Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS
All Rights Reserved by Academic Press, New York and London

Vol. 47, No. 1, October 1993
Printed in Belgium

Estimating a Multidimensional Extreme-Value Distribution

JOHN H. J. EINMAHL

Eindhoven University of Technology, The Netherlands

AND

LAURENS DE HAAN AND XIN HUANG*

Erasmus University, Rotterdam, The Netherlands

Let F and G be multivariate probability distribution functions, each with equal one dimensional marginals, such that there exists a sequence of constants $a_n > 0$, $n \in \mathbb{N}$, with

$$\lim_{n \to \infty} F^n(a_n x_1, ..., a_n x_d) = G(x_1, ..., x_d),$$

for all continuity points $(x_1, ..., x_d)$ of G. The distribution function G is characterized by the extreme-value index (determining the marginals) and the so-called angular measure (determining the dependence structure). In this paper, a non-parametric estimator of G, based on a random sample from F, is proposed. Consistency as well as asymptotic normality are proved under certain regularity conditions. © 1993 Academic Press, Inc.

1. INTRODUCTION

Let F and G be non-degenerate bivariate probability distribution functions, each with equal marginals, such that there exists a sequence of constants $a_n > 0$, $n \in \mathbb{N}$, with

$$\lim_{n \to \infty} F^n(a_n x, a_n y) = G(x, y), \tag{1}$$

Received April 19, 1991; revised February 1, 1993.
AMS 1980 subject classifications: primary 62E20, 62G30; secondary 60F05, 60F15.
Key words and phrases: asymptotic normality, consistency, estimation, multivariate extremes.

* Now at the University of Beijing.

0047-259X/93 $5.00

for all continuity points (x, y) of G. We suppose that $G(x, y) < 1$ for all x, $y \in \mathbb{R}$.

The left hand side is the distribution function of

$$(a_n^{-1} \max(X_1, ..., X_n), a_n^{-1} \max(Y_1, ..., Y_n)),$$

where $(X_1, Y_1), ..., (X_n, Y_n)$ are i.i.d. random vectors with distribution function F. That is, G is a two-dimensional extreme-value distribution and F is in its (simple) domain of attraction.

It is known (see de Haan and Resnick, 1977), that G satisfying (1) must be "max-stable." That is to say, for all $m \in \mathbb{N}$ there exist constants $A_m > 0$, such that for all x, y

$$G^m(A_m x, A_m y) = G(x, y). \tag{2}$$

Consideration of the marginals (using the fact that $G(x, y) < 1$ for all $x, y \in \mathbb{R}$) shows that

$$-\log G(x, \infty) = -\log G(\infty, x) = (\beta x)^{-1/\gamma}, \qquad x > 0, \tag{3}$$

$$\log G(tx, ty) = t^{-1/\gamma} \log G(x, y) \qquad t, x, y > 0. \tag{4}$$

Here β and γ are positive constants. Note that $\{a_n\}$ and β are related: if we replace $\{a_n\}$ in (1) by $\{ca_n\}$, then $G(x, y)$ is replaced by $G(cx, cy)$, hence (3) is replaced by

$$-\log G(cx, \infty) = -\log G(\infty, cx) = (\beta cx)^{-1/\gamma}, \qquad x > 0.$$

So we could choose $\{a_n\}$ in (1) in such a way that (3) holds with $\beta = 1$. However, we normalize $\{a_n\}$ in a different way, explained shortly. Obviously we can rewrite (1) as

$$\lim_{n \to \infty} n\{1 - F(a_n x, a_n y)\} = -\log G(x, y), \tag{5}$$

for all continuity points $(x, y) \in (0, \infty)^2$ of G for which $0 < G(x, y) < 1$. Now (1) implies (de Haan and Resnick, 1977) that there is a measure μ on $[0, \infty)^2$, assigning finite values to Borel sets bounded away from the origin, such that

$$-\log G(x, y) = \mu\{(u, v) \mid u > x \text{ or } v > y\}. \tag{6}$$

Moreover, if (X, Y) is a random vector with distribution function F, then

$$\lim_{n \to \infty} nP\{(X, Y) \in a_n B\} = \mu(B) \tag{7}$$

for all Borel sets in $[0, \infty)^2$ with $\mu(\partial B)=0$ and B bounded away from the origin.

