Estimation of conditional laws given an extreme component

Anne-Laure Fougères, Philippe Soulier

Estimation of conditional laws given an extreme component

Anne-Laure Fougères* Philippe Soulier†

March 7, 2009

Abstract

Let (X, Y) be a bivariate random vector. The estimation of a probability of the form $P(Y \leq y \mid X > t)$ is challenging when t is large, and a fruitful approach consists in studying, if it exists, the limiting conditional distribution of the random vector (X, Y), suitably normalized, given that X is large. There already exists a wide literature on bivariate models for which this limiting distribution exists. In this paper, a statistical analysis of this problem is done. Estimators of the limiting distribution (which is assumed to exist) and the normalizing functions are provided, as well as an estimator of the conditional quantile function when the conditioning event is extreme. Consistency of the estimators is proved and a functional central limit theorem for the estimator of the limiting distribution is obtained. The small sample behavior of the estimator of the conditional quantile function is illustrated through simulations.

1 Introduction

Let (X, Y) be a bivariate random vector for which the conditional distribution of Y given that $X > t$ is of interest, for values of t such that the conditioning event is a rare event. This happens for example when the possible contagion between two dependent market returns X and Y is investigated, see e.g. Bradley and Taqqu (2004) or Abdous et al. (2008). The estimation of a probability of the form $P(Y \leq y \mid X > t)$ starts to be challenging as soon as t is large, since the conditional empirical distribution becomes useless when no observations are available. A fruitful alternative approach consists in studying, if it exists, the limiting distribution of the

*Université Claude Bernard - Lyon 1
†Université Paris Ouest-Nanterre

random vector (X, Y) conditionally on X to be large. This corresponds to assuming that there exist functions m, a and ψ, and a bivariate distribution function (cdf) F such that

$$\lim_{t\to\infty} \mathbb{P}(X \leq t + \psi(t)x \,; Y \leq m(t) + a(t)y \mid X > t) = F(x, y) \,. \qquad (1)$$

This approach was suggested by Heffernan and Tawn (2004) and investigated by Heffernan and Resnick (2007). Models for which condition (1) holds have already been investigated in many references. Eddy and Gale (1981) and Berman (1992) proved that (1) holds for spherical distributions; bivariate elliptic distributions were investigated by Abdous et al. (2005) multivariate elliptic distributions and related distributions by Hashorva (2006); Hashorva et al. (2007). The analysis of the underlying geometric structure (ellipticity of the level sets of the densities) has lead to various generalizations by Barbe (2003) and Balkema and Embrechts (2007). See also Fougères and Soulier (2008) for a recent review on the subject.

An important issue that still has to be addressed is the statistical estimation of the functions a and m that appear in (1), as well as the limiting distribution function F. This is the aim of the present paper. Two problems are considered. The first one is the nonparametric estimation of the limiting distribution and of the normalizing functions. This allows for instance to test for a specific limiting distribution, e.g. the standard Gaussian distribution as in many examples. Since we are also interested in the case where the conditionning event is beyond the range of observations, a semiparametric procedure will be defined to allow this extrapolation. This can only be done under more restrictive assumptions, which are satisfied by most models already investigated in the literature.

The paper is organized as follows. In Section 2, we rephrase (1) in terms of vague convergence of measures in order to use the point process techniques and the results of Heffernan and Resnick (2007). We also introduce moment assumptions which are needed to prove the consistency of the non parametric estimators introduced in Section 3. A functional central limit theorem is obtained under a second order condition, and its use in a goodness of fit test is discussed. A specific analysis of the case of a limiting distribution with product form is done in Section 4. Semi-parametric estimators that allow extrapolations beyond the range of the observations are studied and applied to the estimation of conditional quantiles when the conditioning event is extreme. A simulation study is given in Section 5, which illustrates the behavior of the goodness of fit test proposed in Section 3 and of the estimator of the conditional quantile proposed in Section 4. Section 6 finally collects some proofs.

2 Assumptions and preliminary results

We first rephrase the convergence (1) in terms of vague convergence of measures, in order to use point process techniques and the results of Heffernan and Resnick (2007).

Assumption 1. *There exists monotone functions a, b, m and ψ such that the measure ν_n defined by*

$$\nu_n(\cdot) = n\mathbb{P}\left(\left\{\frac{X - b(n)}{\psi \circ b(n)}, \frac{Y - m \circ b(n)}{a \circ b(n)}\right\} \in \cdot\right) \tag{2}$$

converges vaguely to a measure ν on $(-\infty, \infty] \times [-\infty, \infty]$ such that $\nu([0, \infty) \times (-\infty, \infty)) = 1$.

Define the probability distribution functions F_n and F by

$$F_n(x, y) = \nu_n([0, x] \times (-\infty, y]) \ , \quad F(x, y) = \nu([0, x] \times (-\infty, y]) \ .$$

Assumption 1 implies the convergence of F_n to F at all points of continuity of F.

Remark 1. Assumption 1 implies that the marginal distribution of X, say F_X belongs to a max-domain of attraction. Thus the functions ψ and b are determined by the marginal distribution of X only. The function b can and will be chosen as $b = (1/(1 - F_X))^{\leftarrow}$ where F_X is the distribution function of X. The function ψ satisfies

$$\lim_{x \to x_1} \frac{\psi(x + \psi(x)u)}{\psi(x)} = 1 + \gamma u \ , \tag{3}$$

where x_1 is the right endpoint of the support of the distribution of X and γ is the extreme value index of X. In the sequel, we will assume that $x_1 = +\infty$. The first marginal of ν is the generalized Pareto distribution G_γ, defined if $\gamma \neq 0$ by $G_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}$ for x such that $1 + \gamma x > 0$ and by $G_0(t) = 1 - \mathrm{e}^{-t}$ for $t > 0$ otherwise. Thus, all results concerning only the marginal distribution of X are obtained by applying the usual extreme value theory.

Remark 2. Assumption 1 has little implications on the functions a and m and the second marginal of the measure ν. Let Ψ denote this distribution, defined by

$$\Psi(z) = \int_0^\infty \int_{-\infty}^z \nu(\mathrm{d}x, \mathrm{d}y) \ .$$

If Y is independent of X, then Ψ is the distribution of Y, $a \equiv 1$ and $m \equiv 0$. Thus Ψ can be *any* probability distribution. In particular, it is not necessarily an extreme-value distribution.

Remark 3. If the pair (X, Y) satisfies Assumption 1, then so does any affine transformation of (X, Y). For instance, if X and Y have finite mean and variance, then $((X - \mathbb{E}[X])/\mathrm{var}^{1/2}(X), (Y - \mathbb{E}[Y])/\mathrm{var}^{1/2}(Y))$ also satisfies Assumption 1. But non linear transformations of (X, Y) do not necessarily satisfy the assumption. In particular, the usual (in extreme value theory) transformation of X and Y to random variables with prescribed marginal distributions, is not always possible, as investigated in (Heffernan and Resnick, 2007, Section 7). It is never possible in the cases where the joint limiting distribution is a product measure. Consequently, we do not make any specific assumption on the marginal distributions of X and Y.

Obviously, the functions a and m are defined up to asymptotic equivalence, i.e. if m' and a' satisfy

$$\lim_{x\to\infty} \frac{a'(x)}{a(x)} = 1 \ , \quad \lim_{x\to\infty} \frac{m(x) - m'(x)}{a(x)} = 0 \ ,$$

then the measure ν_n' defined as ν_n but with a' and m' instead of a and m converges vaguely to the same limit measure ν. Beyond this trivial remark, the following result summarizes (Heffernan and Resnick, 2007, Proposition 1 and 2) and contains most of what can be infered from Assumption 1.

Proposition 1. *Under Assumption 1, there exists $\zeta \in \mathbb{R}$ such that the function $a \circ b$ is regularly varying at infinity with index ζ and the function m satisfies*

$$\lim_{t\to\infty} \frac{m \circ b(tx) - m \circ b(t)}{a \circ b(t)} = J_\zeta(x) \ ,$$

with $J_\zeta(x) = (x^\zeta - 1)/\zeta$ if $\zeta \neq 0$ and $J_0(x) = c \log(x)$ for some $c \in \mathbb{R}$, and the convergence is locally uniform on $(0, \infty)$.

Assumption 1 alone is not sufficient for statistical purposes (which typically require bounded convergence arguments), since it has no implication on the joint moments of X of Y. Therefore the following assumption is introduced.

Assumption 2. *There exists $p^* > 0$, $q^* > 0$ such that*

$$\lim_{n\to\infty} \int_0^\infty \int_{-\infty}^\infty x^{p^*} |y|^{q^*} \nu_n(\mathrm{d}x, \mathrm{d}y) = \int_0^\infty \int_{-\infty}^\infty x^{p^*} |y|^{q^*} \nu(\mathrm{d}x, \mathrm{d}y) \ . \qquad (4)$$

For the reason mentioned in Remark 1, Assumption 1 implies the convergence (4) with $q^* = 0$ and $p^* < 1/\gamma$. In applications, it will be assumed that $q^* \geq 2$. The function a and the limiting measure ν are defined up to a change of scale, thus, without loss of generality, we assume henceforth that

$$\int_0^\infty \int_{-\infty}^\infty y^2 \, \nu(\mathrm{d}x, \mathrm{d}y) = \int_{-\infty}^\infty y^2 \Psi(\mathrm{d}y) = 1 \ .$$

Condition (4) can be equivalently written as the convergence of conditional moments. For all $q \leq p^*$ and $p \leq q^*$, it holds that

$$\lim_{t\to\infty} \frac{\mathbb{E}[(X-t)^q |Y - m(t)|^p \mid X > t]}{\psi^q(t) a^p(t)} = \int_0^\infty \int_{-\infty}^\infty x^q |y|^p \nu(\mathrm{d}x, \mathrm{d}y) \,. \tag{5}$$

More generally, for any measurable function g such that $|g(x,y)| \leq C(x \vee 1)^{p^*}(|y| \vee 1)^{q^*}$, it holds that

$$\lim_{t\to\infty} \mathbb{E}[g\{(X-t)/\psi(t), (Y-m(t))/a(t)\} \mid X > t] = \int_0^\infty \int_{-\infty}^\infty g(x,y)\nu(\mathrm{d}x, \mathrm{d}y) \,. \tag{6}$$

This can also be stated equivalently with the measure ν_n defined in Assumption 1. For all measurable function g such that $|g(x,y)| \leq C(x \vee 1)^{p^*}(|y| \vee 1)^{q^*}$, it holds that

$$\lim_{n\to\infty} \int_0^\infty \int_{-\infty}^\infty g(x,y)\nu_n(\mathrm{d}x, \mathrm{d}y) = \int_0^\infty \int_{-\infty}^\infty g(x,y)\nu(\mathrm{d}x, \mathrm{d}y) \,.$$

The following result will be needed for statistical purposes, and illustrates the use of the moment Assumption 2. For a sequence (X_i, Y_i), $1 \leq i \leq n$, let $X_{(n:i)}$ denote the i-th order statistic and $Y_{[n:i]}$ denote its concomitant.

