Journal of Statistical Planning and Inference 139 (2009) 990 – 998

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Testing asymptotic independence in bivariate extremes

Jürg Hüsler[a,*], Deyuan Li[a,b]

[a]Department of Mathematical Statistics, University of Bern, Switzerland
[b]Department of Statistics, School of Management, Fudan University, China

ARTICLE INFO

Article history:
Received 22 February 2007
Received in revised form
9 June 2008
Accepted 12 June 2008
Available online 18 June 2008

Keywords:
Extremes
Bivariate random vectors
Asymptotic independence
Testing
Extreme value distributions

ABSTRACT

Bivariate extreme value condition (see (1.1) below) includes the marginal extreme value conditions and the existence of the (extreme) dependence function. Two cases are of interest: asymptotic independence and asymptotic dependence. In this paper, we investigate testing the existence of the dependence function under the null hypothesis of asymptotic independence and present two suitable test statistics. Small simulations are studied and the application for a real data is shown. The other case with the null hypothesis of asymptotic dependence is already investigated.

1. Introduction

Let (X, Y) be a bivariate random vector with continuous distribution function (d.f.) F. Suppose F belongs to the attraction of the max-domain of an extreme value distribution G, denoted by $F \in D(G)$, i.e. there exist normalized constants $a_n, c_n > 0$ and $b_n, d_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} F^n(a_n x + b_n, c_n y + d_n) = G(x, y) \tag{1.1}$$

for all but countably many x and y. Then, for a suitable choice of a_n, b_n, c_n and d_n, there exist $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$G(x, \infty) = \exp(-(1+\gamma_1 x)^{-1/\gamma_1}), \quad G(\infty, y) = \exp(-(1+\gamma_2 y)^{-1/\gamma_2}).$$

Condition (1.1) is the so-called bivariate extreme value condition, where γ_1 and γ_2 denote the (marginal) extreme value indices.

The extreme value distribution G can be characterized by a measure Λ on $[0,\infty]^2 \backslash \{(\infty,\infty)\}$, such that with

$$l(x, y) := -\log G\left(\frac{x^{-\gamma_1}-1}{\gamma_1}, \frac{y^{-\gamma_2}-1}{\gamma_2}\right) \tag{1.2}$$

we have

$$\begin{aligned} &1.\ l(x,y) = \Lambda(\{(u,v) \in [0,\infty]^2 : u \leqslant x \text{ or } v \leqslant y\}),\\ &2.\ l(tx,ty) = t\,l(x,y) \quad \text{for } t,x,y>0. \end{aligned} \tag{1.3}$$

* Corresponding author.
E-mail address: juerg.husler@stat.unibe.ch (J. Hüsler).

0378-3758/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2008.06.003

Denote the marginal d.f.'s of X and Y by $F_1(x)=F(x,\infty)$ and $F_2(y)=F(\infty,y)$, respectively. Then it follows that by (1.1) and (1.2),

$$\lim_{t\downarrow 0} t^{-1}P(1-F_1(X)\leqslant tx \text{ or } 1-F_2(Y)\leqslant ty)=l(x,y) \tag{1.4}$$

for $(x,y)\in[0,\infty)^2$. More generally

$$\lim_{t\downarrow 0} t^{-1}P((1-F_1(X),1-F_2(Y))\in tA)=\Lambda(A) \tag{1.5}$$

for any Borel set A in $[0,\infty]^2\backslash\{(\infty,\infty)\}$ (with $tA:=\{(tx,ty):(x,y)\in A\}$) provided $\Lambda(\partial A)=0$.

The function l is called the dependence function, which satisfies the well-known inequalities

$$x\vee y\leqslant l(x,y)\leqslant x+y, \quad x,y>0.$$

If $l(x,y)=x+y$ for all $x,y>0$, X and Y are called asymptotically independent (in this case $G(x,y)=G(x,\infty)G(\infty,y)$). Otherwise, X and Y are called asymptotically dependent. In particular, if $l(x,y)=x\vee y$ for all $x,y>0$, X and Y are called asymptotically fully dependent.

The bivariate extreme value condition (1.1) includes the marginal extreme value conditions and the existence of the dependence function l, which implies two cases: asymptotic independence (i.e. $l(x,y)=x+y$ for all $x,y>0$) and asymptotic dependence (i.e. all other cases of $l(x,y)$). As in the one-dimensional case, not all the bivariate distributions satisfy the bivariate extreme value condition. For example, an adaptive distribution in Schlather (2001) (see Section 3) satisfies the marginal extreme value distribution but its dependence function l does not exist. So before applying the bivariate extreme value theory to real data, we have to check assumption (1.1). For testing the marginal extreme value conditions, we refer to Dietrich et al. (2002), Drees et al. (2006) and Hüsler and Li (2006). Recently, testing asymptotic dependence was investigated. Draisma et al. (2004) provide one method to test it by assuming Ledford and Tawn's (1996–1998) sub-model. Einmahl et al. (2006) present another method to test it by comparing the difference of two non-parametric estimators of the dependence function l. In the two papers above, the alternative hypothesis is that the dependence function l does not exist or $l(x,y)=x+y$ for all $x,y>0$. So, rejecting the null hypothesis does not mean that we should reject $F\in D(G)$ since F may belong to $D(G)$ with asymptotic independence. Thus, in order to complete testing $F\in D(G)$, we need to test asymptotic independence, i.e. to test

$$\mathrm{H}_0: l(x,y)=x+y \quad \text{for all } x,y>0.$$

In this paper, we will construct two test statistics for H_0.