The measure μ can be further described as follows (de Haan and Resnick, 1977). If (1) holds, there exists a finite measure, called angular measure, on $[0, \pi/2]$ with distribution function Φ such that

$$\int_0^{\pi/2} \cos^{1/\gamma} \theta \, d\Phi(\theta) = \int_0^{\pi/2} \sin^{1/\gamma} \theta \, d\Phi(\theta), \tag{8}$$

$$\mu\{(x, y) : (x^2+y^2)^{1/2} \geqslant r, \operatorname{arctg}(y/x) \leqslant \theta\} = r^{-1/\gamma}\Phi(\theta), \tag{9}$$

$r>0$, $\theta \in [0, \pi/2]$.

In particular, combining (7) and (9), we get

$$\lim_{n\to\infty} nP\{(X^2+Y^2)^{1/2} \geqslant a_n r\} = r^{-1/\gamma}\Phi\left(\frac{\pi}{2}\right),$$

so that obviously the choice of $\{a_n\}$ determines $\Phi(\pi/2)$.

We write $S(r)=P\{(X^2+Y^2)^{1/2} \geqslant r\}$ and choose a_n to be $S^{\leftarrow}(1/n)$, where $S^{\leftarrow}$ denotes the right continuous inverse of S; to facilitate proofs we assume throughout S to be continuous. Then for the limiting function G in (1) we have $\Phi(\pi/2)=1$, i.e., the angular measure is a *probability* measure.

Clearly the distribution function G is characterized by the positive constant γ (the extreme-value index) and the probability distribution function Φ (determining the angular probability measure). Our object is to estimate the pair (γ, Φ) using a random sample from a distribution F in the domain of attraction of G, i.e., satisfying (1).

Denote with (ρ, Θ) the vector of polar coordinates of (X, Y); let $r>0$ and $0 \leqslant \theta \leqslant \pi/2$ and set $B=\{(x, y): x^2+y^2 \geqslant r^2,\ y/x \leqslant \operatorname{tg}\theta\}$ in (7). Then we get for all $r>0$ and $\theta \in T_\Phi := \{\theta \in [0, \pi/2)$: Φ is continuous at $\theta\} \cup \{\pi/2\}$

$$\lim_{n\to\infty} nP(\rho \geqslant a_n r, \Theta \leqslant \theta) = r^{-1/\gamma}\Phi(\theta) \tag{10}$$

and hence

$$\lim_{n\to\infty} nS(a_n r) = r^{-1/\gamma}, \qquad r>0. \tag{11}$$

By inverting (11) we obtain

$$\lim_{n\to\infty} \frac{S^{\leftarrow}(t/n)}{S^{\leftarrow}(1/n)} = \lim_{n\to\infty} S^{\leftarrow}(t/n)/a_n = t^{-\gamma}, \qquad t>0. \tag{12}$$

Write $R := S(\rho)$. For every $\theta \in T_\Phi$, the convergence in (10) is locally uniform for $r > 0$, therefore we have for all $r > 0$ and $\theta \in T_\Phi$

$$\begin{aligned}\lim_{n\to\infty} nP(R \leqslant r/n, \Theta \leqslant \theta) &= \lim_{n\to\infty} nP\left(\rho \geqslant a_n \frac{S^{\leftarrow}(r/n)}{a_n}, \Theta \leqslant \theta\right) \\ &= \lim_{n\to\infty} nP(\rho \geqslant a_n r^{-\gamma}, \Theta \leqslant \theta) \\ &= r\Phi(\theta). \end{aligned} \tag{13}$$

It is not hard to see that in the left hand side of (13), n may be interpreted as a continuous variable, i.e., running through $\mathbb{R}$.

Our main result is Theorem 2, stating joint asymptotic normality of estimators γ_n for γ and Φ_n for Φ. In fact the two estimators are asymptotically independent. This makes sense since the estimator γ_n is based on the ρ-component of the initial random vector and the estimator Φ_n is based mainly on the Θ-component. The two components are asymptotically independent in the following sense

$$\lim_{n\to\infty} \frac{P\{\rho \geqslant a_n r, \Theta \leqslant \theta\}}{P\{\rho \geqslant a_n\}} = r^{-1/\gamma}\Phi(\theta)$$

weakly (from (10)).