Recall that an intermediate sequence is a sequence of integers k_n such that $\lim_{n\to\infty} k_n = \lim_{n\to\infty} n/k_n = \infty$. In accordance with common use and for the clarity of notation, the dependence on n will be implicit in the sequel.

Define the random measure

$$\tilde{\nu}_n = \frac{1}{k} \sum_{i=1}^n \delta_{(\{X_i - b(n/k)\}/\psi\circ b(n/k), \{Y_i - m\circ b(n/k)\}/a\circ b(n/k))} \,. \tag{7}$$

Proposition 2. *If Assumption 1 and 2 hold, then for any intermediate sequence k and any measurable function g such that $|g(x,y)| \leq C(x\vee 1)^{p^*}(|y|\vee 1)^{q^*}$,*

$$\int_0^\infty \int_{-\infty}^\infty g(x,y)\tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) \to_P \int_0^\infty \int_{-\infty}^\infty g(x,y)\nu(\mathrm{d}x, \mathrm{d}y) \,. \tag{8}$$

The proof is in section 7. For historical interest, we can also mention the following consequence of Assumption 1, first stated in (Eddy and Gale, 1981, Theorem 6.1) in a restricted case of spherical distributions.

Proposition 3. *Under Assumption 1, $\{Y_{[n:n]} - m \circ b(n/k)\}/a \circ b(n/k)$ converges weakly to Ψ. If moreover ν is a product measure, then $\{Y_{[n:n]} - m \circ b(n/k)\}/a \circ b(n/k)$ is asymptotically independent of $X_{(n:n)}$.*

Let us finally mention that Davydov and Egorov (2000) obtained functional limit theorems for sums of concomitants corresponding to a number k of order statistics such that $k/n \nrightarrow 0$. Their problem differs from ours. Their assumptions on the joint distribution of the random pairs are much weaker than Assumption 1, but their results are of a very different nature and it does not seem possible to use them to derive Proposition 2 for instance.

3 Nonparametric estimation of ψ, a, m and F

In this section, we introduce nonparametric estimators of the functions ψ, m, a and F based on i.i.d. observations $(X_1, Y_1), \ldots, (X_n, Y_n)$ of a bivariate distribution which satisfy Assumption 2.

3.1 Definitions and consistency

In order to estimate nonparametrically the limiting distribution F, we first need nonparametric estimators of the quantities $\psi(X_{(n:n-k)})$, $m(X_{(n:n-k)})$ and $a(X_{(n:n-k)})$, with k an intermediate sequence, i.e. such that $k \to \infty$ and $k/n \to 0$. The estimation of $\psi(X_{(n:n-k)})$ is a well known estimation issue, see e.g. Beirlant et al. (2004) or De Haan and Ferreira (2006) If the extreme value index γ of X is less than 1, then ψ can be estimated as the mean residual life. Let $\hat{\gamma}$ be a consistent estimator of γ (see e.g. Beirlant et al. (2004)) and define

$$\hat{\psi}(X_{(n:n-k)}) = \frac{1-\hat{\gamma}}{k} \sum_{i=1}^{k} \{X_{(n:n-i+1)} - X_{(n:n-k)}\} \, . \tag{9}$$

It follows straightforwardly from Proposition 2 that $\hat{\psi}(X_{(n:n-k)})/\psi \circ b(n/k)$ converges weakly to $(1-\gamma)\int_0^\infty x G_\gamma(\mathrm{d}x) = 1$. If it is moreover assumed (as in Section 4 below) that $\gamma = 0$, then the estimator above can be modified accordingly:

$$\hat{\psi}(X_{(n:n-k)}) = \frac{1}{k} \sum_{i=1}^{k} \{X_{(n:n-i+1)} - X_{(n:n-k)}\} \, . \tag{10}$$

In order to estimate m, define

$$\hat{m}(X_{(n:n-k)}) = \frac{\sum_{i=1}^{k} Y_{[n:n-i+1]} \{X_{(n:n-i+1)} - X_{(n:n-k)}\}}{\sum_{i=1}^{k} \{X_{(n:n-i+1)} - X_{(n:n-k)}\}} \, . \tag{11}$$

Proposition 4. *If Assumption 1 holds and Assumption 2 holds with $p^* \geq 1$ and $q^* \geq 1$, then, for any intermediate sequence k, it holds that*

$$\frac{\hat{m}(X_{(n:n-k)}) - m \circ b(n/k)}{a \circ b(n/k)} \to_P \mu \, ,$$

where $\mu = (1-\gamma)\int_0^\infty \int_{-\infty}^\infty xy\nu(\mathrm{d}x, \mathrm{d}y)$. *If moreover* $m(x) = \rho x$ *and if* $\mu = 0$ *or* $a(x) = o(x)$ *then* $\hat{m}(X_{(n:n-k)})/X_{(n:n-k)}$ *is a consistent estimator of* ρ.

Proof. Write

$$\frac{\hat{m}(X_{(n:n-k)}) - m \circ b(n/k)}{a \circ b(n/k)} = \frac{S_n}{T_n} ,$$

with

$$S_n = \frac{1}{k}\sum_{i=1}^{k} \frac{Y_{[n:n-i+1]} - m \circ b(n/k)}{a \circ b(n/k)} \, \frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi \circ b(n/k)} ,$$

$$T_n = \frac{1}{k}\sum_{i=1}^{k} \frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi \circ b(n/k)} .$$

We have already seen that T_n converges weakly to $1/(1-\gamma)$. Denote

$$\tilde{x}_n = \frac{X_{(n:n-k)} - b(n/k)}{\psi \circ b(n/k)} . \tag{12}$$

Under Assumption 1 (which implies that X belongs to a max-domain of attraction), it is well known that $\tilde{x}_n = o_P(1)$. Cf. De Haan and Ferreira (2006, Theorem 2.2.1). Write now (with notation $x_+ = \sup(x, 0)$)

$$\begin{aligned} S_n &= \frac{1}{k}\sum_{i=1}^{n} \frac{Y_i - m \circ b(n/k)}{a(X_{(n:n-k)})} \left\{ \frac{X_i - b(n/k)}{\psi \circ b(n/k)} - \tilde{x}_n \right\}_+ \\ &= \int_{\tilde{x}_n}^{\infty} \int_{-\infty}^{\infty} (x - \tilde{x}_n) y \, \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) \\ &= \int_0^{\infty} \int_{-\infty}^{\infty} xy \, \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) + o_P(1) . \end{aligned}$$

Thus S_n converges weakly to $\mu/(1-\gamma)$ by Proposition 2. □

Remark 4. A sufficient condition for $\mu = 0$ is the symmetry of the measure ν with respect to the second variable. This happens in particular if ν is a product measure, and the distribution Ψ is symmetric.

We now estimate $a(X_{(n:n-k)})$. Many estimators can be defined, each needing an ad hoc moment assumption. The one we have chosen needs $q^* \geq 2$ in Assumption 2. Define

$$\hat{a}(X_{(n:n-k)}) = \left\{ \frac{1}{k}\sum_{i=1}^{k} \{Y_{[n:n-i+1]} - \hat{m}(X_{(n:n-k)})\}^2 \right\}^{1/2} . \tag{13}$$

Proposition 5. *If Assumption 1 holds and Assumption 2 holds with $p^* \geq 1$ and $q^* \geq 2$, and if $\mu = 0$, then, for any intermediate sequence k, it holds that*

$$\hat{a}(X_{(n:n-k)})/a \circ b(n/k) \to_P 1 \ .$$

Proof. We show that $\hat{a}^2(X_{(n:n-k)})/a^2 \circ b(n/k)$ converges weakly to 1. Denote

$$\xi_n = \{\hat{m}(X_{(n:n-k)}) - m \circ b(n/k)\}/a \circ b(n/k) \ . \tag{14}$$

By Proposition 4 $\xi_n = o_P(1)$, and noting that $\tilde{\nu}_n\{[\tilde{x}_n, \infty] \times [-\infty, \infty]\} = 1$, where $\tilde{\nu}_n$ and $\tilde{x}_n$ are respectively defined by (7) and (12), we have

$$\begin{aligned}
&\frac{\hat{a}^2(X_{(n:n-k)})}{a^2 \circ b(n/k)} \\
&\qquad = \frac{1}{k}\sum_{i=1}^{n}\left\{\frac{Y_i - m \circ b(n/k)}{a \circ b(n/k)} - \xi_n\right\}^2 \mathbb{1}_{\{\frac{X_i - b(n/k)}{\psi \circ b(n/k)} \geq \tilde{x}_n\}} \\
&\qquad = \int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty}(y - \xi_n)^2 \, \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) \\
&\qquad = \int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty} y^2 \, \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) - 2\xi_n \int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty} y \, \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) + \xi_n^2 \qquad (15) \\
&\qquad = \int_{0}^{\infty}\int_{-\infty}^{\infty} y^2 \, \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) + o_P(1) \ .
\end{aligned}$$

