau u}, \infty) du - W(1, \infty),$$

Using Step 2:

$$\sqrt{k} \left( \frac{\hat{U}_X(n/k)}{U_X(n/k)} - 1 \right) = \sqrt{k} \left( \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right) \frac{1}{(\tau \log(n/k))} \rightarrow_p 0.$$

and:

$$\sqrt{k} \left( \frac{1}{\hat{\tau}_k} - \frac{1}{\tau} \right) d \rightarrow \int_0^\infty W(e^{\tau u}, \infty) du - \frac{1}{\tau} W(1, \infty).$$

By using second order condition and results in de Hann and Resnick [1996], we get:

$$\sqrt{k} \left( \frac{1}{\hat{\alpha}_k} - \frac{1}{\alpha} \right) d \rightarrow \int_1^\infty W(\infty, v^\alpha) \frac{dv}{v} - \frac{1}{\alpha} W(\infty, 1).$$

Combination of Steps 1 to 3 gives the result.

We can now establish Proposition 3.3. First let us remark that:

$$\begin{aligned} & \hat{\mu}_n^*((x, \infty] \times (y, \infty]) \\ &= \tilde{\mu}_n^* \left( \left[ \exp \left\{ \frac{\tau}{\hat{\tau}} \log x \frac{\hat{U}_X(n/k)}{U_X(n/k)} + \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right\}, \infty \right] \times \left[ y^{\frac{\alpha}{\hat{\alpha}}} \left( \frac{\hat{U}_Y(n/k)}{U_Y(n/k)} \right)^\alpha, \infty \right] \right). \end{aligned}$$

Then:

$$\begin{aligned}
& \sqrt{k} \left( \hat{\mu}_n^*((x, \infty] \times (y, \infty]) - \Phi_\theta(x, y) \right) = I + II \\
& = \sqrt{k} \left( \hat{\mu}_n^* \left( \left[ \exp \left\{ \frac{\tau}{\hat{\tau}_k} \log x \frac{\hat{U}_X(n/k)}{U_X(n/k)} + \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right\}, \infty \right] \times \left[ y^{\frac{\alpha}{\hat{\alpha}_k}} \left( \frac{\hat{U}_Y(n/k)}{U_Y(n/k)} \right)^\alpha, \infty \right] \right) \right. \\
& \quad \left. - k \Phi_\theta \left( \exp \left\{ \frac{\tau}{\hat{\tau}_k} \log x \frac{\hat{U}_X(n/k)}{U_X(n/k)} + \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right\}, y^{\frac{\alpha}{\hat{\alpha}_k}} \left( \frac{\hat{U}_Y(n/k)}{U_Y(n/k)} \right)^\alpha \right) \right. \\
& \quad \left. + \sqrt{k} \left( \Phi_\theta \left( \exp \left\{ \frac{\tau}{\hat{\tau}_k} \log x \frac{\hat{U}_X(n/k)}{U_X(n/k)} + \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right\}, y^{\frac{\alpha}{\hat{\alpha}_k}} \left( \frac{\hat{U}_Y(n/k)}{U_Y(n/k)} \right)^\alpha \right) - \Phi_\theta(x, y) \right) \right)
\end{aligned}$$

Then let us look at the convergence of  $I$ . In this purpose, consider the continuous function  $T : M_+([0, \infty]^2 \setminus \{0\}) \times (\mathbb{R}^+)^5 \rightarrow M_+([0, \infty]^2 \setminus \{0\}) :$

$$T(v, a, b, c, d, e) := v \left( ([\exp\{ad \log x + c\}, \infty] \times [0, b^\alpha \exp\{e \log y\}, \infty]) \right).$$

which gives:

$$\hat{\mu}_n^*((x, \infty] \times (y, \infty]) = T \left( \hat{\mu}_n^*, \frac{\hat{U}_X(n/k)}{U_X(n/k)}, \frac{\hat{U}_Y(n/k)}{U_Y(n/k)}, \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))}, \frac{\tau}{\hat{\tau}_k}, \frac{\alpha}{\hat{\alpha}_k} \right)$$