Finally we compare the framework of our results with that of some other papers. Most authors (e.g., Pickands, 1981; Deheuvels and Tiago de Oliveira, 1989; Tiago de Oliveira, 1989; Smith *et al.* 1990; and Deheuvels, 1991) construct estimators for the extreme value distribution G using observations from G itself rather than from a distribution function F in its domain of attraction.

Another distinction between our work and the mentioned papers is that previous authors assumed that the marginal distributions are known and concentrated on the estimation of the dependence structure (i.e., Φ). In contrast we do estimate Φ but also (jointly) γ, the asymptotic shape parameter of the marginal distributions.

A concrete problem where our approach seems to be relevant, is the two-dimensional sea dike problem. High tide water levels have been monitored in a continuous and reliable way at (at least) five stations on the Dutch coast. One is interested in the approximate distribution function of very high water levels—higher than the available observations—since these can cause floods. In fact, one wants to estimate the probability of a flood at either of the two sites, say Vlissingen (in the province of Zeeland) and Hoek van Holland (in the province of Zuid-Holland). To this end one has to estimate the limiting two-dimensional extreme value distribution G (i.e., γ and Φ) based on observations from the domain of attraction of G (since there is no reason to believe that the high tide water levels themselves follow an extreme-value distribution).

A (very rough) outline of the present paper appeared in de Haan (1985). The interested reader may also consult Geffroy (1958/1959), Tiago de Oliveira (1958), Sibuya (1960), Deheuvels (1978, 1980), and Resnick (1987) for more background.

2. Statistical Results

Let (X_1, Y_1), (X_2, Y_2), ..., be a sequence of independent random vectors with common distribution function F and let F be as in Section 1; recall that X_i, $Y_i \geqslant 0$, $i \in \mathbb{N}$. Denote the polar coordinates of (X_i, Y_i) by (ρ_i, Θ_i). We consider the problem of using the first n of these observations to estimate γ and Φ.

Throughout for an integer $1 \leqslant k \leqslant n$, $\rho_{k:n}$ will denote the kth order statistic from $\rho_1, \rho_2, ..., \rho_n$ and k_n will be a sequence of positive integers satisfying

$$1 \leqslant k_n \leqslant n/2, n \in \mathbb{N}, \qquad \text{and} \qquad k_n \to \infty, k_n/n \to 0 \ (n \to \infty).$$

Given k_n, introduce the estimators for γ and Φ by

$$\gamma_n = (\log \rho_{n-k_n+1:n} - \log \rho_{n-2k_n+1:n})/\log 2 \tag{14}$$

and

$$\Phi_n(\theta) = \frac{1}{k_n} \sum_{i=1}^{n} 1(\Theta_i \leqslant \theta, \rho_i \geqslant \rho_{n-k_n+1:n}), \qquad 0 \leqslant \theta \leqslant \pi/2, \tag{15}$$

respectively. The estimator γ_n is a simplified form of an estimator of γ proposed by Pickands (1975). The purpose of this paper is to prove that γ_n and Φ_n are consistent estimators and that $k_n^{1/2}(\gamma_n - \gamma, \Phi_n - \Phi)$ is jointly asymptotically normal. The results are captured in the following two theorems; the proofs are deferred to the next section.

THEOREM 1. *We have as* $n \to \infty$

$$\gamma_n \xrightarrow[P]{} \gamma$$

and

$$\Phi_n(\theta) \xrightarrow[P]{} \Phi(\theta), \qquad \textit{for} \quad \theta \in T_\Phi.$$

Moreover, if $k_n/\log\log n \to \infty$ $(n \to \infty)$, *then almost surely*

$$\gamma_n \to \gamma, \tag{16}$$

and, if $k_n/\log n \to \infty$ $(n \to \infty)$, *then almost surely*

$$\Phi_n(\theta) \to \Phi(\theta), \qquad \textit{for} \quad \theta \in T_\Phi. \tag{17}$$

Remark 1. Using more delicate arguments (cf. Theorem 4.6 in Einmahl, 1987), we can show that (17) holds under the weaker assumption that $k_n/\log\log n \to \infty$ $(n \to \infty)$. For the sake of brevity, Theorem 1 is proved as it stands.