Thus $\hat{a}(X_{(n:n-k)})/a \circ b(n/k)$ converges weakly to 1 by Proposition 2 and the chosen normalization $\int_{-\infty}^{\infty} y^2 \Psi(\mathrm{d}y) = 1$. □

Remark 5. If $\mu \neq 0$, then $\xi_n \to_P \mu$ and $\hat{a}(X_{(n:n-k)})/a \circ b(n/k) \to_P \tau$, with

$$\tau^2 = 1 - 2\mu \int_{-\infty}^{\infty} y \Psi(\mathrm{d}y) + \mu^2 \ . \tag{16}$$

We can now consider the nonparametric estimator of the limiting joint distribution F. Define

$$\begin{aligned}
\hat{F}(x, y) = \frac{1}{k}\sum_{i=1}^{k} &\mathbb{1}_{\{X_{(n:n-i+1)} \leq X_{(n:n-k)} + \hat{\psi}(X_{(n:n-k)})x\}} \\
&\times \mathbb{1}_{\{Y_{[n:n-i+1]} \leq \hat{m}(X_{(n:n-k)}) + \hat{a}(X_{(n:n-k)})y\}} \ . \qquad (17)
\end{aligned}$$

Theorem 6. *Under Assumptions 1 and 2 with $\gamma < 1$, $p^* \geq 1$ and $q^* \geq 2$, if $\mu = 0$, and if F is continuous, then, for any intermediate sequence k, $\hat{F}(x, y)$ converges weakly to $F(x, y)$.*

Proof. Denote $u_n = \hat{\psi}(X_{(n:n-k)})/\psi \circ b(n/k)$ and

$$v_n = \hat{a}(X_{(n:n-k)})/a \circ b(n/k) \ . \tag{18}$$

Then

$$\hat{F}(x,y) = \tilde{\nu}_n([\tilde{x}_n, \tilde{x}_n + u_n x] \times (-\infty, \xi_n + v_n y]) \ ,$$

where ξ_n is defined by (14). Thus $\hat{F}(x,y)$ converges weakly to $F(x,y)$ by Proposition 2, 4 and 5. □

We can also define an estimator of the second marginal Ψ of F. Denote

$$\hat{\Psi}(y) = \frac{1}{k} \sum_{i=1}^{k} \mathbb{1}_{\{Y_{[n:n-i+1]} \leq \hat{m}(X_{(n:n-k)}) + \hat{a}(X_{(n:n-k)})y\}} \ . \tag{19}$$

Then, under the assumptions of Theorem 6, $\hat{\Psi}_n$ also converges to Ψ. Note that if $\mu \neq 0$, then $\hat{\Psi}(z)$ converges weakly to $\Psi(\mu + \tau z)$, with τ defined in (16).

3.2 Central limit theorems

In order to obtain central limit theorems, we need to strengthen Assumptions 1 and 2.

Assumption 3. *There exist positive real numbers p^* and q^*, a function c and a locally bounded function χ on $\mathbb{R}^2$ such that*

$$|\mathbb{P}(X \leq t + \psi(t)x \ , \ Y \leq m(t) + a(t)y \mid X > t) - F(x,y)| \leq c(t)\chi(x,y) \ , \tag{20}$$

$$\int_0^\infty \int_{-\infty}^\infty x^{p^*+1}|y|^{q^*+1}\chi(x,y)\,\mathrm{d}x\,\mathrm{d}y < \infty \ . \tag{21}$$

Remark 6. Condition (20) is similar to de Haan and Resnick (1993, Condition 4.1), but with a bound instead of a limit. It gives a rate of uniform convergence on compact sets. Condition (21) yields the convergence of moments.

For a sequence k depending on n, define the random measure $\tilde{\mu}_n$ by

$$\tilde{\mu}_n = k^{1/2}\,(\tilde{\nu}_n - \nu)$$

and denote

$$W_n(x,y) = \tilde{\mu}_n((x,\infty) \times (-\infty, y]).$$

Proposition 7. *If Assumption 3 holds with $p^* \geq 2$ and $q^* \geq 4$ and if the sequence k is chosen such that (42) holds, then the sequence of processes W_n converges weakly in $\mathcal{D}$ to a Gaussian process W with autocovariance function*

$$\operatorname{cov}(W(x,y), W(x',y')) = \nu([x \vee x', +\infty) \times (-\infty, y \wedge y']) \,.$$

Moreover, the sequence of random measures $\tilde{\mu}_n$ converges weakly (in the sense of finite dimensional distributions) to an independently scattered Gaussian random measure W with control measure ν on the space of measurable functions g such that $|g(x,y)|^2 \leq C(x \vee 1)^{p^*}(|y| \vee 1)^{q^*}$, i.e. $W(g)$ is a centered Gaussian random variable with variance

$$\int_0^\infty \int_{-\infty}^\infty g^2(s,t)\, \nu(\mathrm{d}s, \mathrm{d}t)$$

and $W(g)$, $W(h)$ are independent if $\int gh \,\mathrm{d}\nu = 0$.

The proof is in section 7. Applying Proposition 7, we easily obtain the following corollary. For $i, j \geq 0$, denote $g_{i,j}(x,y) = x^i y^j \mathbb{1}_{\{x>0\}}$.

Corollary 8. *Under the assumptions of Proposition 7 and if moreover $\mu = 0$, then*

$$k^{1/2} \left\{ \frac{X_{(n:n-k)} - b(n/k)}{\psi \circ b(n/k)} \,, \frac{\hat{m}(X_{(n:n-k)}) - m \circ b(n/k)}{a \circ b(n/k)} \,, \frac{\hat{a}(X_{(n:n-k)})}{a \circ b(n/k)} - 1 \right\}$$

converges jointly with $k^{1/2}(\tilde{\nu}_n - \nu)$ to a Gaussian vector which can be expressed as $(W(g_{0,0}), (1-\gamma)W(g_{1,1}), \frac{1}{2}W(g_{0,2}))$.

Proposition 7 and Corollary 8 straightforwardly yield a functional central limit theorem for the estimator $\hat{\Psi}$ of Ψ defined in (19). Denote $G(x,y) = \nu([x,\infty) \times (-\infty, y])$.

Theorem 9. *If Assumption 3 holds, if $\mu = 0$ and if Ψ is absolutely continuous, then $k^{1/2}(\hat{\Psi} - \Psi)$ converges in $\mathcal{D}$ to the process M defined by*

$$\begin{aligned} M(y) &= W(0,y) \\ &\quad + \frac{\partial}{\partial x} G(0,y) W(g_{0,0}) + \Psi'(z)\{(1-\gamma)W(g_{1,1}) + \frac{1}{2}W(g_{0,2})z\} \,. \end{aligned} \tag{22}$$

Proof of Theorem 9. Recall the definitions of $\tilde{x}_n$, ξ_n and v_n in (12), (14) and (18), respectively. Then

$$\begin{aligned} &k^{1/2}\{\hat{\Psi}(y) - \Psi(y)\} \\ &\quad = k^{1/2}\{\tilde{\nu}_n([\tilde{x}_n, \infty) \times (-\infty, \xi_n + v_n y]) - \Psi(y)\} \\ &\quad = \tilde{\mu}_n([\tilde{x}_n, \infty) \times (-\infty, \xi_n + v_n y]) + k^{1/2}\{G(\tilde{x}_n, \xi_n + v_n y) - \Psi(y)\} \,. \end{aligned}$$

By Proposition 7, the first term in the right hand-side of the last displayed equation converges weakly to $W(0, y)$. By Corollary 8 and the delta method, $k^{1/2}\{G(\tilde{x}_n, \xi_n + v_n y) - \Psi(y)\}$ converges weakly to

$$\frac{\partial}{\partial x} G(0, y) W(g_{0,0}) + \Psi'(y)\{(1-\gamma) W(g_{1,1}) + \frac{1}{2} W(g_{0,2}) y\} \; .$$

□

4 Case of a product measure

In this section, guided by examples (see e.g. Fougères and Soulier (2008)), we make the following additional assumption.

Assumption 4.
The function ψ is an auxiliary function satisfying $\lim_{x\to\infty} \psi(x)/x = 0$, there exists $\rho \in \mathbb{R}$ such that $m(x) = \rho x$ and the measure ν is of the form

$$\nu((x, \infty) \times [-\infty, y]) = \mathrm{e}^{-x} \Psi(y) \; , \tag{23}$$

where Ψ is a distribution function on $\mathbb{R}$.

The condition $\lim_{x\to\infty} \psi(x)/x = 0$ implies that the extreme value index of X is 0. Heffernan and Resnick (2007, Proposition 2) provides a necessary and sufficient condition for ν to be a product measure.