Note that:

$$\left( \hat{\mu}_n^*, \frac{\hat{U}_X(n/k)}{U_X(n/k)}, \frac{\hat{U}_Y(n/k)}{U_Y(n/k)}, \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))}, \frac{\tau}{\hat{\tau}_k}, \frac{\alpha}{\hat{\alpha}_k} \right) \Rightarrow (\tilde{\mu}, 1, 1, 0, 1, 1)$$

in  $M_+([0, \infty]^2 \setminus \{0\}) \times (\mathbb{R}^+)^5$ . Apply the map  $T$  and the continuous mapping theorem to obtain:

$$I \Rightarrow W(x, \infty) + W(\infty, y) - W(x, y)$$

Here we have used the fact that since  $W$  has sample continuous, the convergence is the same as as weak convergence with respect to the topology of local uniform convergence.

We finish this proof with the convergence of  $II$ . Since  $\Phi_\theta$  has continuous first order partial derivatives, Corollary (3.1) implies that:

$$\begin{aligned}
& \sqrt{k} \left( \Phi_\theta \left( \exp \left\{ \frac{\tau}{\hat{\tau}} \ln x \frac{\hat{U}_X(n/k)}{U_X(n/k)} + \frac{\hat{U}_X(n/k) - U_X(n/k)}{U_X(n/k)/(\tau \log(n/k))} \right\}, y^{\frac{\alpha}{\hat{\alpha}}} \left( \frac{\hat{U}_Y(n/k)}{U_Y(n/k)} \right)^\alpha \right) - \Phi_\theta(x, y) \right) \\
& \Rightarrow \frac{\partial \Phi_\theta(x, y)}{\partial x} x (W(1, \infty) + \tau \log x \Gamma_\tau) + \frac{\partial \Phi_\theta(x, y)}{\partial y} y (W(\infty, 1) + \alpha \log y \Gamma_\alpha),
\end{aligned}$$

which gives final result.

**PROOF OF PROPOSITION 3.3:**

The proof is similar as Proposition 3.2.

PROOF OF COROLLARY 3.4:

A simple computation of the variance of:

$$W(x, y) - \frac{1}{x}(W(1, \infty) + \tau \log x \Gamma_\tau) - \frac{1}{y}(W(\infty, 1) + \alpha \log y \Gamma_\alpha)$$

gives the result.

## Appendix 2

PROOF OF PROPOSITION 4.1.:

To prove consistency, we make the following technical assumptions:

H0: the sequence  $k$  satisfies  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$ ,  $k/\log n \rightarrow \infty$ , as  $n \rightarrow \infty$ .

H1: the variables  $(X_i, Y_i)$ ,  $i = 1, \dots, n$ , are independent and identically distributed..

H2:  $\hat{\mu}_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty])$  converges almost surely to  $\Phi_{\theta_0}(\mathbf{x}, \mathbf{y})$ .

H3:  $S_n$  converges almost surely to a non-random matrix  $S_0$ .

H4: the parameter is identifiable from the constraints, i.e.:

$$\left[ \mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_\theta \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_0 \left[ \mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_\theta \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right] = 0 \Rightarrow \theta = \theta_0.$$

H5: The function  $\Phi_\theta(\mathbf{x}, \mathbf{y})$  is continuous in  $\theta$ .

H6:  $\theta_0$  is known to belong to a compact set  $A$ .

The function:

$$\theta \mapsto \left[ \hat{\mu}_{n_j}^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_\theta \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_n \left[ \hat{\mu}_{n_j}^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_\theta \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]$$

is continuous on the compact set  $A$ . It follows that the estimator  $\theta_n(S_n)$  exists because the associated optimisation problem has a solution.

Let us consider the sequence  $\theta_n(S_n)$ . Since it belongs to the compact set  $A$ , there exists a sub-sequence  $\theta_{n_j}(S_{n_j})$  converging to a value, which implies the existence of  $\theta_\infty(X, Y) \in A$ .