The proof of the asymptotic normality requires some regularity conditions. These conditions are a natural strengthening of (1), more precisely of (13) and (11). Write $R_i = S(\rho_i)$, $i \in \mathbb{N}$. Assume there exists a $0 < \delta < 1$ such that

$$\lim_{n\to\infty} k_n^{1/2} \sup_{\substack{0 \leqslant \theta \leqslant \pi/2 \\ 1-\delta \leqslant r \leqslant 1+\delta}} \pi \left| \frac{n}{k_n} P\left(\Theta_1 \leqslant \theta, R_1 \leqslant \frac{k_n}{n} r\right) - r\Phi(\theta) \right| = 0. \tag{C1}$$

It readily follows from (11) that

$$\lim_{t\to\infty} S(tx)/S(t) = x^{-1/\gamma}, \tag{18}$$

for all $x > 0$. Inverting (18) we obtain

$$\lim_{t \downarrow 0} S^{\leftarrow}(tx)/S^{\leftarrow}(t) = x^{-\gamma}, \qquad x > 0. \tag{19}$$

Taking the logarithm on both sides yields

$$\lim_{t \downarrow 0} \log S^{\leftarrow}(tx) - \log S^{\leftarrow}(t) = -\gamma \log x. \tag{20}$$

Moreover this convergence is uniform for $x \in [\frac{1}{4}, 2\frac{1}{2}]$, which implies that there exists a positive function b satisfying (as $t \downarrow 0$), $b(t) \to 0$ and

$$\log S^{\leftarrow}(tx) - \log S^{\leftarrow}(t) = -\gamma \log x + \mathcal{O}(b(t)), \tag{21}$$

uniformly on $[\frac{1}{2}, 2\frac{1}{2}]$. We will need the following condition:

$$\lim_{n\to\infty} k_n^{1/2} b(k_n/n) = 0. \tag{C2}$$

THEOREM 2. *If* (C1) *and* (C2) *hold, then as* $n \to \infty$, $k_n^{1/2}(\gamma_n - \gamma, \Phi_n - \Phi)$ *converges weakly to* (Z, Λ) *on* $\mathbb{R} \times D[0, \pi/2]$ (*the D-space is equipped with the Skorohod* J_1*-topology*), *where* Z *is a mean zero Gaussian random variable with variance* $\gamma^2/(2\log^2 2)$ *and* Λ *is a mean zero Gaussian process with covariance structure*

$$E\Lambda(\theta_1)\,\Lambda(\theta_2) = \Phi(\theta_1 \wedge \theta_2) - \Phi(\theta_1)\,\Phi(\theta_2), \qquad 0 \leqslant \theta_1, \theta_2 \leqslant \pi/2.$$

Moreover, Z *and* Λ *are stochastically independent.*

Remark 2. Note that $\Lambda =^d B \circ \Phi$, where B is a Brownian bridge.

Remark 3. For convenience of writing we have chosen the bivariate set-up in this paper. Note however, that Theorems 1 and 2 remain true, *mutatis mutandis*, in the multivariate setting. To be more precise, when we denote the dimension with d ($\geqslant 3$), then the major changes are that $[0, \pi/2]$ has to be replaced by $[0, \pi/2]^{d-1}$ (see de Haan and Resnick, 1977, p. 320) and that, when dealing with θ's, "$\leqslant$," and "$\wedge$" have to be understood componentwise in the definitions and Theorem 2.

3. PROOFS

Before we begin with the proof of Theorem 1 we state a useful inequality. Let $Z_1, Z_2, \ldots$ be independent $\mathbb{R}^d$-valued random vectors with common distribution function C. For any Borel subset of $\mathbb{R}^d$ set

$$C(B) = P(Z_1 \in B),$$

$$C_n(B) = \frac{1}{n} \sum_{i=1}^{n} 1(Z_i \in B).$$

Let A denote any half-open rectangle in $\mathbb{R}^d$ of the form $\prod_{i=1}^{d} (a_i, b_i]$, $a_i < b_i$, $i = 1, \ldots, d$. For any $\lambda > 0$, let

$$\psi(\lambda) = 2\lambda^{-2}((1+\lambda)\log(1+\lambda) - \lambda).$$

This function has the property $\psi(\lambda) \uparrow 1$ as $\lambda \downarrow 0$.