Proposition 10. *The measure ν is a product measure if and only if $a \circ b$ is slowly varying at infinity and*

$$\lim_{t\to\infty} \frac{b(tx) - b(t)}{a \circ b(t)} = 0 \; . \tag{24}$$

The main consequence of Assumption 4 and of this result is that

$$\psi(x) = o(a(x)) \; ,$$

(by application of De Haan and Ferreira (2006, Theorem B.2.21)) and this implies that the deterministic centering in (2) can be replaced by a random centering. Specifically, Assumptions 2 and 4 imply

$$\lim_{t\to\infty} \mathbb{P}(X > t + \psi(t)x \, , \; Y - \rho X \le a(x) y \mid X > t) = \mathrm{e}^{-x} \Psi(y) \; .$$

4.1 Nonparametric estimation

Under Assumption 4, we can define new estimators of ρ, a and the marginal distribution Ψ as follows:

$$\hat{\rho} = \frac{\sum_{i=1}^{k} Y_{[n:n-i+1]}\{X_{(n:n-i+1)} - X_{(n:n-k)}\}}{\sum_{i=1}^{k} X_{(n:n-i+1)}\{X_{(n:n-i+1)} - X_{(n:n-k)}\}} , \tag{25}$$

$$\check{a}(X_{(n:n-k)}) = \left[\frac{1}{k}\sum_{i=1}^{k}\{Y_{[n:n-i+1]} - \hat{\rho}X_{(n:n-i+1)}\}^2\right]^{1/2} , \tag{26}$$

$$\check{\Psi}(z) = \frac{1}{k}\sum_{i=1}^{k} \mathbb{1}_{\{Y_{[n:n-i+1]} \leq \hat{\rho}X_{(n:n-i+1)} + \check{a}(X_{(n:n-k)})z\}} . \tag{27}$$

Proposition 11. *If Assumptions 1, 2 and 4 hold and if $\mu = 0$, then for any intermediate sequence k, $b(n/k)(\hat{\rho} - \rho)/a \circ b(n/k)$ converges weakly to 0, $\check{a}(X_{(n:n-k)})/a \circ b(n/k)$ converges weakly to 1 and $\check{\Psi}$ is a consistent estimator of Ψ. If $a(x) = o(x)$ then $\hat{\rho}$ converges weakly to ρ.*

In order to prove central limit theorems, we now introduce a second order assumption which is a modification of Assumption 3 in order to account for the random centering.

Assumption 5. *There exist positive real numbers p^* and q^*, a fonction $\tilde{c}$ and a locally bounded function $\tilde{\chi}$ on $\mathbb{R}^2$ such that*

$$|\mathbb{P}(X \leq t + \psi(t)x; Y - \rho X \leq a(t)y \mid X > t) - F(x,y)| \leq \tilde{c}(t)\tilde{\chi}(x,y) , \tag{28}$$

$$\int_0^\infty \int_{-\infty}^\infty x^{p^*+1}|y|^{q^*+1}\chi(x,y)\,\mathrm{d}x\,\mathrm{d}y < \infty . \tag{29}$$

The only difference with Assumption 3 is the presence of the term ρX instead of ρx. To illustrate the importance of this change, consider the case of a bivariate Gaussian vector with standard marginals and correlation ρ. Abdous et al. (2005) have shown that $\lim_{x\to\infty} \mathbb{P}(Y \leq \rho x + \sqrt{1-\rho^2}y \mid X > x) = \Phi(y)$ (where Φ is the distribution function of the standard Gaussian law), and a rate of convergence of order x^{-1} has been proved in Abdous et al. (2008). But of course, since $(Y - \rho X)/\sqrt{1-\rho^2}$ is standard Gaussian and independent of X, for all x it holds that $\mathbb{P}(Y \leq \rho X + \sqrt{1-\rho^2}y \mid X > x) = \Phi(y)$. For general elliptical bivariate random vectors, it is also proved in Abdous et al. (2008) that the rate of convergence with random centering can be the square of the rate with deterministic centering. Assumption 5 can also be checked for the generalized elliptical distributions studied in Fougères and Soulier (2008).

Theorem 12. *If Assumptions 4 and 5 hold and if the intermediate sequence k is chosen such that $\lim_{n\to\infty} k^{1/2}\tilde{c} \circ b(n/k) = 0$, then $k^{1/2}\{\check{\Psi} - \Psi\}$ converges weakly in $\mathcal{D}$ to the process M defined in (22).*

The proof of Proposition 11 and Theorem 12 is similar to the proof of the previous results. The only difference is that instead of the random measure $\tilde{\nu}_n$ defined in (7) we use the measure $\check{\nu}_n$, defined by

$$\check{\nu}_n = \frac{1}{k}\sum_{i=1}^{n} \delta_{(\{X_i - b(n/k)\}/\psi\circ b(n/k), \{Y_i - \rho X_i\}/a\circ b(n/k))} \ , \tag{30}$$

which converges weakly to the measure ν, and if k satisfies the condition stated in Theorem 12, then $k^{1/2}(\check{\nu}_n - \nu)$ converges to the Gaussian measure W defined in Proposition 7. We briefly sketch the proof of Proposition 11 at the end of Section 7. Theorem 12 is proved along the same lines as Theorem 9.

Kolmogorov-Smirnov Test

In the case $\gamma = 0$ and when the limiting measure ν has product form, then $\frac{\partial}{\partial x}G(0,y) = -\Psi(y)$. Define $B(t) = W(0, \Psi^{-1}(t))$. Then B is a standard Brownian motion on $[0,1]$ and

$$W(0,y) + \frac{\partial}{\partial x}G(0,y)W(g_{0,0}) = B\circ\Psi(y) - \Psi(y)B(1) = \mathcal{B}\circ\Psi(y)$$

where $\mathcal{B}$ is a standard Brownian bridge. Thus, $\sup_{y\in\mathbb{R}} |M(y)|$ has the same distribution as

$$Z = \sup_{t\in[0,1]} \left| \mathcal{B}(t) + \Psi'\circ\Psi^{-1}(t)\{U + \frac{1}{2}\Psi^{-1}(t)V\} \right| \ , \tag{31}$$

where U and V are jointly Gaussian random variables that can be expressed as

$$U = \int_0^1 \Psi^{-1}(s)\,\mathrm{d}B(s) + N \ , \quad V = \int_0^1 \{\Psi^{-1}(s)\}^2\,\mathrm{d}B(s) \ .$$

where N is a standard Gaussian random variable independent of the Brownian motion B. These extra terms come from the estimation of the functions a and m. If they were known, the limiting distribution would be the Brownian bridge as expected. Nevertheless, this distribution depends only on Ψ, so it can be used for a goodness-of-fit test. See Section 5.1 for a numerical illustration.

4.2 Semi-parametric estimation

Two problems arise in practice: the estimation of the conditional probability $\theta(x,y) = \mathbb{P}(Y \le y \mid X > x)$ and of the conditional quantile $y = \theta^{\leftarrow}(x,p)$ for

some fixed $p \in (0,1)$ and for some extreme x, i.e. beyond the range of the observations.

If x lies *within* the range of the observations, then $\theta(x,y)$ can be estimated empirically by

$$\hat{\theta}_{\text{emp}}(x,y) = \frac{1}{k}\sum_{i=1}^{n} \mathbb{1}_{\{Y_i \le y\}} \mathbb{1}_{\{X_i > x\}} \,,$$

for $x = X_{(n:n-k)}$. The most interesting situation for using the limit distributions that arise in Assumption 1 is when x is outside the range of the observations, so that an empirical estimate is no longer available. In such a situation, a semi-parametric approach will be needed to extrapolate the functions $a(x)$, $m(x)$ and $\psi(x)$ for values x beyond $X_{(n:n)}$. This requires some modeling restrictions. We still assume that Assumption 4 holds and we assume moreover that there exists $\sigma > 0$ such that

$$a(x) = \sigma\sqrt{x\psi(x)} \,. \tag{32}$$

We will also assume that the limiting distribution function Ψ in (23) is known. These assumptions hold in particular for bivariate elliptical distribution, see Abdous et al. (2008). There, and in many other examples, Ψ is the distribution function of the standard Gaussian law. See also Fougères and Soulier (2008). Assumption 4 and (32) imply that

$$\lim_{x\to\infty} \theta(x, \rho x + \sigma\sqrt{x\psi(x)}z) = \Psi(z)\,, \tag{33}$$

so that $\theta(x,y)$ can be approximated for x large enough by

$$\Psi\left(\frac{y-\rho x}{\sigma\sqrt{x\psi(x)}}\right) \,.$$

Thus, in order to estimate θ, we need a semi-parametric estimator of ψ. For this purpose, we make the following assumption on the marginal distribution of X.

Assumption 6. *The distribution function H of X satisfies*

$$1 - H(x) = \mathrm{e}^{-x^{\beta}\{c+O(x^{\beta\eta})\}}$$

with $\beta > 0$ and $\eta < 0$.

Under Assumption 6, an admissible auxiliary function is given by

$$\psi(x) = \frac{1}{c\beta} x^{1-\beta} \,. \tag{34}$$

Under (32), the normalizing function a is then

$$a(x) = \frac{\sigma}{\sqrt{c\beta}} x^{1-\beta/2} .$$

Let k and k_1 be intermediate sequences. Semi-parametric estimators of β and $a(x)$ are given by

$$\hat{\beta} = \frac{\sum_{i=1}^{k} \log\log(n/i) - \log\log(n/k)}{\sum_{i=1}^{k} \log(X_{(n:n-i+1)}) - \log(X_{(n:n-k)})} , \tag{35}$$

$$\breve{a}(x) = \breve{a}(X_{(n:n-k_1+1)}) \left(\frac{x}{X_{(n:n-k_1+1)}} \right)^{1-\hat{\beta}/2} . \tag{36}$$

Proposition 13. *If Assumption 6 holds, and if k is an intermediate sequence such that*

$$\lim_{n\to\infty} \log(k)/\log(n) = \lim_{n\to\infty} k \log^{2\eta}(n) = 0 ,$$

then $k^{1/2}(\hat{\beta} - \beta)$ converges weakly to $\mathbb{N}(0, \beta^{-2})$.