Since

$$\begin{aligned} & \left[ \hat{\mu}_{n_j}^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_{n_j}(S_{n_j})} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_{n_j} \left[ \hat{\mu}_{n_j}^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_{n_j}(S_{n_j})} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right] \\ & \leq \left[ \hat{\mu}_{n_j}^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_{n_j} \left[ \hat{\mu}_{n_j}^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right], \end{aligned}$$

because  $\theta_{n_j}(S_{n_j})$  is a solution to the minimization problem, taking the limit, it follows that:

$$\begin{aligned} & \left[ \mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_\infty} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_0 \left[ \mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_\infty} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right] \\ & \leq \left[ \mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]' S_0 \left[ \mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \frac{1}{\mathbf{y}} \chi_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right]. \end{aligned}$$

Because the right-hand side of the latter inequality is zero, it follows that the left-hand side must also be equal to zero. Then assumption H4 implies that  $\theta_\infty = \theta_0$ . In particular  $\theta_\infty(X, Y)$  is nonrandom. Finally it suffices to note that every converging subsequence of  $\theta_n(S_n)$  converges to  $\theta_\infty$  because  $A$  is compact.

**PROOF OF PROPOSITION 4.2:**

In addition to the previous conditions, it is assumed that:

H8:  $\theta_0$  belongs to the interior of  $A$ .

H9: the gradient  $\frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta}$  exists and is continuous in  $\theta_0$ .

H10:  $\lim_{n \rightarrow \infty} \sqrt{k} c(U_X(n/k), U_Y(n/k)) = 0$ .

We have:

$$\frac{\partial \Phi'_{\theta_n(S_n)}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_n \left[ \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \Phi_{\theta_n(S_n)}(\mathbf{x}, \mathbf{y}) \right] = 0.$$

We can take the Taylor expansion because  $\theta_n(S_n)$  converge to  $\theta_0$ , as  $n$  increases to infinity:

$$\Phi_{\theta_n(S_n)}(\mathbf{x}, \mathbf{y}) = \Phi_{\theta_0}(\mathbf{x}, \mathbf{y}) + \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} (\theta_n(S_n) - \theta_0) + o_p(\theta_n(S_n) - \theta_0).$$

We obtain:

$$\frac{\partial \Phi'_{\theta_n(S_n)}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_n \left[ \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \Phi_{\theta_0}(\mathbf{x}, \mathbf{y}) - \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} (\theta_n(S_n) - \theta_0) + o_p(\theta_n(S_n) - \theta_0) \right] = 0,$$

or equivalently:

$$\begin{aligned} \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_n \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} (\theta_n(S_n) - \theta_0) &= \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_n \left[ \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \Phi_{\theta_0}(\mathbf{x}, \mathbf{y}) \right] \\ &\quad + o_p(\theta_n(S_n) - \theta_0), \end{aligned}$$

Then it suffices to use the asymptotic normality of

$$\sqrt{k} \left( \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \frac{1}{\mathbf{y}} \mathcal{X}_{\theta_0} \left( \frac{\mathbf{y}}{\mathbf{x}} \right) \right) \rightarrow_d N(0, \Omega_{\theta_0}(\mathbf{x}, \mathbf{y})),$$

to establish that:

$$\sqrt{k} (\theta_n(S_n) - \theta_0) \rightarrow_d N(0, \Sigma(S_0, \mathbf{x}, \mathbf{y})),$$

with:

$$\begin{aligned} \Sigma(S_0, \mathbf{x}, \mathbf{y}) &= \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1} \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \Omega_{\theta_0}(\mathbf{x}, \mathbf{y}) S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right] \\ &\quad \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1}. \end{aligned}$$

PROOF OF PROPOSITION 4.3:

We need to prove that  $\Sigma(S_0^*, \mathbf{x}, \mathbf{y}) \circ \Sigma(S_0, \mathbf{x}, \mathbf{y})$  for every other symmetric positive matrix  $S_0$ , where the order relation  $\circ$  means that the symmetric matrix  $\Sigma(S_0^*, \mathbf{x}, \mathbf{y}) - \Sigma(S_0, \mathbf{x}, \mathbf{y})$  is negative. This condition can be written as:

$$\begin{aligned} & \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y}) \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1} \\ & \circ \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1} \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \Omega_{\theta_0}(\mathbf{x}, \mathbf{y}) S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right] \\ & \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1}. \end{aligned}$$

Let:

$$\begin{aligned} A &= \Omega_{\theta_0}^{-1/2}(\mathbf{x}, \mathbf{y}) \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta}, \\ B &= \Omega_{\theta_0}^{1/2}(\mathbf{x}, \mathbf{y}) S_0 \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta}. \end{aligned}$$

The above condition becomes:

$$(A'A)^{-1} \circ (A'B)^{-1}(B'B)(A'B)^{-1},$$

i.e.

$$A'A \pm A'B(B'B)^{-1}B'A,$$

i.e.

$$A'(Id - B(B'B)^{-1}B')A \pm 0.$$

The latter inequality is satisfied because  $Id - B(B'B)^{-1}B'$  is an orthogonal projection matrix.

PROOF OF PROPOSITION 4.4:

By using a Taylor expansion around  $\theta_0$ , we have:

$$\begin{aligned} \sqrt{k} \left( \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \Phi_{\theta_n^*}(\mathbf{x}, \mathbf{y}) \right) &= \sqrt{k} \left( \hat{\mu}_n^*((\mathbf{x}, \infty) \times (\mathbf{y}, \infty)) - \Phi_{\theta_0}(\mathbf{x}, \mathbf{y}) \right) \\ &\quad - \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \sqrt{k} (\theta_n^* - \theta_0) + o_p \left( \sqrt{k} (\theta_n^* - \theta_0) \right). \end{aligned}$$

On the other hand, the proof of Proposition 4.2. implies:

$$\begin{aligned}\sqrt{k}(\theta_n^* - \theta_0) &= \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y}) \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1} \\ &\quad \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y}) \sqrt{k} \left[ \hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \Phi_{\theta_0}(\mathbf{x}, \mathbf{y}) \right] + o_p \left( \sqrt{k}(\theta_n^* - \theta_0) \right),\end{aligned}$$

hence, we obtain:

$$\begin{aligned}&\sqrt{k} \left( \hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \Phi_{\theta_n^*}(\mathbf{x}, \mathbf{y}) \right) \\ &= \left( Id - \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y}) \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1} \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y}) \right) \\ &\quad \times \sqrt{k} \left[ \hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \Phi_{\theta_0}(\mathbf{x}, \mathbf{y}) \right] + o_p \left( \sqrt{k}(\theta_n^* - \theta_0) \right).\end{aligned}$$

But,

$$M_0 = \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \left[ \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y}) \frac{\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \right]^{-1} \frac{\partial \Phi'_{\theta_0}(\mathbf{x}, \mathbf{y})}{\partial \theta} \Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y})$$

is the orthogonal projection matrix onto the subspace generated by columns of  $\partial \Phi_{\theta_0}(\mathbf{x}, \mathbf{y}) / \partial \theta$ , with respect to the scalar product  $\Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y})$ .

Thus the asymptotic distribution  $\sqrt{k} \left( \hat{\mu}_n^*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \Phi_{\theta_n^*}(\mathbf{x}, \mathbf{y}) \right)$  is the singular multivariate normal distribution  $N(0, (Id - M_0) \Omega_{\theta_0}(\mathbf{x}, \mathbf{y}) (Id - M_0)')$ . Since  $\Omega_{\theta_0}^{-1}(\mathbf{x}, \mathbf{y})$  is the generalized inverse of  $(Id - M_0) \Omega_{\theta_0}(\mathbf{x}, \mathbf{y}) (Id - M_0)'$ , it follows that the asymptotic distribution of  $\xi_n$  is a  $\chi^2$  distribution with the number of degrees of freedoms equal to the rank of  $(Id - M_0) \Omega_{\theta_0}(\mathbf{x}, \mathbf{y}) (Id - M_0)'$ , which is  $H - K$ .

If there does not exist a value of the parameter  $\theta$ ,  $\theta_n^*$  converges to  $\theta_\infty$  such that  $\mu_*((\mathbf{x}, \infty] \times (\mathbf{y}, \infty]) - \Phi_{\theta_\infty}(\mathbf{x}, \mathbf{y})) \neq 0$ . It follows that  $\xi_n$  converges to  $+\infty$  and the test is convergent.