Inequality 1 (Einmahl, 1987). Let $d \in \mathbb{N}$, $0 < C(A) \leqslant \frac{1}{2}$ and $0 < \delta < 1$. Then there exists a constant $K > 0$ depending only on d and δ such that

$$P(\sup_{\tilde{A} \subset A} |n^{1/2}(C_n(\tilde{A}) - C(\tilde{A}))| \geqslant \lambda)$$
$$\leqslant K \exp\left(\frac{-(1-\delta)\lambda^2}{2C(A)}\,\psi\left(\frac{\lambda}{n^{1/2}C(A)}\right)\right), \qquad \lambda > 0, \tag{22}$$

where $\tilde{A}$ denotes any half-open rectangle (of the same form as A).

We also need some more notation. Write

$$F_n(\theta, r) = \frac{n}{k_n} P\left(\Theta_1 \leqslant \theta,\, R_1 \leqslant \frac{k_n}{n} r\right), \qquad 0 \leqslant \theta \leqslant \pi/2,\, r \geqslant 0, \tag{23}$$

and denote the corresponding tail empirical distribution function by

$$\mathbb{F}_n(\theta, r) = \frac{1}{k_n} \sum_{i=1}^{n} 1\left(\Theta_i \leqslant \theta,\, R_i \leqslant \frac{k_n}{n} r\right). \tag{24}$$

Note that, with the obvious notation, almost surely,

$$\Phi_n(\theta) = \mathbb{F}_n\left(\theta, \frac{n}{k_n} R_{k_n:n}\right), \qquad 0 \leqslant \theta \leqslant \pi/2. \tag{25}$$

Proof of Theorem 1. The proofs of the "in probability" statements are omitted because they are similar to, but easier than, those of the "almost sure" statements, which follow now.

If $k_n/\log\log n \to \infty$, it readily follows from Wellner (1978) that almost surely

$$R_{2k_n:n} \to 0 \qquad \text{and} \qquad R_{k_n:n}/R_{2k_n:n \to 1/2}. \tag{26}$$

Combining this with (20) yields almost surely

$$\lim_{n\to\infty} \log S^{\leftarrow}(R_{k_n:n}) - \log S^{\leftarrow}(R_{2k_n:n}) \to \gamma \log 2. \tag{27}$$

Noting that $S^{\leftarrow}(R_i) = \rho_i$, almost surely, (16) follows.

Now assume $k_n/\log n \to \infty$. We have, writing $e_n = (n/k_n)\, R_{k_n:n}$,

$$\begin{aligned} &\sup_{0 \leqslant \theta \leqslant \pi/2} |\mathbb{F}_n(\theta, e_n) - F_n(\theta, 1)| \\ &\quad \leqslant \sup_{0 \leqslant \theta \leqslant \pi/2} |\mathbb{F}_n(\theta, e_n) - \mathbb{F}_n(\theta, 1)| \\ &\qquad + \sup_{0 \leqslant \theta \leqslant \pi/2} |\mathbb{F}_n(\theta, 1) - F_n(\theta, 1)| =: \Delta_{n1} + \Delta_{n2}. \end{aligned} \tag{28}$$

But almost surely

$$\begin{aligned} \Delta_{n1} &= |\mathbb{F}_n(\pi/2, e_n) - \mathbb{F}_n(\pi/2, 1)| \\ &= |\mathbb{F}_n(\pi/2, 1) - 1| = |\mathbb{F}_n(\pi/2, 1) - F_n(\pi/2, 1)| \leqslant \Delta_{n2}. \end{aligned} \tag{29}$$

Applying Inequality 1 with $\delta=\frac{1}{2}$, we obtain for any $\varepsilon>0$ and n large enough

$$P(\Delta_{n2}\geqslant\varepsilon)\leqslant K\exp(-k_n\varepsilon^2\psi(\varepsilon)/4)\leqslant n^{-2}. \tag{30}$$