Suppose moreover that Assumptions 4, 5 and (32) hold. Let (x_n) be a sequence and k_1 be an intermediate sequence such that

$$\lim_{n\to\infty} k_1^{1/2} \tilde{c} \circ b(n/k_1) = \lim_{n\to\infty} k/k_1 = 0 ,$$
$$\lim_{n\to\infty} \log(b(n))/\log(x_n) = 1 ,$$
$$\lim_{n\to\infty} k^{-1/2} \log(x_n) = 0 .$$

Then $\dfrac{k^{1/2}}{\log(x_n)} \left\{ \dfrac{\breve{a}(x_n)}{a(x_n)} - 1 \right\}$ *converges weakly to* $\mathbf{N}(0, \beta^{-2})$.

Remark 7. A similar result could hold if $\lim_{n\to\infty} \log(b(n))/\log(x_n) = 0$, but this assumption may not be consistent with the other assumptions.

Proof of Proposition 13. The asymptotic normality of $\hat{\beta}$ is proved (under more general conditions) in Gardes and Girard (2006, Corollary 1). Consider now $\breve{a}(x_n)$.

$$\begin{aligned}
\frac{\breve{a}(x_n)}{a(x_n)} - 1 &= \frac{\hat{a}(X_{(n:n-k_1+1)})}{a(X_{(n:n-k_1+1)})} X_{(n:n-k_1)}^{(\hat{\beta}-\beta)/2} x_n^{(\beta-\hat{\beta})/2} - 1 \\
&= \left\{ \frac{\hat{a}(X_{(n:n-k_1+1)})}{a(X_{(n:n-k_1+1)})} - 1 \right\} X_{(n:n-k_1)}^{(\hat{\beta}-\beta)/2} x_n^{(\beta-\hat{\beta})/2} \\
&+ \left\{ X_{(n:n-k_1)}^{(\hat{\beta}-\beta)/2} - 1 \right\} \left\{ x_n^{(\beta-\hat{\beta})/2} - 1 \right\} + X_{(n:n-k_1)}^{(\hat{\beta}-\beta)/2} - 1 + x_n^{(\beta-\hat{\beta})/2} - 1 .
\end{aligned}$$

Since $\hat{\beta} - \beta = O_P(k^{-1/2})$, $\log(x_n) = o(k^{1/2})$ and $k/k_1 \to 0$, we obtain

$$x_n^{(\beta-\hat{\beta})/2} - 1 \sim (\beta - \hat{\beta}) \log(x_n)/2 \ ,$$
$$X_{(n:n-k_1+1)}^{(\beta-\hat{\beta})/2} - 1 \sim (\beta - \hat{\beta}) \log(X_{(n:n-k_1+1)})/2 \sim (\beta - \hat{\beta}) \log(b(n))/2 \ ,$$

where the equivalence relations above hold in probability. Finally,

$$\frac{k^{1/2}}{\log(x_n)} \left\{ \frac{\breve{a}(x_n)}{a(x_n)} - 1 \right\} = k^{1/2}(\beta - \hat{\beta}) + o_P(1) \ ,$$

and the proof follows from the asymptotic normality of $k^{1/2}(\beta - \hat{\beta})$. □

The previous results lead to two natural estimators of the conditional probability $\theta(x, y) = \mathbb{P}(Y \leq y \mid X > x)$ and of the conditional quantile $y = \theta^{\leftarrow}(x, p)$. Define

$$\hat{\theta}(x, y) = \Psi \left(\frac{y - \hat{\rho} x}{\breve{a}(x)} \right) \ . \tag{37}$$

Under Assumptions 2, 4 and (32), Proposition 11 implies that for *fixed* x and y, $\hat{\theta}(x, y)$ is a consistent estimator of $\Psi\left((y - \rho x)/a(x)\right)$, but a biased estimator of $\theta(x, y)$. The remaining bias, which is an approximation error due to the asymptotic nature of equation (33), can be bounded thanks to the second order Assumption 5. For more details, see (Abdous et al., 2008, Section 3.2) for a treatment in the elliptical case.

We now investigate more thoroughly the estimation of the conditional quantile $y_n = \theta^{\leftarrow}(x_n, p)$ for some fixed $p \in (0, 1)$ and some extreme sequence x_n, i.e. beyond the range of the observations, or equivalently, $x_n > b(n)$. An estimator $\hat{y}_n$ is defined by

$$\hat{y}_n = \hat{\rho} x_n + \breve{a}(x_n) \Psi^{-1}(p) \ . \tag{38}$$

Corollary 14. *Under the assumptions of Proposition 13, if* $\Psi'(p) > 0$*, if* $b(n/k_1) \asymp b(n) \asymp x_n$ *and if*

$$\lim_{n \to \infty} \frac{k^{1/2} \tilde{c} \circ b(x_n)}{\log(x_n)} = 0 \ , \tag{39}$$

then

$$\frac{k^{1/2} x_n}{\log(x_n) a(x_n)} \left\{ \frac{\hat{y}_n}{y_n} - 1 \right\}$$

converges weakly to a centered Gaussian law with variance $\left\{ \Psi^{-1}(p)/\rho\beta \right\}^2$.

Proof. Define $\tilde{y}_n = \rho x_n + a(x_n)\Psi^{-1}(p)$. Then

$$\begin{aligned}\hat{y}_n - y_n &= \hat{y}_n - \tilde{y}_n + \tilde{y}_n - y_n \\ &= (\hat{\rho} - \rho)x_n + (\hat{a}(x_n) - a(x_n))\Psi^{-1}(p) + \tilde{y}_n - y_n \ .\end{aligned}$$

In order to study $\tilde{y}_n - y_n$, denote $z_n = (y_n - \rho x_n)/a(x_n)$. Then $\lim_{n\to\infty} z_n = \Psi^{-1}(p)$. Indeed, if the sequence z_n is unbounded, then it tends to infinity at least along a subsequence. Choose $z > \psi^{-1}(p)$. Then, for large enough n,

$$\begin{aligned}p &= \mathbb{P}(Y \le \rho x_n + a(x_n)z_n \mid X > x_n) \ge \mathbb{P}(Y \le \rho x_n + a(x_n)z \mid X > x_n) \\ &\to \Psi(z) > p \ .\end{aligned}$$

Thus the sequence z_n is bounded, and if it converges to z (along a subsequence), it necessarily holds that $\psi(z) = p$, thus z_n converges to $\Psi^{-1}(p)$. Since we have assumed that $a(x) = o(x)$, this implies that $y_n \sim \rho x_n$ and

$$\frac{\tilde{y}_n - y_n}{y_n} \sim \frac{a(x_n)\{\Psi^{-1}(p) - z_n\}}{\rho x_n} \to 0 \ .$$

Moreover, since $\Psi'(p) > 0$, by a first order Taylor expansion, we have

$$\Psi^{-1}(p) - z_n = \frac{1}{\Psi'(\xi_n)}\{\theta(x_n, y_n) - \Psi(z_n)\} \ ,$$

where $\xi_n = \Psi^{-1}(p) + u\{z_n - \Psi^{-1}(p)\}$ for some $u \in (0,1)$. By Assumption 5, $\|\theta(x_n,\cdot) - \Psi\|_\infty = O(\tilde{c} \circ b(x_n))$. Since we have already shown that z_n converges to $\Psi^{-1}(p)$, $1/\Psi'(\xi_n)$ is bounded for large enough n, so $\Psi^{-1}(p) - z_n = O(\tilde{c} \circ b(x_n))$. Thus, by (39),

$$\frac{k^{1/2}x_n}{\log(x_n)a(x_n)}\frac{\tilde{y}_n - y_n}{y_n} = O\left(\frac{k^{1/2}\tilde{c} \circ b(x_n)}{\log(x_n)}\right) = o(1) \ .$$

Next,

$$\frac{\hat{y}_n - \tilde{y}_n}{y_n} = \frac{\hat{\rho} - \rho}{\rho} + \frac{a(x_n)\Psi^{-1}(p)}{\rho x_n}\left\{\frac{\breve{a}(x_n)}{a(x_n)} - 1\right\} \ .$$

Applying Proposition 11 and 13, if $k/k_1 \to 0$ and $x_n \asymp b(n)$, we obtain

$$\frac{k^{1/2}}{\log(x_n)}\frac{x_n}{a(x_n)}\frac{\hat{y}_n - \tilde{y}_n}{y_n} = \frac{k^{1/2}x_n(\hat{\rho} - \rho)}{\rho a(x_n)\log(x_n)} + \frac{\Psi^{-1}(p)}{\rho}\frac{k^{1/2}}{\log(x_n)}\left\{\frac{\breve{a}(x_n}{a(x_n)} - 1\right\} \ .$$

The first term in the right-hand side tends to zero by Proposition 11 and the assumptions on the sequences k_1, k and x_n. The second term converges weakly to $\Psi^{-1}(p)/(\rho\beta)\mathbf{N}(0,1)$ by Proposition 13. □

5 Numerical Illustration

In this section, we perform a small sample simulation study with two purposes. We analyze the behavior of the Kolmogorov-Smirnov test proposed in Section 4.1 and we illustrate the behavior of the estimator of the conditional quantile proposed in Section 4.2.

5.1 Goodness-of-fit test for the distribution Ψ

Assume that the hypotheses of Section 4 hold, so that the nonparametric estimation procedure described in Section 4.1 can be used. Three types of distributions are considered, each of them restricted to the positive quadrant for convenience. These distributions are:
(a) the elliptical distribution with radial survival function $P(R > t) = \mathrm{e}^{-t}$;
(b) the distribution with radial representation $R(\cos(T), \sin(T))$, where $P(R > t) = \mathrm{e}^{-t^2/2}$ and T has a non uniform concave density function $f_T(t) = 4/\left\{\pi + \pi(2t-1)^2\right\}$;
(c) the distribution with radial representation $R(\cos(T), \sin(T))$, where $P(R > t) = \mathrm{e}^{-t^2/2}$ and T has a non uniform convex density function $f_T(t) = 2 - 4/\left\{\pi(1 + (2t-1)^2\right\}$.

Case **(a)** is an example of the standard elliptical case, for which estimation results already exist (see Abdous et al. (2008)), whereas **(b)** and **(c)** illustrate the situation where the density level lines are "asymptotically elliptic" (see Fougères and Soulier (2008)). In these three cases, Ψ is the Normal distribution function (denoted by Φ). Figure 1 illustrates the estimation of Ψ via the nonparametric estimator $\check{\Psi}$ defined by (27) for one sample $(n = 1000, k = 100)$ of distribution **(b)**.

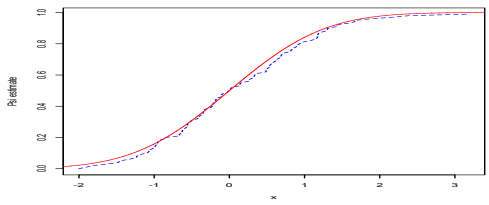

Figure 1: Estimation of Ψ via the nonparametric estimator $\check{\Psi}_n$ for one sample $(n = 1000, k = 100)$ of distribution **(b)**.

The Kolmogorov-Smirnov goodness-of-fit test performed here admits therefore as test statistic

$$T_{KS} = \sup_{y \in \mathbb{R}} \sqrt{k} |\check{\Psi}(y) - \Phi(y)| \, . \tag{40}$$

As shown in Section 4.1, T_{KS} has asymptotically the same distribution as the random variable Z defined in (31). Quantiles of this distribution have been obtained numerically and are listed in Table 1.

Table 1: *Quantiles q_α of order $1-\alpha$ of Z.*

α	0.01	0.05	0.10	0.15	0.20	0.25
q_α	1.598	1.297	1.174	1.076	1.029	0.980

We have compared these theoretical levels to the empirical levels obtained by simulation. In the three cases **(a)** to **(c)**, 10000 samples of size $n = 10^3$, 10^4 and 10^5, are simulated. The k observations having the largest first component are kept (for $k = 50, 100, 200$), and the nonparametric estimate $\check{\Psi}$ given in (27) is computed with this reduced sample. The observed values of the test statistic T_{KS} are compared to the quantiles listed in Table 1. For brevity, we present only the results corresponding to the two theoretical levels $\alpha = (0.05, 0.1)$. These empirical levels are shown in Table 2.

Table 2: *Empirical levels $(\hat{\alpha}_{0.05}, \hat{\alpha}_{0.1})$ associated to theoretical levels $(0.05, 0.1)$ for the goodness-of-fit test with statistic T_{KS}. The original sample size is denoted by n, and the number of observations used for the estimation is denoted by k. Notation* **(a)**–**(c)** *refers to the three bivariate distributions listed above.*

n	k	**(a)**	**(b)**	**(c)**
1000	50	$(0.057, 0.104)$	$(0.031, 0.062)$	$(0.02, 0.048)$
	100	$(0.131, 0.224)$	$(0.055, 0.102)$	$(0.04, 0.083)$
	200	$(0.599, 0.741)$	$(0.144, 0.241)$	$(0.157, 0.326)$
10000	50	$(0.042, 0.076)$	$(0.026, 0.055)$	$(0.027, 0.043)$
	100	$(0.062, 0.114)$	$(0.037, 0.073)$	$(0.04, 0.086)$
	200	$(0.096, 0.176)$	$(0.054, 0.993)$	$(0.074, 0.149)$
100000	50	$(0.036, 0.071)$	$(0.027, 0.054)$	$(0.026, 0.043)$
	100	$(0.034, 0.073)$	$(0.035, 0.069)$	$(0.045, 0.07)$
	200	$(0.065, 0.115)$	$(0.031, 0.069)$	$(0.078, 0.124)$

A common feature for the three distributions is that the results are rather

sensitive to the reduced number of observations k. However, the value of k leading to the best adequation between empirical and theoretical levels depends more on n than on the distribution, in most cases that we studied.

5.2 Semi-parametric estimation of the conditional quantile function

Assume from now on that the limiting distribution Ψ is the standard Gaussian distribution Φ and that Assumptions 4, 6 and equation (32) hold. The small sample behavior of the semi-parametric estimator $\hat{y}_n(p)$ of the quantile function $\theta^{\leftarrow}(x_n, p)$ defined by Equation (38) is illustrated in Figure 2 for the three distributions presented in Section 5.1. In each case, 100 samples of size 10000 are simulated. A proportion of 1% of the observations is used, which are the 100 observations with largest first component. For each sample, the conditional quantile function $\theta^{\leftarrow}(x, p)$ is estimated for two values of x corresponding to the theoretical X-quantiles of order $1 - \epsilon$, where $\epsilon = 10^{-4}$ and $\epsilon = 10^{-5}$. Figure 2 summarizes the quality of these estimations by showing the median, and the 2.5%- and 97.5%-quantiles of $\hat{y}_n(p)$ for the two fixed values of x specified above.

The estimation results are globally good, and the best ones are obtained for cases **(a)** and **(c)**, see rows 1 and 3 of Figure 2. Besides, one can observe a slight improvement as the conditioning event becomes more extreme.

6 Data analysis

To illustrate the use of the new procedures, and more specifically the Kolmogorov-Smirnov goodness-of-fit test proposed in Section 4.1, the hypothesis of $\Psi = \Phi$, where Φ is the standard Gaussian cdf, is tested using the series of monthly returns for the 3M stock and the Dow Jones Industrial Average from January 1970 to January 2008 ($n = 457$ values). These data were used by Levy and Duchin (2004) and revisited by Abdous et al. (2008). In the latter paper, the hypothesis of bivariate ellipticity was accepted through a test of elliptical symmetry proposed by Huffer and Park (2007) and the contagion from the Dow Jones to the 3M stock was tested. As shown in Abdous et al. (2005), ellipticity implies that Condition (1) holds and that the limiting distribution is the Gaussian law. The present procedure allows to test for the Gaussian conditional limit law without assuming ellipticity, but the weaker assumption (1). The observed values of the test statistic T_{KS} defined by (40) in terms of different choices of threshold k (or equivalently in terms of the proportion r of observations used, $k = nr$) are summarized in Table 3. According to Table 1, all these observed values correspond to a p-value greater than 0.25, which leads to accept the hypothesis $\Psi = \Phi$.

Table 3: *Observed values t_{KS} of the test statistic T_{KS} defined by (40) in terms of the proportion r or number k of observations used.*

p	0.05	0.10	0.15	0.20
k	22	45	68	91
t_{KS}	0.842	0.847	0.777	0.948

7 Proofs

Proof of Proposition 2. By (Resnick, 2007, Theorem 4.1), Assumption (2) implies that $\tilde{\nu}_n$ converges weakly to ν. We now prove (8). The convergence of $\hat{\nu}_n$ implies that for all compact set K of $[0,\infty)\times(-\infty,\infty)$ and any bounded function h, it holds that

$$\lim_{n\to\infty}\iint_K h(x,y)\tilde{\nu}_n([x,\infty)\times[y,\infty))\,\mathrm{d}x\,\mathrm{d}y = \iint_K h(x,y)\nu([x,\infty)\times[y,\infty))\,\mathrm{d}x\,\mathrm{d}y\ .$$

For $M>0$, define $K=[0,M]\times[0,M]$ and $K^c=[0,\infty)\times(-\infty,\infty)\setminus K$. Let h be a function on $[0,\infty)\times(-\infty,\infty)$ such that $h(x,y)\le C(x\vee 1)^{q^*-1}(|y|\vee 1)^{p^*-1}$. We must prove

$$\limsup_{M\to\infty}\lim_{n\to\infty}\iint_{K^c} h(x,y)\tilde{\nu}_n([x,\infty)\times[y,\infty))\,\mathrm{d}x\,\mathrm{d}y = 0\ , \tag{41}$$

in probability. Denote $G_n(x,y)=\nu_{n/k}((x,\infty)\times(-\infty,y])$. Then

$$\mathbb{E}[\tilde{\nu}_n(x,\infty)\times(y,\infty))] = G_n(x,y)\ .$$

Assumption 1 implies that G_n converges weakly to the function G defined by $G(x,y)=\nu([x,\infty]\times(-\infty,y])$. Il follows from Assumption 2 and integration by parts that

$$\int_0^\infty\int_{-\infty}^\infty h(x,y)G(x,y)\,\mathrm{d}x\,\mathrm{d}y<\infty\ .$$

This implies

$$\begin{aligned}\lim_{M\to\infty}\limsup_{n\to\infty}\iint_{K^c} h(x,y)\mathbb{E}[\tilde{\nu}_n((x,\infty)\times(y,\infty))]\,\mathrm{d}x\,\mathrm{d}y&\\ = \lim_{M\to\infty}\limsup_{n\to\infty}\iint_{K^c} h(x,y)G_n((x,y)\,\mathrm{d}x\,\mathrm{d}y&\\ = \lim_{M\to\infty}\iint_{K^c} h(x,y)G(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0\ .&\end{aligned}$$

This yields (41) and concludes the proof of Proposition 2. □

Lemma 15. *If Assumption 3 holds and if the sequence k is such that*

$$\lim_{n\to\infty} k^{1/2} c \circ b(n/k) = 0 \,, \tag{42}$$

then, for all measurable locally bounded functions h on $(0,\infty)\times(-\infty,\infty)$ such that

$$|h(x,y)| \le (x\vee 1)^{p^*}(|y|\vee 1)^{q^*} \,, \tag{43}$$

it holds that

$$\lim_{n\to\infty} \int_0^\infty \int_{-\infty}^\infty h(x,y)\,\mathrm{d}\mu_n = 0 \,. \tag{44}$$

Proof of Lemma 15. Denote $R_n(x,y) = \mu_n([0,x]\times(-\infty,y])$. With this notation, (20) implies that for all x, y,

$$|R_n(x,y)| \le k^{1/2} c\circ b(n/k)\chi(x,y) \,.$$

This bound and (42) yield

$$\lim_{n\to\infty} \int_0^\infty \int_{-\infty}^\infty (x\vee 1)^{p^*+1}(|y|\vee 1)^{q^*+1} R_n(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0 \,.$$

By integration by parts, this in turn implies that

$$\lim_{n\to\infty} \int_0^\infty \int_{-\infty}^\infty (x\vee 1)^{p^*}(|y|\vee 1)^{q^*} \mu_n(\mathrm{d}x,\mathrm{d}y) = 0 \,.$$

This implies (44) for any function h such that (43) holds. □

Proof of Proposition 7. We start by proving the convergence of the finite dimensional distributions of W_n. For the sake of concision, we introduce more notation. Denote $\tilde{X}_i = \{X_i - b(n/k)\}/\psi\circ b(n/k)$, $\tilde{Y}_i = \{Y_i - m\circ b(n/k)\}/a\circ b(n/k)$ and

$$\xi_{n,i}(x,y) = k^{-1/2}\{\mathbb{1}_{\{\tilde{X}_i > x\,,\,\tilde{Y}_i \le y\}} - kn^{-1}G_n(x,y)\} \,.$$

Then for each n, the random variables $\xi_{n,i}$, $1\le i\le n$ are i.i.d., centered,

$$\mathrm{cov}(\xi_{n,i}(x,y),\xi_{n,i}(x',y')) = \frac{1}{n}G_n(x\vee x', y\wedge y')\{1-\frac{k}{n}G_n(x\wedge x', y\vee y')\} \,,$$

and

$$W_n(x,y) = \sum_{i=1}^n \xi_{n,i}(x,y) + k^{1/2}\{G_n(x,y) - G(x,y)\} \,.$$

Assumption 3 and (42) imply that $k^{1/2}(G_n - G)$ converges to zero locally uniformly. The Lindeberg central limit theorem (cf. Araujo and Giné (1980)) and (42) yield the convergence of finite dimensional distributions of $\sum_{i=1}^{n} \xi_{n,i}(x,y)$. Tightness can be obtained as in Einmahl et al. (1993) by using an exponential inequality such as Inequality 1 in the aforementioned reference.

We now prove the second part of Proposition 7. Let h be C^∞ function with compact support in $\mathbb{R}^2$. The weak convergence of W_n in $\mathcal{D}$ implies that $\iint h(x,y)W_n(x,y)\,\mathrm{d}x\,\mathrm{d}y$ converges weakly to $\iint h(x,y)W(x,y)\,\mathrm{d}x\,\mathrm{d}y$. Thus, by integration by parts, it also holds that $\iint h(x,y)W_n(\mathrm{d}x,\mathrm{d}y)$ converges weakly to $\iint h(x,y)W(\mathrm{d}x,\mathrm{d}y)$. Let now g be a measurable function such that $|g(x,y)|^2 \leq C(x \vee 1)^{p^*}(|y| \vee 1)^{q^*}$. Then, for all $\epsilon > 0$, there exists a C^∞ function h with compact support such that $\int (g-h)^2\,\mathrm{d}\nu \leq \epsilon$. Then,

$$\int g\,\mathrm{d}\tilde{\mu}_n = \int h\,\mathrm{d}\tilde{\mu}_n + \int (g-h)\,\mathrm{d}\tilde{\mu}_n\ .$$

The first term in the right hand side converges weakly to $W(h)$ and we prove now that the second one converges in probability to 0. Denote $u = g - h$ and

$$\mu_n = k^{1/2}\{\nu_{n/k} - \nu\}\ .$$

Then,

$$\int u\,\mathrm{d}\tilde{\mu}_n = k^{-1/2}\sum_{i=1}^{n}\{u(\tilde{X}_i,\tilde{Y}_i) - \mathbb{E}[u(\tilde{X}_i,Y_i)]\} + \int u\,\mathrm{d}\mu_n\ .$$

This yields

$$\mathrm{var}(\int u\,\mathrm{d}\mu_n) \leq \int u^2\,\mathrm{d}\nu_n + \left\{\int u\,\mathrm{d}\mu_n\right\}^2\ .$$

By assumption on g, and since h has compact support, it also holds that $u^2(x,y) \leq C(x \vee 1)^{p^*}(|y| \vee 1)^{q^*}$. Thus, by Lemma 15,

$$\limsup_{n\to\infty} \mathrm{var}(\int u\,\mathrm{d}\mu_n) \leq \int u^2\,\mathrm{d}\nu \leq \epsilon\ .$$

Taking into account that $\mathrm{var}(W(g) - W(h)) = \mathrm{var}(W(g-h)) = \int (g-h)^2\,\mathrm{d}\nu \leq \epsilon$, we conclude that $W_n(g)$ converges weakly to $W(g)$. □

Proof of Corollary 8. We prove separately the claimed limit distributions. The joint convergence is obvious. We start with $\tilde{x}_n$, defined in (12). Denote $\mathbb{G}_n(x) = \tilde{\nu}_n((x,\infty)\times(-\infty,+\infty))$ and $\bar{G}_\gamma(x) = (1+\gamma x)^{-1/\gamma}$ if $\gamma > 0$ and $\bar{G}_0(x) = \mathrm{e}^{-x}$. By Proposition 7, $k^{1/2}(\mathbb{G}_n - \bar{G}_\gamma)$ converges weakly in $\mathcal{D}$ to the

process $B \circ \bar{G}_\gamma$, where B is a standard Brownian motion on $[0,1]$. By Verwaat's Lemma (de Haan and Resnick, 1993, Lemma A.1), $k^{1/2}\{\mathbb{G}_n^{\leftarrow} - \bar{G}_\gamma^{\leftarrow}\}$ jointly converges weakly in $\mathcal{D}$ to $-(\bar{G}_\gamma^{\leftarrow})'B$. Since $\mathbb{G}_n^{\leftarrow}(1) = \tilde{x}_n$, $\bar{G}_\gamma^{\leftarrow}(1) = 0$ and $(\bar{G}_\gamma^{\leftarrow})'(1) = -1$, we get the claimed limit distribution for $k^{1/2}\tilde{x}_n$.

We now consider ξ_n, defined in (14). By definition,

$$\begin{aligned}\xi_n &= \frac{\sum_{i=1}^k \{X_{(n:n-i+1)} - X_{(n:n-k)}\}\{Y_{[n:n-i+1]} - m \circ b(n/k)\}}{k\psi \circ b(n/k) a \circ b(n/k)} \\ &\qquad \div \frac{\sum_{i=1}^k \{X_{(n:n-i+1)} - X_{(n:n-k)}\}}{k\psi \circ b(n/k)} \\ &= \frac{\int_{\tilde{x}_n}^\infty \int_{-\infty}^\infty (x - \tilde{x}_n) y \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y)}{\int_{\tilde{x}_n}^\infty \int_{-\infty}^\infty (x - \tilde{x}_n) \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y)} \, .\end{aligned}$$

Since $\mu = 0$ by assumption, we obtain

$$k^{1/2}\xi_n = \frac{\int_{\tilde{x}_n}^\infty \int_{-\infty}^\infty (x - \tilde{x}_n) y \tilde{\mu}_n(\mathrm{d}x, \mathrm{d}y)}{\int_{\tilde{x}_n}^\infty \int_{-\infty}^\infty (x - \tilde{x}_n) \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y)} \, .$$

Since $\tilde{x}_n = O_P(k^{-1/2})$, it is easily seen that

$$k^{1/2}\xi_n = \frac{\int_0^\infty \int_{-\infty}^\infty x y \tilde{\mu}_n(\mathrm{d}x, \mathrm{d}y) + o_P(1)}{\int_0^\infty \int_{-\infty}^\infty x \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) + o_P(1)} \, .$$

Applying Propositions 2 and 7, we obtain that $k^{1/2}\xi_n$ converges weakly to $(1-\gamma)W(g_{1,1})$. Consider now $\hat{a}(X_{(n:n-k)})$. As in (15), we have

$$\frac{\hat{a}^2(X_{(n:n-k)})}{a^2 \circ (n/k)} = \int_{\tilde{x}_n}^\infty \int_{-\infty}^\infty y^2 \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) - 2\xi_n \int_{\tilde{x}_n}^\infty \int_{-\infty}^\infty y \tilde{\nu}_n(\mathrm{d}x, \mathrm{d}y) + \xi_n^2 \, .$$

Since $\tilde{x}_n = O_P(k^{-1/2})$ and $\xi_n = O_P(k^{-1/2})$, we get

$$k^{1/2}\left\{\frac{\hat{a}^2(X_{(n:n-k)})}{a^2 \circ b(n/k)} - 1\right\} = \int_0^\infty \int_{-\infty}^\infty y^2 \tilde{\mu}_n(\mathrm{d}x, \mathrm{d}y) + o_P(1) \, .$$

Proposition 7 and the delta method yield that $k^{1/2}\{\hat{a}(X_{(n:n-k)})/a \circ b(n/k) - 1\}$ converges weakly to $\frac{1}{2}W(g_{0,2})$. □

Proof of Proposition 11. Write $b(n/k)(\check{\rho}-\rho)/a\circ b(n/k) = S_n/T_n$ with

$$\begin{aligned}
S_n &= \frac{1}{k}\sum_{j=1}^{n}\left(\frac{X_i - b(n/k)}{\psi\circ b(n/k)} - \tilde{x}_n\right)_+ \left(\frac{Y_i-\rho X_i}{a\circ b(n/k)}\right) \\
&= \int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty}(x-\tilde{x}_n)y\,\check{\nu}_n(\mathrm{d}x,\mathrm{d}y) \to_P \mu\ , \\
T_n &= \frac{1}{k}\sum_{j=1}^{n}\left(\frac{X_i - b(n/k)}{\psi\circ b(n/k)} - \tilde{x}_n\right)_+ \\
&\quad + \frac{\psi\circ b(n/k)}{b(n/k)}\frac{1}{k}\sum_{j=1}^{n}\frac{X_i-b(n/k)}{\psi\circ b(n/k)}\left(\frac{X_i - b(n/k)}{\psi\circ b(n/k)} - \tilde{x}_n\right)_+ \\
&= \int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty}(x-\tilde{x}_n)\,\check{\nu}_n(\mathrm{d}x,\mathrm{d}y) \\
&\quad + \frac{\psi\circ b(n/k)}{b(n/k)}\int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty}x(x-\tilde{x}_n)\,\check{\nu}_n(\mathrm{d}x,\mathrm{d}y) = 1+o_P(1)\ .
\end{aligned}$$

The last term tends to zero since $\psi(x) = o(x)$.