It should be mentioned that several articles investigate testing the independence in bivariate extremes. But all these tests are based on the sample from the extreme value distribution G, not from F itself, for example Deheuvels and Martynov (1996).

The rest of this paper is organized as follows. We present the main results in Section 2, and provide small simulations and an application in Section 3. The proofs of main results are shown in Section 4.

2. Main results

Let $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ be independent and identically distributed (i.i.d.) bivariate random vectors from continuous d.f. F such that $F\in D(G)$. In view of (1.4), one reasonable non-parametric estimator of l is

$$\hat{l}(x,y)=\frac{1}{k}\sum_{i=1}^{n} I_{\{X_i>X_{n,n+1-[kx]} \text{ or } Y_i>Y_{n,n+1-[ky]}\}}$$

(see Huang, 1992; Drees and Huang, 1998; Einmahl et al., 2006), where $X_{n,n+1-[kx]}$ and $Y_{n,n+1-[ky]}$ are the $[kx]$-th largest order statistics of $X_1,\ldots,X_n$ and the $[ky]$-th largest order statistics of $Y_1,\ldots,Y_n$, respectively, where $[x]$ denotes the largest integer smaller than or equal to x. We are going to use $k=k(n)$ as an intermediate sequence such that $k\to\infty$ and $k/n\to 0$ as $n\to\infty$. One natural way for testing H_0 is to construct the consistency and asymptotic normality of $\hat{l}$. But unfortunately, by Theorem 2.2 in Einmahl et al. (2006), it follows that under H_0

$$\sup_{x,y\in[0,1]} \sqrt{k}|\hat{l}(x,y)-(x+y)|\xrightarrow{\mathrm{P}} 0.$$

Thus $\hat{l}(x,y)$ cannot be used to construct a test statistic. The reason why the asymptotic variance of $\sqrt{k}(\hat{l}(x,y)-(x+y))$ vanishes is the dependence between X_i's and $X_{n,n+1-[kx]}$ and the dependence between Y_i's and $Y_{n,n+1-[ky]}$. Note that $X_{n,n+1-[kx]}$ and $Y_{n,n+1-[ky]}$ are the estimators of $F_1^{\leftarrow}(1-kx/n)$ and $F_2^{\leftarrow}(1-ky/n)$, respectively. To avoid the asymptotic variance to vanish, we may use other estimators for $F_1^{\leftarrow}(1-kx/n)$ and $F_2^{\leftarrow}(1-ky/n)$, which are independent with (X_i,Y_i)'s. For example, if n is an even, we can divide the sample into two sub-samples with equal sub-sample size. Then we use the second sub-sample to estimate the tail quantiles of F_1 and F_2, and define a new estimator of $l(x,y)$ using the first sub-sample. Since the two sub-samples are independent, the asymptotic variance does not vanish. We call the second sub-sample an auxiliary sample.

To generalize this idea and specify it, we assume $(X_1, Y_1), (X_2, Y_2), \ldots, (X_{n+m}, Y_{n+m})$ are i.i.d. bivariate random vectors with continuous d.f. F such that $F \in D(G)$. Denote $(\tilde{X}_i, \tilde{Y}_i) = (X_{n+i}, Y_{n+i})$ for $i = 1, 2, \ldots, m$. In addition, we assume $m = m(n)$ satisfying

$$m/n \to \theta > 0 \quad \text{and} \quad m/n - \theta = \mathrm{o}(n^{-1/2}). \tag{2.1}$$

Define a new estimator of l by

$$l_n(x, y) = \frac{m}{k}\frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \geqslant \tilde{X}_{m,m+1-[kx]} \text{ or } Y_i \geqslant \tilde{Y}_{m,m+1-[ky]}\}}, \tag{2.2}$$

where $\tilde{X}_{m,m+1-[kx]}$ and $\tilde{Y}_{m,m+1-[ky]}$ are the $[kx]$-th largest order statistics of $\tilde{X}_1, \ldots, \tilde{X}_m$ and the $[ky]$-th largest order statistics of $\tilde{Y}_1, \ldots, \tilde{Y}_m$, respectively. Under H_0 (asymptotic independence), it is easy to see that $l_n(x, y) \xrightarrow{\mathrm{P}} x + y$ for fixed $x, y > 0$, which motivates to consider

$$T_n(x, y) := \sqrt{k}(l_n(x, y) - (x + y)), \quad x, y \in [0, 1].$$

We will investigate the weak convergence of the stochastic process T_n and apply it to derive two test statistics for testing H_0. Before doing this, we assume some extra conditions as below.