Combination of (28)–(30) and an application of the Borel–Cantelli lemma now yields that the left hand side of (28) converges to zero almost surely. Hence, using (13), we see that almost surely

$$\Phi_n(\theta)=\mathbb{F}_n(\theta, e_n)\to\Phi(\theta) \qquad (n\to\infty), \tag{31}$$

for all $\theta\in T_\Phi$. ▮

The proof of Theorem 2 requires more notation. Write

$$H_n(\theta)=F_n(\theta,1);\,\mathbb{H}_n(\theta)=\mathbb{F}_n(\theta,1), \qquad 0\leqslant\theta\leqslant\pi/2, \tag{32}$$

$$\mathbb{J}_n(r)=\mathbb{F}_n(\pi/2,r), \qquad 0\leqslant r\leqslant 3, \tag{33}$$

and introduce the corresponding tail empirical processes

$$W_{1n}(\theta)=k_n^{1/2}(\mathbb{H}_n(\theta)-H_n(\theta)), \tag{34}$$

$$W_{2n}(r)=k_n^{1/2}(\mathbb{J}_n(r)-r). \tag{35}$$

Let (W_1, W_2) be a mean zero Gaussian process on $D[0,\pi/2]\times D[0,3]$ (we again tacitly assume that, as usual, the D-spaces are equipped with the Skorohod J_1-topology and the corresponding σ-algebra) with covariance structure

$$\begin{aligned}
&EW_1(\theta_1)\,W_1(\theta_2)=\Phi(\theta_1\wedge\theta_2),\\
&EW_2(r_1)\,W_2(r_2)=r_1\wedge r_2,\\
&EW_1(\theta_1)\,W_2(r_1)=(r_1\wedge 1)\,\Phi(\theta_1), \qquad 0\leqslant\theta_1,\theta_2\leqslant\pi/2,\ 0\leqslant r_1,r_2\leqslant 3.
\end{aligned} \tag{36}$$

The following result and its corollary are crucial for the proof of Theorem 2.

PROPOSITION 1. *If* (C1) *holds, then the process* (W_{1n}, W_{2n}) *converges weakly in* $D[0,\pi/2]\times D[0,3]$ *to* (W_1, W_2).

Write

$$W_{2n}^*(r)=k_n^{1/2}(\mathbb{J}_n^{\leftarrow}(r)-r), \qquad \tfrac{1}{2}\leqslant r\leqslant 2\tfrac{1}{2}, \tag{37}$$

where $\mathbb{J}_n^{\leftarrow}$ *is the left continuous inverse of* $\mathbb{J}_n$.

COROLLARY 1. *On a suitable probability space there exist a sequence of probabilistically equivalent versions* $(\tilde{W}_{1n}, \tilde{W}_{2n}^*)$ *of the processes* (W_{1n}, W_{2n}^*)

and a mean zero Gaussian process (W_1, W_2) *with covariance structure as in* (36) *such that, as* $n \to \infty$,

$$d(\tilde{W}_{1n}, W_1) \to 0 \qquad a.s., \tag{38}$$

$$\| \tilde{W}_{2n}^* + W_2 \|_{1/2}^{5/2} \to 0 \qquad a.s. \tag{39}$$

Here d *denotes "the" Skorohod metric and* $\| \cdot \|_a^b$ *the supremum–norm on* $[a, b]$.

Proof of Proposition 1. We have to show tightness of the sequence of processes (W_{1n}, W_{2n}) and the weak convergence of the finite dimensional distributions

$$(W_{1n}(\theta_1), ..., W_{1n}(\theta_k), W_{2n}(r_1), ..., W_{2n}(r_m)),$$

for all $\theta_1, ..., \theta_k \in T_\Phi$, $0 \leqslant r_1, ..., r_m \leqslant 3$, $k, m \in \mathbb{N}$, to

$$(W_1(\theta_1), ..., W_1(\theta_k), W_2(r_1), ..., W_2(r_m)).$$

The convergence of these finite dimensional distributions is standard and follows from (13) in combination with a multivariate version of Lindeberg's theorem (see Araujo and Giné, 1980, p. 41, Exercise 4)).