Next, denote $\alpha_n = \psi\circ b(n/k)/b(n/k)$, $\zeta_n = b(n/k)(\check{\rho}-\rho)/a\circ b(n/k)$, $\tilde{Y}_i = (Y_i-\rho X_i)/a\circ b(n/k)$ and $\tilde{X}_i = \{X_i - b(n/k)\}/\psi\circ b(n/k)$. Then $\alpha_n\to 0$, $\zeta_n\to_P 0$ and

$$\begin{aligned}
\frac{\check{a}^2(X_{(n:n-k)})}{a^2\circ b(n/k)} &= \frac{1}{k}\sum_{i=1}^{n}\left\{\tilde{Y}_i + \alpha_n\zeta_n\tilde{X}_i + \zeta_n\right\}^2 \mathbb{1}_{\{\tilde{X}_i\geq\tilde{x}_n\}} \\
&= \int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty}\{y+\alpha_n\zeta_n x+\zeta_n\}^2\ \check{\nu}_n(\mathrm{d}x,\mathrm{d}y) \\
&= \int_0^{\infty}\int_{-\infty}^{\infty}y^2\nu(\mathrm{d}x,\mathrm{d}y) + o_P(1) = 1+o_P(1)\ .
\end{aligned}$$

Denote $\omega_n = \hat{a}(X_{(n:n-k)})/a\circ b(n/k) - 1$. Then $\omega_n\to_P 0$ and we can write

$$\begin{aligned}
\check{\Psi}(z) &= \frac{1}{k}\sum_{i=1}^{n}\mathbb{1}_{\{\tilde{Y}_i\leq\alpha_n\zeta_n\tilde{X}_i+\zeta_n+\omega_n z+z\}}\mathbb{1}_{\{\tilde{X}_i\geq\tilde{x}_n\}} \\
&= \int_{\tilde{x}_n}^{\infty}\int_{-\infty}^{\infty}\mathbb{1}_{\{y\leq\alpha_n\zeta_n x+\zeta_n+\omega_n z+z\}}\,\check{\nu}_n(\mathrm{d}x,\mathrm{d}y) \\
&\to_P \int_0^{\infty}\int_{-\infty}^{\infty}\mathbb{1}_{\{y\leq z\}}\,\nu(\mathrm{d}x,\mathrm{d}y) = \Psi(z)\ .
\end{aligned}$$

□

References

Belkacem Abdous, Anne-Laure Fougères, and Kilani Ghoudi. Extreme behaviour for bivariate elliptical distributions. *Revue Canadienne de Statistiques*, 33(2):1095–1107, 2005.

Belkacem Abdous, Anne-Laure Fougères, Kilani Ghoudi, and Philippe Soulier. Estimation of bivariate excess probabilities for elliptical models. *Bernoulli*, 14(4):1065–1088, 2008.

Aloisio Araujo and Evarist Giné. *The central limit theorem for real and Banach valued random variables*. John Wiley & Sons, New York-Chichester-Brisbane, 1980. Wiley Series in Probability and Mathematical Statistics.

Guus Balkema and Paul Embrechts. *High risk scenarios and extremes. A geometric approach.* Zurich Lectures in Advanced Mathematics. Zürich: European Mathematical Society, 2007.

Philippe Barbe. Approximation of integrals over asymptotic sets with applications to probability and statistics. http://arxiv.org/abs/math/0312132, 2003.

Jan Beirlant, Yuri Goegebeur, Jozef Teugels, and Johan Segers. *Statistics of extremes*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2004.

Simeon M. Berman. *Sojourns and extremes of stochastic processes*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1992.

B. O. Bradley and M. Taqqu. Framework for analyzing spatial contagion between financial markets. *Finance Letters*, 2(6):8–16, 2004.

Yu. Davydov and V. Egorov. Functional limit theorems for induced order statistics. *Mathematical Methods of Statistics*, 9(3):297–313, 2000.

Laurens De Haan and Ana Ferreira. *Extreme value theory. An introduction.* Springer Series in Operations Research and Financial Engineering. New York, NY: Springer., 2006.

Laurens de Haan and Sidney I. Resnick. Estimating the limit distribution of multivariate extremes. *Communications in Statistics. Stochastic Models*, 9(2):275–309, 1993.

William F. Eddy and James D. Gale. The convex hull of a spherically symmetric sample. *Advances in Applied Probability*, 13(4):751–763, 1981.

John H. J. Einmahl, Laurens de Haan, and Xin Huang. Estimating a multidimensional extreme-value distribution. *Journal of Multivariate Analysis*, 47(1):35–47, 1993.

Anne-Laure Fougères and Philippe Soulier. Modeling conditional excesses. Preprint avalaible as arXiv:0806.2426, 2008.

Laurent Gardes and Stéphane Girard. Comparison of Weibull tail-coefficient estimators. *REVSTAT*, 4(2):163–188, 2006.

Enkelejd Hashorva. Gaussian approximation of conditional elliptic random vectors. *Stoch. Models*, 22(3):441–457, 2006.

Enkelejd Hashorva, Samuel Kotz, and Alfred Kume. L_p-norm generalised symmetrised Dirichlet distributions. *Albanian Journal of Mathematics*, 1 (1):31–56 (electronic), 2007.

Janet E. Heffernan and Sidney I. Resnick. Limit laws for random vectors with an extreme component. *Annals of Applied Probability*, 17(2):537–571, 2007.

Janet E. Heffernan and Jonathan A. Tawn. A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society. Series B*, 66(3):497–546, 2004.

F. Huffer and C. Park. A test for elliptical symmetry. *Journal of Multivariate Analysis*, 98:256–281, 2007.

Haim Levy and Ran Duchin. Asset return distributions and the investment horizon. *Journal of Portfolio Management*, 30(3):47–62, 2004.

Sidney I. Resnick. *Heavy-Tail Phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.

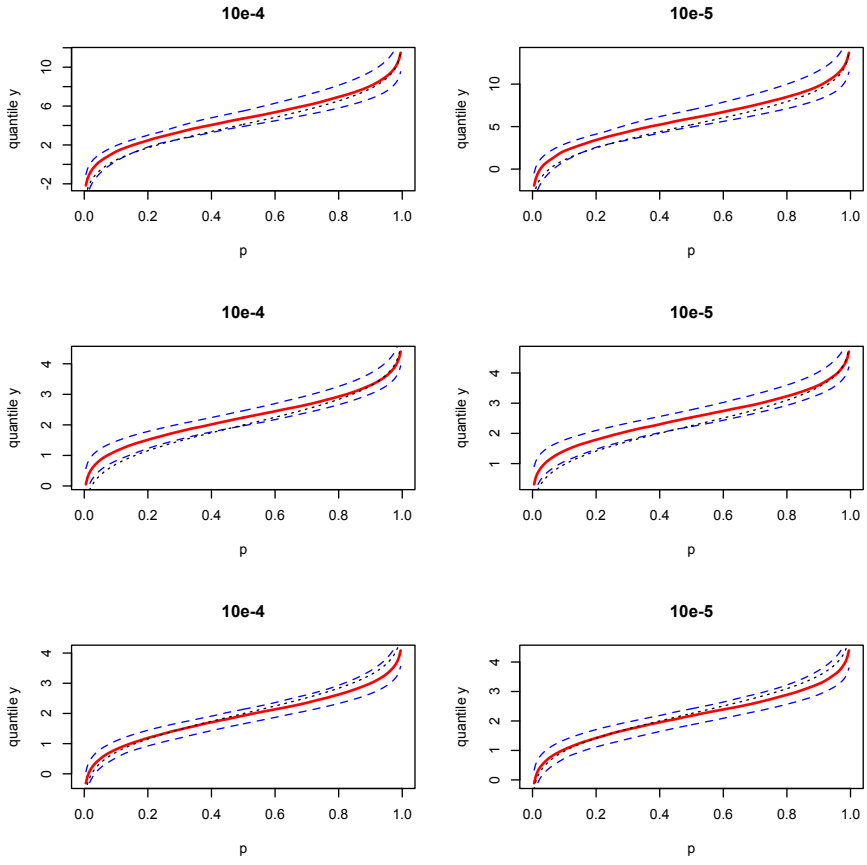

Figure 2: Median (solid line), 2.5%- and 97.5%-quantiles (dashed lines) of the estimated conditional quantile function $\hat{y} = \theta^{\leftarrow}(x, p)$ defined in (38) and theoretical conditional quantile function y (dotted line) as a function of the probability $p \in (0, 1)$. Each row (from 1 to 3) corresponds to a distribution (from **(a)** to **(c)**) as described in Section 5.1. Each column refers to a different value of x, respectively corresponding to the theoretical X-quantiles of order $1 - \epsilon$, where $\epsilon = 10^{-4}$ and $p = 10^{-5}$.