Let $U = 1 - F_1(X)$ and $V = 1 - F_2(Y)$. Denote the d.f. of (U, V) by $C(x, y)$, i.e. $C(x, y) = \mathrm{P}(U \leqslant x, V \leqslant y)$. The function C is a two-dimensional copula. Under H_0, we obtain that

$$\lim_{t \downarrow 0} t^{-1}\mathrm{P}(U \leqslant tx \text{ or } V \leqslant ty) = l(x, y) = x + y$$

and that

$$\lim_{t \downarrow 0} t^{-1}\mathrm{P}(U \leqslant tx \text{ and } V \leqslant ty) = \lim_{t \downarrow 0} t^{-1}C(tx, ty) = 0.$$

Similar to Huang (1992) and Einmahl et al. (2006), we assume some weak conditions on C: there exists $\alpha > 0$ such that

$$t^{-1}C(tx, ty) = \mathrm{O}(t^{\alpha}) \quad \text{as } t \downarrow 0 \tag{2.3}$$

uniformly for $x, y \in [0, 1]$. Further, we require that there exist $0 < \varepsilon < 1$ and $0 < c_0 < 1$ such that

$$x - C(x, y) \geqslant c_0 x \quad \text{and} \quad y - C(x, y) \geqslant c_0 y \quad \text{for } x, y \in (0, \varepsilon] \tag{2.4}$$

The statement of weak convergence for the sequence of stochastic processes $\{T_n\}$ is as follows.

Theorem 2.1. *Assume $(X_1, Y_1), \ldots, (X_n, Y_n), (\tilde{X}_1, \tilde{Y}_1), \ldots, (\tilde{X}_m, \tilde{Y}_m)$ are i.i.d. bivariate random vectors from continuous d.f. F such that $F \in D(G)$ and its corresponding copula function C satisfies* (2.3) *and* (2.4) *for some $\alpha > 0$. Let k be an intermediate sequence such that $k = \mathrm{o}(n^{2\alpha/(1+2\alpha)})$ and m be a sequence such that* (2.1) *holds for some $\theta > 0$. Then under* H_0

$$\{T_n(x, y), x, y \in [0, 1]\} \xrightarrow{\mathrm{d}} \{W_1((1 + \theta)x) + W_2((1 + \theta)y),\ x, y \in [0, 1]\}$$

as $n \to \infty$, where W_1 and W_2 are two independent Brownian motions.

By Theorem 2.1 and continuous mapping theorem we obtain the following results.

Theorem 2.2. *Assume the conditions in Theorem* 2.1. *Then under* H_0

$$\int_0^1 \int_0^1 k(l_n(x, y) - (x + y))^2 \,\mathrm{d}x\,\mathrm{d}y \xrightarrow{\mathrm{d}} \int_0^1 \int_0^1 (W_1((1 + \theta)x) + W_2((1 + \theta)y))^2 \,\mathrm{d}x\,\mathrm{d}y$$

as $n \to \infty$, where W_1, W_2 are two independent Brownian motions.

Theorem 2.3. *Assume the conditions in Theorem* 2.1. *Then under* H_0

$$\sup_{x,y \in [0,1]} \sqrt{k}|l_n(x, y) - (x + y)| \xrightarrow{\mathrm{d}} \sup_{x,y \in [0,1]} |W_1((1 + \theta)x) + W_2((1 + \theta)y)|$$

as $n \to \infty$, where W_1, W_2 are two independent Brownian motions.

Remark 2.1. Theorem 2.2 also holds if we put some weight function $w(x, y)$ in the integrals, where $w(x, y)$ is such that $\int_0^1 \int_0^1 w(x, y)\,\mathrm{d}x\,\mathrm{d}y < \infty$. Theorems 2.2 and 2.3 will be used to construct the test for H_0, which is discussed in the next section. The two tests are similar to the Cramér–von Mises test and the Kolmogorov–Smirnov test, respectively.

3. Simulation and application

In this section we apply Theorems 2.2 and 2.3 to construct the test for H_0 and present a simulation study and an application for a real data set.

We set $n=m$. Then $\theta=1$. Let $I_n=\int_0^1\int_0^1 k(l_n(x,y)-(x+y))^2\,dx\,dy$, $S_n=\sup_{x,y\in[0,1]}\sqrt{k}|l_n(x,y)-(x+y)|$, $I=\int_0^1\int_0^1(W_1(2x)+W_2(2y))^2\,dx\,dy$ and $S=\sup_{x,y\in[0,1]}|W_1(2x)+W_2(2y)|$. Then $I_n\xrightarrow{d}I$ and $S_n\xrightarrow{d}S$ as $n\to\infty$. In application, we can calculate the values of I_n (or S_n) and compare it with the $(1-\alpha)$-th quantile of I (or S), denoted by $Q_{1-\alpha}^I$ (or $Q_{1-\alpha}^S$). If I_n (or S_n) is smaller than $Q_{1-\alpha}^I$ (or $Q_{1-\alpha}^S$), then we have no reason to reject H_0 with confidence level $1-\alpha$. Otherwise we may reject H_0. We call the two tests as integral test and supremum test, respectively. In the following simulation we set $\alpha=0.05$.

We first simulate the quantiles of the limiting distribution I and S. In order to do so, we generated 200,000 Brownian motions, approximate I and S, and use the empirical quantiles as the estimates for those of I and S. Tables 1 and 2 list the estimated quantiles of the limiting distribution I and S, respectively. The critical values for the integral test and the supremum test are 6.237 and 4.956, respectively.

Table 1
Quantiles of the limiting distribution I.

p	0.10	0.25	0.50	0.75	0.90	**0.95**	0.975	0.99
Q_p^I	0.416	0.668	1.259	2.517	4.573	**6.237**	8.040	10.449

Table 2
Quantiles of the limiting distribution S.

p	0.10	0.25	0.50	0.75	0.90	**0.95**	0.975	0.99
Q_p^S	2.041	2.409	2.956	3.669	4.444	**4.956**	5.420	6.024

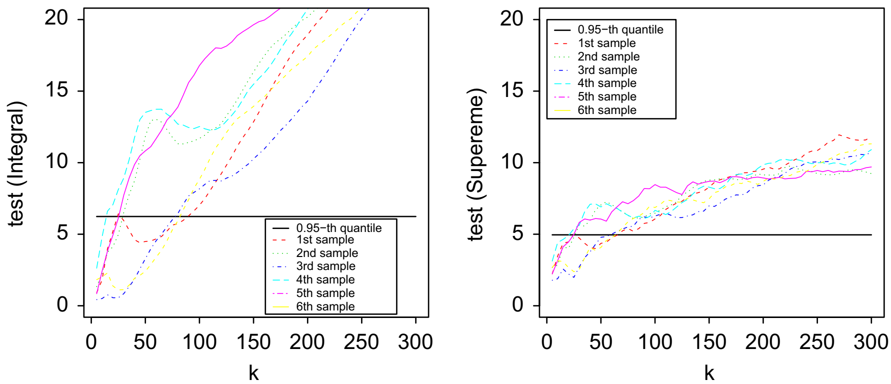

Fig. 1. Test statistics for the Cauchy distribution which satisfies the bivariate extreme value condition with *asymptotic dependence*: (left) integral test and (right) supremum test.

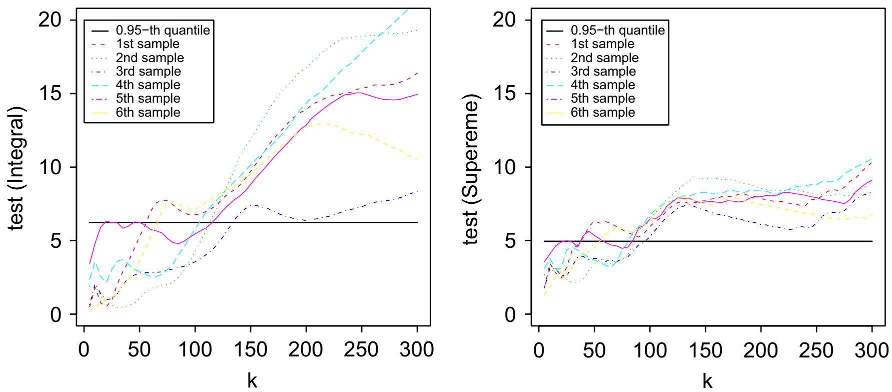

Fig. 2. Test statistics for the adaptive distribution which does not satisfy the bivariate extreme value condition: (left) integral test and (right) supremum test.

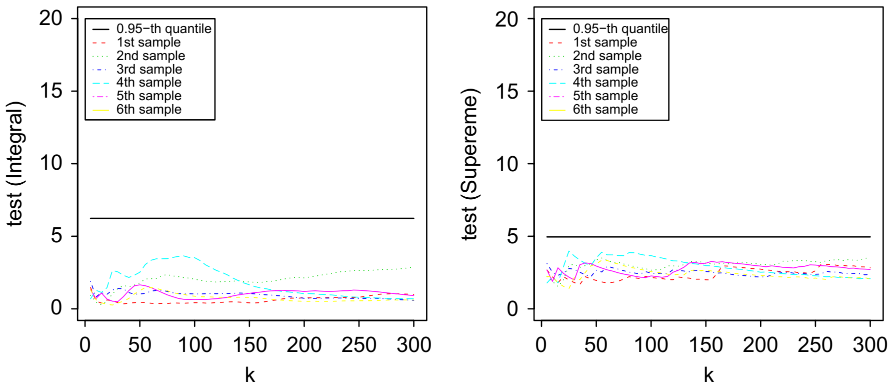

Fig. 3. Test statistics for the transformed Gumbel copula which satisfies the bivariate extreme value condition with *asymptotic independence*: (left) integral test and (right) supremum test.

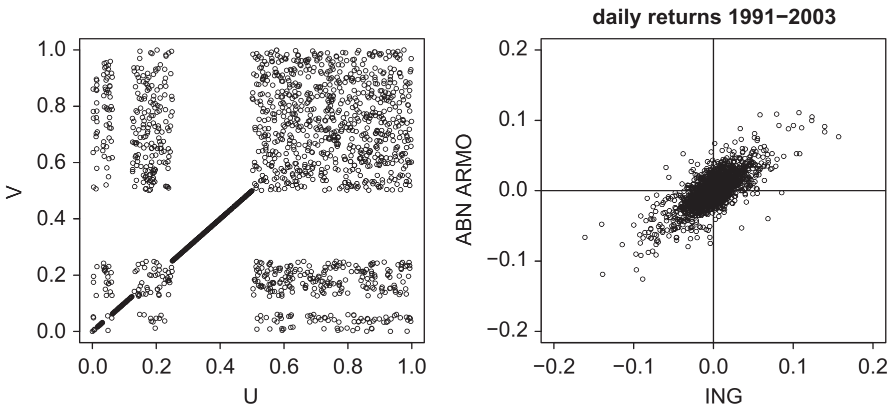

Fig. 4. (Left) One sample of (U,V) corresponding to the adaptive distribution with sample size 2000 and (right) daily equity returns of two Dutch banks with sample size 3283.

Secondly, we consider the following three distributions, which are also simulated in Einmahl et al. (2006). The first distribution is the bivariate Cauchy distribution restricted to the first quadrant with density

$$f(x,y)=\frac{2}{\pi(1+x^2+y^2)^{3/2}},\quad x,y>0.$$

This distribution satisfies the bivariate extreme value condition with asymptotic dependence. The second one is an adaption of a distribution in Schlather (2001), considering the distribution of $(1-U,1-V)$ where (U,V) has a density of $\frac{3}{2}$ on the following rectangles: $[2^{-(2s+1)},2^{-(2s)}]\times[2^{-(2t+1)},2^{-(2t)}]$, for $s=0,1,2,\dots$ and $t=0,1,2,\dots$; in this way a probability mass of $\frac{2}{3}$ is assigned. The remaining $\frac{1}{3}$ is assigned by taking the uniform distribution on the line segments from $(2^{-(2s+2)},2^{-(2s+2)})$ to $(2^{-(2s+1)},2^{-(2s+1)})$, $s=0,1,2,\dots$, such that the mass of the s-th segment is equal to $2^{-(2s+2)}$. Fig. 4 (left) shows one sample of (U,V) with sample size 2000. This adaptive distribution satisfies the marginal extreme value conditions but its dependence function l does not exist. The last example considers the distribution of $(U,1-V)$, where (U,V) has a Gumbel copula as d.f., i.e. the distribution is given by

$$C(u,v)=\exp(-[(-\log u)^{\beta}+(-\log v)^{\beta}]^{1/\beta}),\quad \beta\geqslant 1.$$

We take $\beta=10$. It is easy to check that the distribution of $(U,1-V)$ satisfies the bivariate extreme value condition with asymptotic independence.

For each distribution above, we generate 6 samples with sample size 2000 (i.e. $n=1000$ and using the second n observations as the auxiliary sample). The test statistics for different k and for each distribution are plotted in the following figures. As we see from Figs. 1 and 2, for most k the test statistics (both I_n and S_n) of the Cauchy distribution and the adaptive distribution are larger than the critical values (6.237 and 4.956, respectively). Thus we have no reason to accept H_0 for the two distributions. But for

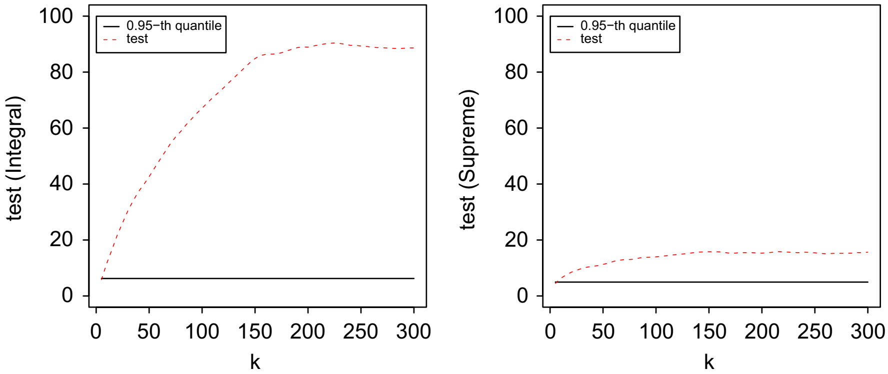

Fig. 5. Test statistics for the daily equity returns of two Dutch banks: (left) integral test and (right) supremum test.

the third distribution (with transformed Gumbel copula), the test statistics (both I_n and S_n) are smaller than the critical values for most k, see Fig. 3. So we may accept H_0 for this distribution. By the simulations, we see that both the integral test and the supremum test work quite well.

Finally, we apply Theorems 2.2 and 2.3 to a real data set: 3283 daily logarithmic equity returns over the period 1991–2003 for two Dutch banks, ING and ABN AMRO bank, see Fig. 4 (right). This data is also applied in Einmahl et al. (2006) and is accepted to satisfy the bivariate extreme value condition with asymptotic dependence. Fig. 5 shows that the test statistics (both I_n and S_n) are much larger than the critical values for most k. So we have no reason to accept H_0. This conclusion does not contradict the conclusion in Einmahl et al. (2006).

4. Proofs

Let $(U_i,V_i)=(1-F_1(X_i),1-F_2(Y_i))$ for $i=1,2,\dots,n$ and $(\tilde U_i,\tilde V_i)=(1-F_1(\tilde X_i),1-F_2(\tilde Y_i))$ for $i=1,2,\dots,m$. Then $(U_1,V_1),\dots,(U_n,V_n)$, $(\tilde U_1,\tilde V_1),\dots,(\tilde U_m,\tilde V_m)$ are i.i.d. random vectors with the common distribution C, which is defined in Section 2. We write

$$l_n(x,y)=\frac{m}{k}\frac{1}{n}\sum_{i=1}^{n} I_{\{U_i\leqslant \tilde U_{m,[kx]}\ \text{or}\ V_i\leqslant \tilde V_{m,[ky]}\}}$$

and so

$$\begin{aligned}T_n(x,y)&=\sqrt{k}(l_n(x,y)-(x+y))\\&=\sqrt{k}\left(\frac{m}{k}\frac{1}{n}\sum_{i=1}^{n} I_{\{U_i\leqslant \tilde U_{m,[kx]}\}}-\frac{m}{k}\tilde U_{m,[kx]}\right)+\sqrt{k}\left(\frac{m}{k}\frac{1}{n}\sum_{i=1}^{n} I_{\{V_i\leqslant \tilde V_{m,[ky]}\}}-\frac{m}{k}\tilde V_{m,[ky]}\right)\\&\quad+\sqrt{k}\left(\frac{m}{k}\tilde U_{m,[kx]}-x\right)+\sqrt{k}\left(\frac{m}{k}\tilde V_{m,[ky]}-y\right)-\sqrt{k}\left(\frac{m}{n}\frac{1}{k}\sum_{i=1}^{n} I_{\{U_i\leqslant \tilde U_{m,[kx]},V_i\leqslant \tilde V_{m,[ky]}\}}\right)\\&=:T_{n1}(x)+T_{n2}(y)+T_{n3}(x)+T_{n4}(y)-T_{n5}(x,y).\end{aligned}\tag{4.1}$$

In order to derive the weak convergence of T_n, it suffices to consider those of T_{ni}'s. The following two propositions play the central role.

Proposition 4.1. *Assume* $(U_1,V_1),\dots,(U_n,V_n),(\tilde U_1,\tilde V_1),\dots,(\tilde U_m,\tilde V_m)$ *are i.i.d. random vectors with marginal uniform distributions and its continuous d.f.* C *satisfies* (2.3) *and* (2.4) *for some* $\alpha>0$. *Let* k *be an intermediate sequence such that* $k=o(n^{2\alpha/(1+2\alpha)})$ *and* m *be a sequence such that* (2.1) *holds for some* $\theta>0$. *Then under* H_0

$$(T_{n1}(u),u\in[0,1],T_{n2}(v),v\in[0,1],T_{n3}(s),s\in[0,1],T_{n4}(t),t\in[0,1])$$

converges in distribution to

$$(W_1(\theta u),u\in[0,1],W_2(\theta v),v\in[0,1],W_3(s),s\in[0,1],W_4(t),t\in[0,1])$$

as $n\to\infty$, *where* W_1–W_4 *are four independent Brownian motions.*

Proposition 4.2. *Assume the conditions in Proposition* 4.1. *Then under* H_0

$$\sup_{x,y\in[0,1]} T_{n5}(x,y) \xrightarrow{\mathrm{P}} 0.$$

Proof of Proposition 4.1. By Proposition 3.1 in Einmahl et al. (2006), we obtain that under H_0 (asymptotic independence) there exist two independent sequences of Brownian motions W_{n1} and W_{n2} such that for any positive constant T,

$$\sup_{0<t\leqslant T}\left|k^{1/2}\left(\frac{1}{k}\sum_{i=1}^{n} I_{\{U_i\leqslant kt/n\}}-t\right)-W_{n1}(t)\right|\xrightarrow{\mathrm{P}} 0 \tag{4.2}$$

and

$$\sup_{0<t\leqslant T}\left|k^{1/2}\left(\frac{1}{k}\sum_{i=1}^{n} I_{\{V_i\leqslant kt/n\}}-t\right)-W_{n2}(t)\right|\xrightarrow{\mathrm{P}} 0 \tag{4.3}$$

as $n\to\infty$ for all intermediate sequences k such that $k\to\infty$ and $k/n\to 0$. This follows by considering the last two components of the random vector with three components in the mentioned proposition, and choosing $\eta=0$.

Relations (4.2) and (4.3) can also be obtained from other references, e.g. see Csörgő and Horváth (1993, Theorem 5.1.5), Einmahl (1997, Corollary 3.3). But from Proposition 3.1 in Einmahl et al. (2006), we can easily have the independence between W_{n1} and W_{n2} by assumption (2.3).

By the approximations of the tail quantile process (e.g. see Csörgő and Horváth, 1993, Theorem 5.2.5), there exist two sequences of Brownian motions $\tilde{W}_{m1}$ and $\tilde{W}_{m2}$ such that

$$\sup_{0<x\leqslant 1}\left|k^{1/2}\left(\frac{m}{k}\tilde{U}_{m,[kx]}-x\right)-\tilde{W}_{m1}(x)\right|\xrightarrow{\mathrm{P}} 0 \tag{4.4}$$

and

$$\sup_{0<y\leqslant 1}\left|k^{1/2}\left(\frac{m}{k}\tilde{V}_{m,[ky]}-y\right)-\tilde{W}_{m2}(y)\right|\xrightarrow{\mathrm{P}} 0 \tag{4.5}$$

as $m\to\infty$ for all intermediate sequences k such that $k\to\infty$ and $k/m\to 0$. Because of assumption (2.3) and the relation between the tail empirical process and tail quantile process, $\tilde{W}_{m1}$ and $\tilde{W}_{m2}$ can be chosen to be independent. Since $\{(U_i,V_i)\}_{i=1}^{n}$ and $\{(\tilde{U}_i,\tilde{V}_i)\}_{i=1}^{m}$ are independent, obviously we may assume (W_{n1},W_{n2}) is independent with $(\tilde{W}_{m1},\tilde{W}_{m2})$ for $n,m\in\mathbb{N}$.

Replacing t by $n/k\tilde{U}_{m,[kx]}$ in (4.2), we have

$$\sup_{0<x\leqslant 1}\left|k^{1/2}\left(\frac{1}{k}\sum_{i=1}^{n} I_{\{U_i\leqslant \tilde{U}_{m,[kx]}\}}-\frac{n}{k}\tilde{U}_{m,[kx]}\right)-W_{n1}\left(\frac{n}{k}\tilde{U}_{m,[kx]}\right)\right| I_{\{n/k\tilde{U}_{m,[kx]}\leqslant 2n/m\}}\xrightarrow{\mathrm{P}} 0.$$

Thus

$$\sup_{0<x\leqslant 1}\left|T_{n1}(x)-\frac{m}{n}W_{n1}\left(\frac{n}{k}\tilde{U}_{m,[kx]}\right)\right| I_{\{m/k\tilde{U}_{m,[kx]}\leqslant 2\}}\xrightarrow{\mathrm{P}} 0.$$

On the other hand, by the modulus of continuity for Brownian motion, (2.1), and the fact $\sup_{0<x\leqslant 1}|m/k\tilde{U}_{m,[kx]}-x|=o_{\mathrm{P}}(k^{-1/2})$ (by (4.4)), it follows that

$$\sup_{0<x\leqslant 1}\left|\frac{m}{n}W_{n1}\left(\frac{n}{k}\tilde{U}_{m,kx}\right)-\theta W_{n1}(\theta^{-1}x)\right| I_{\{m/k\tilde{U}_{m,[kx]}\leqslant 2\}}=o_{\mathrm{P}}(1)$$

and so

$$\sup_{0<x\leqslant 1}|T_{n1}(x)-\theta W_{n1}(\theta^{-1}x)| I_{\{m/k\tilde{U}_{m,[kx]}\leqslant 2\}}\xrightarrow{\mathrm{P}} 0.$$

Since $\mathrm{P}(m/k\tilde{U}_{m,[kx]}>2)\to 0$, we obtain

$$\sup_{0<x\leqslant 1}|T_{n1}(x)-\theta W_{n1}(\theta^{-1}x)|\xrightarrow{\mathrm{P}} 0. \tag{4.6}$$

Similarly, we also obtain

$$\sup_{0<y\leqslant 1}|T_{n2}(y)-\theta W_{n2}(\theta^{-1}y)|\xrightarrow{\mathrm{P}}0. \tag{4.7}$$

Relations (4.4)–(4.7) imply that there exist four independent Brownian motions W_1–W_4 on the same probability space such that

$$(T_{n1}(u),u\in[0,1],T_{n2}(v),v\in[0,1],T_{n3}(s),s\in[0,1],T_{n4}(t),t\in[0,1])$$

converges in distribution to

$$(W_1(\theta u),u\in[0,1],W_2(\theta v),v\in[0,1],W_3(s),s\in[0,1],W_4(t),t\in[0,1]).$$

Thus the statement of Proposition 4.1 follows. □

Proof of Proposition 4.2. Since $0\leqslant\sup_{x,y\in[0,1]}T_{n5}(x,y)=T_{n5}(1,1)$ a.s., it suffices to show $T_{n5}(1,1)\xrightarrow{\mathrm{P}}0$.

For $r=1,2$

$$\begin{aligned}EC^r(\tilde U_{m,k},\tilde V_{m,k})&=E(C^r(\tilde U_{m,k},\tilde V_{m,k})I_{\{m/k\tilde U_{m,k}\leqslant 2,m/k\tilde V_{m,k}\leqslant 2\}})\\&\quad+E(C^r(\tilde U_{m,k},\tilde V_{m,k})I_{\{m/k\tilde U_{m,k}>2\text{ or }m/k\tilde V_{m,k}>2\}})\\&\leqslant C^r(2k/m,2k/m)+2\mathrm{P}(\tilde U_{m,k}>2k/m).\end{aligned}$$

Note that, for large k such that $k/m\to 0$,

$$\begin{aligned}P\left(\tilde U_{m,k}>\frac{2k}{m}\right)&\leqslant P\left(\tilde U_{m,m-k}<1-\frac{2k}{m}\right)\\&=\sum_{i=m-k}^{m}\binom{m}{i}\left(1-\frac{2k}{m}\right)^{m-i}\left(\frac{2k}{m}\right)^{i}\leqslant m^{k+1}\left(\frac{2k}{m}\right)^{m-k}.\end{aligned}$$

Hence

$$\begin{aligned}ET_{n5}(1,1)&=E(ET_{n5}(1,1)|(\tilde U_{m,k},\tilde V_{m,k}))=Ek^{-1/2}\frac{m}{k}C(\tilde U_{m,k},\tilde V_{m,k})\\&\leqslant k^{-1/2}\frac{m}{k}C\left(\frac{2k}{m},\frac{2k}{m}\right)+2k^{-1/2}\frac{m}{k}\mathrm{P}(\tilde U_{m,k}>2k/m).\end{aligned}$$

By (2.1), (2.3) and $k=o(n^{2\alpha/(1+2\alpha)})$, we obtain $k^{-1/2}(m/k)C(2k/m,2k/m)\to 0$ and $k^{-1/2}(m/k)\mathrm{P}(\tilde U_{m,k}>2k/m)\to 0$, which imply $ET_{n5}(1,1)\to 0$.

On the other hand,

$$T^2_{n5}(1,1)=\frac{m^2}{n^2}\frac{1}{k}\sum_{i=1}^{n}I_{\{U_i\leqslant\tilde U_{m,k},V_i\leqslant\tilde V_{m,k}\}}+\frac{m^2}{n^2}\frac{1}{k}\sum_{i\neq j}I_{\{U_i\leqslant\tilde U_{m,k},V_i\leqslant\tilde V_{m,k}\}}I_{\{U_j\leqslant\tilde U_{m,k},V_j\leqslant\tilde V_{m,k}\}}$$

and

$$\begin{aligned}ET^2_{n5}(1,1)&=E(ET^2_{n5}(1,1)|(\tilde U_{m,k},\tilde V_{m,k}))\\&=E\left(\frac{m}{n}\frac{m}{k}C(\tilde U_{m,k},\tilde V_{m,k})\right)+E\left(\frac{n(n-1)}{2n^2}\frac{m^2}{k}C^2(\tilde U_{m,k},\tilde V_{m,k})\right).\end{aligned}$$

Similar to above we can show that $E((m^2/k)C^2(\tilde U_{m,k},\tilde V_{m,k}))\to 0$ and hence $ET^2_{n5}(1,1)\to 0$. By Tchebychev inequality, $T_{n5}(1,1)\xrightarrow{\mathrm{P}}0$ follows. □

Proof of Theorem 2.1. Theorem 2.1 follows immediately from Propositions 4.1 and 4.2, the continuous mapping theorem, Slutsky theorem and the fact

$$\{W_1(\theta x)+W_2(x),x\in[0,1]\}\stackrel{\mathrm{d}}{=}\{W_1((1+\theta)x),x\in[0,1]\}$$

where W_1 and W_2 are two independent Brownian motions. □

Proof of Theorem 2.2. Theorems 2.2 follows from Theorem 2.1 and the continuous mapping theorem. □

Proof of Theorem 2.3. Theorem 2.3 follows from Theorem 2.1 and the continuous mapping theorem. □

Acknowledgements

The authors thank the associate editor and the referee for a valuable comment that made the proof very brief. Part of the second author's work was done when he worked at the University of Bern, supported by a grant of the Swiss National Science foundation.

References

Csörgő, M., Horváth, L., 1993. Weighted Approximations in Probability and Statistics. Wiley, New York.
Deheuvels, P., Martynov, G., 1996. Cramér–von Mises-type tests with applications to tests of independence for multivariate extreme-value distributions. Comm. Statist. Theory Methods 25, 871–908.
Dietrich, D., de Haan, L., Hüsler, J., 2002. Testing extreme value conditions. Extremes 5, 71–85.
Draisma, G., Drees, H., Ferreira, A., de Haan, L., 2004. Bivariate tail estimation: dependence in asymptotic independence. Bernoulli 10, 251–280.
Drees, H., Huang, X., 1998. Best attainable rates of convergence for estimators of the stable tail dependence function. J. Multivariate Anal. 64, 25–47.
Drees, H., de Haan, L., Li, D., 2006. Approximations to the tail empirical distribution function with application to testing extreme value conditions. J. Statist. Plann. Inference 136, 3498–3538.
Einmahl, J.H.J., 1997. Poisson and Gaussian approximation of weighted local empirical processes. Stochastics Process. Appl. 70, 31–58.
Einmahl, J., de Haan, L., Li, D., 2006. Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. Ann. Statist. 34, 1987–2014.
Huang, X., 1992. Statistics of bivariate extremes, series no. 22. Thesis. Erasmus University Rotterdam. Tinbergen Institute.
Hüsler, J., Li, D., 2006. On testing extreme value condition. Extremes 9, 69–86.
Ledford, A., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83, 169–187.
Ledford, A., Tawn, J.A., 1997. Modelling dependence within joint tail regions. J. Roy. Statist. Soc. Ser. B 59, 475–499.
Ledford, A., Tawn, J.A., 1998. Concomitant tail behaviour for extremes. Adv. in Appl. Probab. 30, 197–215.
Schlather, M., 2001. Examples for the coefficient of tail dependence and the domain of attraction of a bivariate extreme value distribution. Statist. Probab. Lett. 53, 325–329.