For the tightness condition it is enough to establish tightness for the marginal processes W_{1n} and W_{2n}. The proof for W_{2n} is similar to, but easier than, the one for W_{1n}. Therefore we only establish tightness for the sequence of processes W_{1n}. To do so we have to show (see Billingsley, 1968, p. 125) that for each $\varepsilon > 0$, there exists an M such that for all $n \in \mathbb{N}$

$$P(\| W_{1n} \|_0^{\pi/2} > M) \leqslant \varepsilon, \tag{40}$$

and that for each $\varepsilon > 0$ there exist $0 = \theta_0 < \theta_1 < \cdots < \theta_m = \pi/2$ such that for large n

$$P(\max_{1 \leqslant i \leqslant m} \sup_{\theta \in [\theta_{i-1}, \theta_i)} |W_{1n}(\theta) - W_{1n}(\theta_{i-1})| \geqslant \varepsilon) \leqslant \varepsilon. \tag{41}$$

Assertion (40) follows from an easy application of Inequality 1, using $\psi(\lambda) \sim (2 \log \lambda)/\lambda$ $(\lambda \to \infty)$. Now consider (41). We can obviously bound the left hand side by

$$\sum_{i=1}^{m} P(\sup_{\theta \in [\theta_{i-1}, \theta_i)} |W_{1n}(\theta) - W_{1n}(\theta_{i-1})| \geqslant \varepsilon). \tag{42}$$

Let $m \geqslant 1$ be an integer which will be specified later on and choose

$0 = \theta_0 < \theta_1 < \cdots < \theta_m = \pi/2$ such that for $1 \leqslant i \leqslant m$, $\Phi_-(\theta_i) - \Phi(\theta_{i-1}) \leqslant 1/m$. From (C1), it easily follows that for large n

$$\lim_{\theta \uparrow \theta_i} F_n(\theta, 1) - F_n(\theta_{i-1}, 1) \leqslant 2/m. \tag{43}$$

Now applying again Inequality 1 (the fact that not all rectangles involved are half-open causes no problems) and using $I\psi$ is increasing, we find for large n as an upper bound for the left hand side of (41)

$$Km \exp(-\varepsilon^2 m\psi(\varepsilon m/(2k_n^{1/2}))/8) \leqslant Km \exp(-\varepsilon^2 m/16). \tag{44}$$

This last expression is bounded by ε for m large enough. Thus we have shown that the sequence W_{1n} is tight. ▮

Proof of Corollary 1. The existence of a probability space on which we have (38) and (with the obvious notation)

$$\|\tilde{W}_{2n} - W_2\|_0^3 \to 0 \qquad \text{a.s.} \quad (n \to \infty) \tag{45}$$

is a consequence of the Skorohod representation theorem. Assertion (39) follows from (45) and Lemma 1 in Vervaat (1972). ▮

The next step is relating γ_n and Φ_n to the processes W_{1n} and W_{2n}^*. This is achieved in a couple of lemmas.

LEMMA 1. *If* (C2) *holds, then as* $n \to \infty$

$$\left| k_n^{1/2}(\gamma_n - \gamma) - \frac{\gamma}{\log 2}\left(\frac{1}{2} W_{2n}^*(2) - W_{2n}^*(1)\right)\right| \xrightarrow[P]{} 0. \tag{46}$$

Proof. Note that the left hand side of (46) is almost surely bounded from above by

$$\begin{aligned}
&\left| k_n^{1/2}\left(\frac{\log S^{\leftarrow}(R_{k_n:n}) - \log S^{\leftarrow}(R_{2k_n}:n)}{\log 2} - \gamma\right)\right.\\
&\quad \left. - \frac{\gamma k_n^{1/2}}{\log 2}\left(\left(\log\left(\frac{n}{k_n} R_{2k_n:n}\right) - \log 2\right) - \log\left(\frac{n}{k_n} R_{k_n:n}\right)\right)\right|\\
&\quad + \left|\frac{\gamma k_n^{1/2}}{\log 2}\left(\log\left(\frac{n}{k_n} R_{2k_n:n}\right) - \log 2\right) - \log\left(\frac{n}{k_n} R_{k_n:n}\right)\right)\\
&\quad \left. - \frac{\gamma k_n^{1/2}}{\log 2}\left(\frac{1}{2}\left(\frac{n}{k_n} R_{2k_n:n} - 2\right) - \left(\frac{n}{k_n} R_{k_n:n} - 1\right)\right)\right|.
\end{aligned} \tag{47}$$

It is easily seen that the first term tends to zero in probability from (21), (C2), and the fact that, as $n \to \infty$,

$$\frac{n}{k_n} R_{k_n : n} \xrightarrow[P]{} 1 \qquad \text{and} \qquad \frac{n}{k_n} R_{2k_n : n} \xrightarrow[P]{} 2. \tag{48}$$

An application of the mean value theorem in conjunction with Corollary 1 yields that the second term is also $o_P(1)$. ∎

LEMMA 2. *If* (C1) *holds, then as* $n \to \infty$

$$\| k_n^{1/2}(\Phi_n - \Phi) - (W_{1n} + W_{2n}^*(1)\Phi) \|_0^{\pi/2} \xrightarrow[P]{} 0. \tag{49}$$

Proof. Let $\varepsilon > 0$. First we show that for large n

$$P\Big(\sup_{0 \leqslant \theta \leqslant \pi/2} | k_n^{1/2}(\mathbb{F}_n(\theta, e_n) - F_n(\theta, e_n)) - W_{1n}(\theta)| \geqslant \varepsilon \Big) \leqslant \varepsilon. \tag{50}$$

Note that the left hand side of (50) is bounded from above by

$$P(k_n^{1/4}\,|e_n - 1| \geqslant 1) + P\Big(\sup_{\substack{0 \leqslant \theta \leqslant \pi/2 \\ |r-1| \leqslant k_n^{-1/4}}} | k_n^{1/2}(\mathbb{F}_n(\theta, r) - F_n(\theta, r)) - W_{1n}(\theta)| \geqslant \varepsilon \Big). \tag{51}$$

It easily follows from Corollary 1 that $k_n^{1/2}(e_n - 1) = \mathcal{O}_P(1)$, which implies that the first term is less than ε for large n. From Inequality 1 we see that the second term can be bounded by

$$K \exp(-\varepsilon^2 k_n^{1/4} \psi(\varepsilon/(2k_n^{1/4}))/8), \tag{52}$$

which is less than ε for large n.

Combining (50) with (C1) we readily obtain (49). ∎

Proof of Theorem 2. Combination of Proposition 1, Corollary 1, and Lemmas 1 and 2 reduces the proof to straightforward covariance calculations, which are omitted for the sake of brevity. ∎

REFERENCES

[1] ARAUJO, A., AND GINÉ, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables.* Wiley, New York.

[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] DEHEUVELS, P. (1978). Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes. *Publ. Inst. Statist. Univ. Paris* **23** 1–36.

[4] DEHEUVELS, P. (1980). Some applications of the dependence functions to statistical inference: nonparametric estimates of extreme values distributions, and a Kiefer type universal bound for the uniform test of independence. *Colloq. Math. Soc. János Bolyai* **32** 183–201.

[5] DEHEUVELS, P. (1991). On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions. *Statist. Probab. Lett.* **12** 429–439.

[6] DEHEUVELS, P., AND TIAGO DE OLIVEIRA, J. (1989). On the non-parametric estimation of the bivariate extreme-value distributions. *Statist. Probab. Lett.* **8** 315–323.

[7] EINMAHL, J. H. J. (1987). *Multivariate Empirical Processes.* CWI Tract 32, Mathematisch Centrum, Amsterdam.

[8] GEFFROY, J. (1958/1959). Contribution à la théorie des values extrêmes. *Publ. Inst. Statist. Univ. Paris* **7–8** 37–185.

[9] DE HAAN, L. (1985). Extremes in higher dimensions: the model and some statistics. In *Proc. 45th Session ISI (Amsterdam)*, 26.3.

[10] DE HAAN, L., AND RESNICK, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrsch. Verw. Gebiete* **40** 317–337.

[11] PICKANDS III, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3** 119–131.

[12] PICKANDS III, J. (1981). Multivariate extreme value distributions. In *Proc. 43rd Session ISI (Buenos Aires)*, pp. 859–878.

[13] RESNICK, S. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[14] SIBUYA, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math.* **11** 195–210.

[15] SMITH, R. L., TAWN, J. A., AND YUEN, H. K. (1990). Statistics of multivariate extremes. *Internat. Statist. Rev.* **58** 47–58.

[16] TIAGO DE OLIVEIRA, J. (1958). Extremal distributions. *Rev. Fac Ciencias Lisboa, Ser. 2A, Mat.* **7** 215–227.

[17] TIAGO DE OLIVEIRA, J. (1989). Intrinsic estimation of the dependence structure for bivariate extremes. *Statist. Probab. Lett.* **8** 213–218.

[18] VERVAAT, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. Verw. Gebiete* **23** 245–253.

[19] WELLNER, J. A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete* **45** 73–88.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium