INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2002-06

Weak Convergence of Empirical Copula Processes

J. D. FERMANIAN[1]
D. RADULOVIC[2]
M. WEGKAMP[3]

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

[1] CREST-ENSAE? France, email : fermania@ensae.fr
[2] Yale University, USA.
[3] Yale University, USA.

WEAK CONVERGENCE OF EMPIRICAL COPULA PROCESSES

JEAN-DAVID FERMANIAN (ENSAE-CREST, FRANCE)
AND DRAGAN RADULOVIĆ, MARTEN WEGKAMP (YALE UNIVERSITY, USA)

Abstract : Weak convergence of the empirical copula process has been established in the case of independent marginal distributions (Deheuvels, 1979, 1981). Van der Vaart and Wellner (1996) utilize the functional delta method to show convergence in $\ell^\infty([a,b]^2)$ for some $0 < a < b < 1$, under restrictions on the distribution functions.

We extend their results by proving the weak convergence of this process in $\ell^\infty([0,1]^2)$ under minimal conditions on the copula function. In addition, we consider smoothed versions of the empirical copula process, and show that they tend weakly towards the same Gaussian limit. Some applications to semi-parametric models are considered as well.

Résumé : La convergence faible du processus des copules empiriques a été établie par P. Deheuvels dans le cas de distributions marginales indépendantes (Deheuvels, 1979, 1981). Van der Vaart et Wellner (1996) utilisent la delta méthode pour montrer la convergence dans $\ell^\infty([a,b]^2)$ pour tous réels $0 < a < b < 1$, sous des conditions de régularité.

Nous étendons ces résultats en prouvant la convergence faible de ce processus dans $\ell^\infty([0,1]^2)$ sous des hypothèses minimales concernant la copule elle-même. De plus, nous étudions des versions lissées du processus des copules empiriques et nous montrons qu'elles tendent faiblement vers le même processus limite gaussien. Des applications aux modèles semi-paramétriques sont également considérées.

1. Introduction

It is well-known (cf., *e.g.*, Nelsen (1999)) that every multivariate cumulative distribution function (cdf.) H on $\mathbb{R}^p$ can be put in the form

$$H(x_1, \dots, x_p) = C(F_1(x_1), \dots, F_p(x_p)), \tag{1.1}$$

where $F_1, \cdots, F_p$ denote the marginal cumulative distribution functions. The function C is called the *copula* or *dependence function* associated to H, and in itself is a distribution function on $[0,1]^p$ with uniform margins. The representation (1.1) is unique on the range of $(F_1, \dots, F_p)$, a result due to Sklar (1959). For some historic notes, we refer to Schweizer (1991) and the recent surveys by Joe (1997) and Nelsen (1999).

Copulas capture the dependence structure among the components X_j of the random vector $(X_1, \cdots, X_p)$, irrespectively of their marginal distributions F_j. In fact, Lemma 3 below asserts that we may assume without loss of generality the X_j to be uniformly distributed on [0,1]. In other words, copulas allow us to model separately the marginal distributions and the dependence structure, and because of this ability, they have been rediscovered recently, *e.g.*, to study the joint probability to default of several borrowers in finance and actuarial sciences, or more generally some correlated extreme events. See, *e.g.*, Schweizer and Sklar (1974), Genest and McKay (1986a, 1986b), Genest and Rivest (1993), Genest *et al.* (1995), Capéraà *et al.* (1997), Bouyé *et al.* (2000), Schönbucher and Schubert (2001), and Embrechts *et al.* (2002).

In order to simplify our notation and exposition, we will consider only two-dimensional copulas in this paper ($p = 2$). The general case, however, can be easily deduced from our results. Let (X, Y) be a bivariate random vector with joint cdf. H and marginal cdf.'s F and G. Its associated copula C is defined by, for every real numbers x and y,

$$H(x, y) = C(F(x), G(y)). \tag{1.2}$$

If F and G are continuous, then the copula C satisfying (1.2) is unique, and we may write

$$C(u, v) = H(F^-(u), G^-(v)), \quad 0 \le u, v \le 1 \tag{1.3}$$

where F^- and G^- are the *generalized* quantile functions of F and G, respectively. Recall that the generalized inverse of a cdf. F is defined as

$$F^-(u) = \inf\{t \in \mathbb{R} \mid F(t) \ge u\} \text{ for } 0 \le u \le 1.$$

Given a sample of independent pairs $(X_1, Y_1), \cdots, (X_n, Y_n)$ with the same distribution H, there are various ways to estimate the associated copula C of H. They all follow the same methodology: firstly, estimate H and its margins F and G by $\widehat{H}$, $\widehat{F}$ and $\widehat{G}$, respectively. Secondly, compute the (generalized) quantile functions

$$\widehat{F}^{-}(u) = \inf\{t \in \mathbb{R} \mid \widehat{F}(t) \geq u\} \quad \text{and} \quad \widehat{G}^{-}(v) = \inf\{t \in \mathbb{R} \mid \widehat{G}(t) \geq v\} \text{ for } 0 \leq u, v \leq 1.$$

Thirdly, invoking identity (1.3), set

$$\widehat{C}(u,v) = \widehat{H}(\widehat{F}^{-}(u), \widehat{G}^{-}(v)).$$

Throughout this paper we will make the blanket assumption that H has continuous marginals.

We will consider mainly three variations on this methodology. The main goal is to state weak convergence of the process $\sqrt{n}(\widehat{C} - C)$ in a convenient function space. Section 2 considers the empirical copula process based on the ordinary empirical distribution H_n, which puts mass $1/n$ at each observation (X_i, Y_i), and its marginal distributions F_n and G_n. We extend the result obtained by Van der Vaart and Wellner (1996) by showing that weak convergence holds true in a larger space under weaker assumptions. In particular, we only impose restrictions on the underlying copula function, rather than differentiability assumptions on H and its marginals. We show that the needed regularity on C, to wit, C has continuous partial derivatives, cannot be dispensed with. Statistical applications in hypothesis testing for independence, asymptotic normality of rank statistics, and the bootstrap are provided.

Section 3 deals with the smoothed empirical copula process that is obtained by taking kernel estimates for $\widehat{H}, \widehat{F}$ and $\widehat{G}$ in lieu of the ordinary empirical cdf.'s considered in Section 2.

In Section 4 the margins F and G are assumed to be known up to finite dimensional parameters, but the copula function is entirely unknown to us. The converse situation where C is assumed to be known up to a finite parameter, and the distributions are unknown and estimated by the empirical cdf.'s has been studied by, *e.g.*, Genest *et al.* (1995), Klaassen and Wellner (1997) and Genest and Werker (2000).

In the Appendix, we prove a uniform law of large numbers for the empirical copula process indexed by a class of functions.

2. Weak convergence of empirical copula processes

Let (X, Y) be a pair of random variables with distribution function H and continuous marginals F and G. Based on independent copies $(X_1, Y_1), \cdots, (X_n, Y_n)$, we construct the empirical distribution function

$$H_n(x,y) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i \le x, Y_i \le y\}},\ x,y \in \mathbb{R},$$

and let $F_n(x)$ and $G_n(y)$ be its associated marginal distributions, that is,

$$F_n(x) = H_n(x,+\infty) \text{ and } G_n(y) = H_n(+\infty, y),\ x,y \in \mathbb{R}.$$

We define the *empirical* copula function C_n by

$$C_n\,(u,v) = H_n\left(F_n^-(u), G_n^-(v)\right), \tag{2.1}$$

where, for every univariate cdf. F, we define the generalized quantile function F^- as usual by

$$F^-(u) = \inf\{t : F(t) \ge u\},\ 0 \le u \le 1.$$

The function C_n has been briefly discussed by Frits Ruymgaart in the introduction of his doctoral thesis (Ruymgaart (1973), pp. 6 – 13). Paul Deheuvels investigated the consistency of C_n (*cf.* Deheuvels (1979)), and he obtained the exact law and the limiting process of $\sqrt{n}(C_n - C)(u,v)$ when the two margins are independent (*cf.* Deheuvels (1981a and 1981b)).

Notice that our definition slightly differs from the one proposed by Genest *et al.*(1995) who define

$$\overline{C}_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{F_n(X_i) \le u, G_n(Y_i) \le v\}},\ u,v \in [0,1]. \tag{2.2}$$

It is easily seen that C_n and $\overline{C}_n$ coincide on the grid $\{(i/n, j/n),\ 1 \le i,j \le n\}$. The subtle difference lies in the fact that C_n is left-continuous with right-hand limits, whereas $\overline{C}_n$ on the other hand is right-continuous with left-hand limits. The difference between C_n and $\overline{C}_n$, however, is small:

$$\sup_{0 \le u,v \le 1} \left|C_n(u,v) - \overline{C}_n(u,v)\right| \le \max_{1 \le i,j \le n} \left|C_n\left(\frac{i}{n}, \frac{j}{n}\right) - C_n\left(\frac{i-1}{n}, \frac{j-1}{n}\right)\right| \le \frac{2}{n}. \tag{2.3}$$

As a consequence, all weak convergence results established for C_n hold for $\overline{C}_n$ as well.

By an elegant application of the functional delta-method, Van der Vaart and Wellner (1996, p.389) proved the weak convergence of the (ordinary) *empirical copula process*

$$Z_n(u,v) \equiv \sqrt{n}(C_n - C)(u,v),\ 0 \le u,v \le 1$$

to a Gaussian process in $\ell^\infty([a,b]^2)$ when $a > 0$ and $b < 1$. More precisely,

Theorem 1. *Let $0 < a < b < 1$. Suppose that $H(x,y)$ has some marginal distribution functions $F(x)$ and $G(y)$ that are continuously differentiable on the intervals $[F^{-}(a)-\varepsilon, F^{-}(b)+\varepsilon]$ and $[G^{-}(a)-\varepsilon, G^{-}(b)+\varepsilon]$ with positive derivatives for some $\varepsilon > 0$. Furthermore, assume that $H(x,y)$ is continuously differentiable on the product of these intervals. Then the copula process $\{\sqrt{n}(C_n - C)(x,y),\ a \leq x, y \leq b\}$ converges in distribution in $\ell^\infty([a,b]^2)$ to a tight Gaussian process $\{\mathbb{G}_C(x,y),\ a \leq x,y \leq b\}$.*

The next result is another application of the delta-method:

Theorem 2. *Let $H(x,y)$ have compact support $[0,1]^2$, and marginal distributions $F(x)$ and $G(y)$ that are continuously differentiable on its support with strictly positive densities $f(x)$ and $g(y)$, respectively. Furthermore, assume that $H(x,y)$ is continuously differentiable on $[0,1]^2$. Then the copula process $\{\sqrt{n}(C_n - C)(x,y),\ 0 \leq x,y \leq 1\}$ converges in distribution in $\ell^\infty([0,1]^2)$ to a tight Gaussian process $\{\mathbb{G}_C(x,y),\ 0 \leq x,y \leq 1\}$.*

Proof. As in Van der Vaart and Wellner (1996), page 389, we observe that mapping H into its copula function can be decomposed as

$$H \mapsto (H,F,G) \mapsto (H,F^{-},G^{-}) \mapsto H \circ (F^{-},G^{-}).$$

The first map and the third map are Hadamard differentiable as pointed out in the proof of Lemma 3.9.28 in Van der Vaart and Wellner (1996). The second map is Hadamard differentiable as a consequence of Lemma 3.9.23, page 386 in Van der Vaart and Wellner (1996), which states that the inverse mapping $F \mapsto F^{-}$ as a mapping $D_2 \subset D[0,1] \mapsto \ell^\infty[0,1]$ is Hadamard differentiable at F, tangentially to $C[0,1]$. Here D_2 is the collection of distribution functions of measures that concentrate on $[0,1]$. Apply the chain rule to show that $H \mapsto H \circ (F^{-},G^{-})$ is Hadamard differentiable, followed by the delta-method, Theorem 3.9.4 in Van der Vaart and Wellner (1996). $\square$

Both results are unsatisfactory in two respects. Firstly, the weak convergence is in $\ell^\infty([a,b]^2)$ rather than $\ell^\infty([0,1]^2)$, unless $H(x,y)$ has compact support. Secondly, $H(x,y)$, $F(x)$ and $G(y)$ are assumed to have continuous derivatives. This is needed to establish compact differentiability of the inverse map $F \mapsto F^{-}$.

A more direct proof of the weak convergence of the empirical copula process along the lines of Theorem 10 in Section 3 below, shows that actually $\{Z_n(x,y),\ 0 \leq x,y \leq 1\}$ converges weakly to a Gaussian process in $\ell^\infty([0,1]^2)$, provided $C(x,y)$ has continuous partial derivatives only. We do not follow this route since Lemma 3 below allows us to obtain the stronger result using Theorem 2 directly.

We first introduce some more notation. Define the pseudo variables

$$(X^*, Y^*) = (F(X), G(Y)),$$

with distribution function

$$H^*(x, y) = \mathbb{P}\{X_1^* \le x, Y_1^* \le y\} = H(F^-(x), G^-(y)),$$

and marginal cdf.'s $F^*(x) = H^*(x, +\infty)$ and $G^*(y) = H^*(+\infty, y)$. Notice that $F^*(x)$ and $G^*(y)$ are both uniform distributions on $[0, 1]$. The copula function associated to $H^*(x, y)$ is denoted by $C^*(u, v) = H^*(F^{*-}u, G^{*-}v)$ for $0 \le u, v \le 1$. Finally, let $H_n^*(x, y)$ be the empirical distribution function based on $(X_1^*, Y_1^*), \ldots, (X_n^*, Y_n^*)$ with marginal distributions $F_n^*(x) = H_n^*(x, +\infty)$ and $G_n^*(y) = H_n^*(+\infty, y)$, and let $C_n^*(x, y)$ be its associated empirical copula function.

Lemma 3. *We have*

$$C(x, y) = C^*(x, y) = H^*(x, y) \text{ for all } x, y \in [0, 1].$$

Moreover,

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = C_n^*\left(\frac{i}{n}, \frac{j}{n}\right) \text{ for } i, j = 0, 1, \ldots, n.$$

Remark: The first assertion is well-known. The fact that C_n and C_n^* agree on the grid points $(i/n, j/n)$ is more surprising.

Proof. The first claim follows from

$$\begin{aligned} C(x, y) &= H(F^- x, G^- y) \\ &\quad \text{by definition of } C \\ &= H^*(x, y) \\ &\quad \text{by definition of } H^* \\ &= C^*(x, y) \\ &\quad \text{since } F^* \text{ and } G^* \text{ are uniform.} \end{aligned}$$

We prove the second display by the following reasoning. Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics of the sample $X_1, \cdots, X_n$. Similarly, $Y_{(1)} < \cdots < Y_{(n)}$ are the order statistics of the sample $Y_1, \cdots, Y_n$. Define $X_{(0)} = Y_{(0)} = -\infty$ and $X_{(n+1)} = Y_{(n+1)} = +\infty$, and set

$i_n = i/n$ and $j_n = j/n$. Hence

$$\begin{aligned}
C_n(i_n, j_n) &= H_n(X_{(i)}, Y_{(j)}) \\
&\quad \text{as } F_n(X_{(i)}) = i_n \text{ and } G_n(Y_{(j)}) = j_n \\
&= H_n(F^- F(X_{(i)}), G^- G(Y_{(j)})) \\
&= H_n^*(F(X_{(i)}), G(Y_{(j)})) \\
&\quad \text{since } H_n^*(x,y) = H_n(F^- x, G^- y) \\
&= H_n^*(X_{(i)}^*, Y_{(j)}^*) \\
&\quad \text{where } X_{(i)}^* = F(X_{(i)}) \text{ and } Y_{(j)}^* = G(Y_{(j)}) \\
&= C_n^*(i_n, j_n) \\
&\quad \text{since } F_n^*(X_{(i)}^*) = i_n \text{ and } G_n^*(Y_{(j)}^*) = j_n.
\end{aligned}$$

This conclues the proof of the lemma. □

Combination of Theorem 2 and Lemma 3 immediately yields

Theorem 4. *Let the copula function $C(x,y)$ have continuous partial derivatives. Then the empirical copula process converges weakly to the Gaussian process $\mathbb{G}_C$ in $\ell^\infty([0,1]^2)$.*

Proof. First notice that for all $x, y \in [0,1]$, there exist i_n, j_n such that $C_n(x,y) = C_n(i_n, j_n)$, which coupled with Lemma 3 yields $\sqrt{n}(C_n - C)(x,y) = \sqrt{n}(C_n^* - C^*)(x,y)$. Since $H^*(x,y) = C(x,y)$ satisfies the conditions of Theorem 2, invoke Theorem 2 to conclude the proof. □

The limiting Gaussian process can be written as

$$\mathbb{G}_C(u,v) = \mathbb{B}_C(u,v) - \partial_1 C(u,v)\mathbb{B}_C(u,1) - \partial_2 C(u,v)\mathbb{B}_C(1,v),$$

where $\mathbb{B}_C$ is a Brownian bridge on $[0,1]^2$ with covariance function

$$\mathbb{E}\left[\mathbb{B}_C(u,v) \cdot \mathbb{B}_C(u',v')\right] = C(u \wedge u', v \wedge v') - C(u,v)C(u',v')$$

for each $0 \le u, u', v, v' \le 1$.

Let $\overline{C}_n$ be the *càdlàg* version of the empirical copula function defined in (2.2). Theorem 4 also implies that the related process $\overline{Z}_n \equiv \sqrt{n}(\overline{C}_n - C)$ converges weakly to $\mathbb{G}_C$, provided C has continuous partial derivatives. Invoke inequality (2.3) to see that $\overline{Z}_n$ converges weakly if and only if Z_n converges weakly.

As already mentioned – since the smoothness of derivatives of F, G and H easily implies the smooth partial derivatives of C – Theorem 4 offers another improvement of Theorem

1 besides the extension from the space $\ell^\infty([a,b]^2)$ to $\ell^\infty([0,1]^2)$. We argue that this is not merely a cosmetic improvement. Copulas are designed in order to capture the dependency structure of the vector (X,Y) independently of the marginal distributions, which makes it desirable to state the results with the assumptions on C rather than on the marginal distributions. Lemma 3 supports this approach as well. For example, if X and Y are independent, $C(s,t) = st$ and this copula has clearly continuous partial derivatives, regardless of degree of smoothness of marginal distributions. Numerous other statistically interesting cases follow this pattern, where we are interested in the estimation or in testing the hypothesis of a particular smooth copula function, and where we are not concerned with the smoothness of marginal distributions. See, for example, Nelsen (1999), Chapter 3.

Regarding the assumption on C, we note that every copula C is Lipschitz and its partial derivatives exist for almost all points in $[0,1]^2$ (*cf.*, *e.g.*, Nelsen (1999)). Careful inspection of the proof for Theorem 4 reveals that we require smoothness of the partial derivatives only in order to apply the functional delta method (Theorems 1 and 2). This observation, coupled with the above example for independent X and Y, suggests that one might perhaps be able to relax this assumption. That would be useful, since there are many statistically relevant cases where the desired copula function does not have continuous partial derivatives. For example, let X be any (continuous) symmetric random variable and let vector $(X,Y) = (X,-X)$. In this case $C(s,t) = \max(0, s+t-1)$ does not have continuous partial derivatives. Unfortunately, the following result, which in a sense is the converse of Theorem 4, indicates that there is actually very little we can do.

Theorem 5. *Assume that the inverses F^{-1} and G^{-1} exist and that there exists at least one point $(s^*,t^*) \in (0,1)^2$ for which the four quantities*

$$A_1 \equiv \lim_{h \nearrow 0} \frac{C(s^*+h,t^*) - C(s^*,t^*)}{h}, \quad A_2 \equiv \lim_{h \searrow 0} \frac{C(s^*+h,t^*) - C(s^*,t^*)}{h},$$
$$\overline{A}_1 \equiv \lim_{h \nearrow 0} \frac{C(s^*,t^*+h) - C(s^*,t^*)}{h}, \quad \overline{A}_2 \equiv \lim_{h \searrow 0} \frac{C(s^*,t^*+h) - C(s^*,t^*)}{h}$$

are not all equal. Then $\sqrt{n}(C_n(s,t) - C(s,t))$, $s,t \in [0,1]^2$ does not converge to a tight Gaussian process.

Proof. We have

$$
\begin{aligned}
&Z_n(s,t)\\
&= \sqrt{n}(C_n(s,t) - C(s,t))\\
&= \sqrt{n}\left[H_n(F_n^-(s), G_n^-(t)) - H(F^{-1}(s), G^{-1}(t))\right]\\
&= \sqrt{n}\left[(H_n - H)(F_n^-(s), G_n^-(t)) - (H_n - H)(F^{-1}(s), G^{-1}(t))\right] +\\
&\quad + \sqrt{n}[H(F_n^-(s), G_n^-(t)) - H(F^{-1}(s), G^{-1}(t))] + \sqrt{n}(H_n - H)(F^{-1}(s), G^{-1}(t)).
\end{aligned}
$$

The first term on the right is $o_P(1)$ since the process $\sqrt{n}(H_n - H)$ is stochastically equicontinuous. The third part converges to a normal random variable for every fixed (s,t). Evaluated at the point (s^*, t^*) the second part of the above equation is equal to

$$
\begin{aligned}
&\sqrt{n}[C(F \circ F_n^-(s^*), G \circ G_n^-(t^*)) - C(s^*, t^*)]\\
&= A_1\sqrt{n}(F \circ F_n^-(s^*) - s^*)\mathbb{I}_{\left\{F_n^-(s^*) < F^{-1}(s^*)\right\}} +\\
&\quad + A_2\sqrt{n}(F \circ F_n^-(s^*) - s^*)\mathbb{I}_{\left\{F_n^-(s^*) > F^{-1}(s^*)\right\}} +\\
&\quad + \overline{A}_1\sqrt{n}(G \circ G_n^-(t^*) - t^*)\mathbb{I}_{\left\{G_n^-(t^*) < G^{-1}(t^*)\right\}} +\\
&\qquad + \overline{A}_2\sqrt{n}(G \circ G_n^-(t^*) - t^*)\mathbb{I}_{\left\{G_n^-(t^*) > G^{-1}(t^*)\right\}} + o_P(1).
\end{aligned}
$$

Thus, the asymptotic behavior of the previous term depends on the values of the left - and right hand limits. If one of the four constants $A_1, A_2, \overline{A}_1$ or $\overline{A}_2$ differs from the others, the above expression does not converge for $n \to \infty$. Notice that the events in the indicator functions occur with a nonzero probability, and as a result the limiting process (if it exists) is not Gaussian. Moreover, a very similar argument reveals that

$$
\begin{aligned}
&Z_n(s^* + \delta, t^*) - Z_n(s^* - \delta, t^*)\\
&= A_1\xi_1^\delta\mathbb{I}_{\left\{F_n^-(s^*+\delta) < F^{-1}(s^*+\delta)\right\}} + A_2\xi_1^\delta\mathbb{I}_{\left\{F_n^-(s^*+\delta) > F^{-1}(s^*+\delta)\right\}} +\\
&\quad - A_1\xi_2^\delta\mathbb{I}_{\left\{F_n^-(s^*-\delta) < F^{-1}(s^*-\delta)\right\}} - A_2\xi_2^\delta\mathbb{I}_{\left\{F_n^-(s^*-\delta) > F^{-1}(s^*-\delta)\right\}} + o_P(1),
\end{aligned}
$$

where

$$
\xi_1^\delta = \lim_{n\to\infty}\sqrt{n}(F_n(s^* + \delta) - F(s^* + \delta)) \text{ and } \xi_2^\delta = \lim_{n\to\infty}\sqrt{n}(F_n(s^* - \delta) - F(s^* - \delta)).
$$

Assume that $A_1 \neq A_2$. Then the right hand side of the above equation does not converge to 0 as $\delta \to 0$ and $n \to \infty$ because

$$
\lim_{\delta \searrow 0}\liminf_{n\to\infty} \mathbb{P}\left\{F_n^-(s^* + \delta) < F^{-1}(s^* + \delta); F_n^-(s^* - \delta) > F^{-1}(s^* - \delta)\right\} > 0.
$$

This implies that the process $Z_n(s,t)$ is not stochastically equicontinuous, which establishes the claim. $\square$

The covariance structure of $\sqrt{n}(C_n - C)(u,v)$ might be complicated to estimate, and the bootstrap methodology provides an attractive alternative to estimate the finite sample distribution of C_n. We will show that the bootstrap "works", but first we need some additional notation. Let $(X_{1,B}, Y_{1,B}), \cdots, (X_{n,B}, Y_{n,B})$ be the bootstrap sample obtained by sampling with replacement from the original observations $(X_1, Y_1), \cdots, (X_n, Y_n)$. We write $H_{n,B}(x,y)$ for the empirical cdf. based on the bootstrap sample, and denote its associated empirical copula function by $C_{n,B}$.

Theorem 6. *The conditional distribution of* $\{\sqrt{n}(C_{n,B} - C_n)(x,y),\ 0 \le x, y \le 1\}$ *converges to the same limiting Gaussian process of* $\{\sqrt{n}(C_n - C)(x,y),\ 0 \le x, y \le 1\}$ *in* $\ell^\infty([0,1]^2)$ *in probability.*

Proof. We can invoke the same uniform transformation trick of Lemma 3. We already know that $C_n(i_n, j_n) = C_n^*(i_n, j_n)$, and it is easily verified that $C_{n,B}(i_n, j_n) = C_{n,B}^*(i_n, j_n)$ as well, where $C_{n,B}^*$ is the empirical copula function based on $(F(X_{i,B}), G(Y_{i,B}))$. Hence

$$\sqrt{n}(C_{n,B} - C_n)(i_n, j_n) = \sqrt{n}(C_{n,B}^* - C_n^*)(i_n, j_n).$$

The conclusion follows by observing that the proof of Theorem 2 is based on an application of the delta method, by showing that the map $H \mapsto C_H$ is Hadamard differentiable, and hence $\sqrt{n}(\phi(H_{n,B}) - \phi(H_n))$ converges weakly if and only if $\sqrt{n}(\phi(H_n) - \phi(H))$ is weakly convergent by Theorem 3.9.11 in Van der Vaart and Wellner (1996, page 378). $\square$

We mention some consequences of our results. Deheuvels (1981) proposed among other related procedures the Kolmogorov-Smirnov type statistic

$$T \equiv \sup_{0 \le s,t \le 1} |\sqrt{n}(C_n - C)(s,t)|$$

for testing the independence hypothesis $H_0 : C(s,t) = s \cdot t$. He calculated the distribution of T under this null hypothesis. The results established here are useful to compute the power of this test under various alternatives.

In the early 1970's there was quite some interest in multivariate rank order statistics (see, for example, Ruymgaart, Shorack and Van Zwet (1972), Ruymgaart (1974), and Rüschendorf (1976)). Such statistics are of the form

$$R_n = \frac{1}{n} \sum_{i=1}^{n} J(F_n(X_i), G_n(Y_i)),$$

and asymptotic normality of R_n has been established under regularity assumptions on $J : [0,1]^2 \to \mathbb{R}$. However, simply by observing that

$$\frac{1}{n}\sum_{i=1}^{n} J(F_n(X_i), G_n(Y_i)) = \int_{[0,1]^2} J(u,v)\, d\overline{C}_n(u,v), \tag{2.4}$$

where $\overline{C}_n$ is the *càdlàg* version of the empirical copula function defined in (2.2), and

$$\mathbb{E}J(F(X), G(Y)) = \int_{[0,1]^2} J(u,v)\, dC(u,v),$$

it follows immediately from Theorem 4 and the functional delta method that

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{J(F_n(X_i), G_n(Y_i)) - \mathbb{E}J(F(X_i), G(Y_i))\}\\
&= \sqrt{n}\int_{[0,1]^2} J(u,v)\, d(\overline{C}_n - C)(u,v)\\
&= \int_{[0,1]^2} \sqrt{n}(\overline{C}_n - C)(u,v)\, dJ(u,v)\\
&\qquad \text{using integration by parts}\\
&\xrightarrow{D} \int_{[0,1]^2} \mathbb{G}_C(u,v)\, dJ(u,v)
\end{aligned}$$

for every function J of bounded variation, provided C has continuous partial derivatives. Since a continuous, linear transformation of a tight Gaussian process is normally distributed, the limit has a Gaussian distribution and we reach the same conclusion under weak assumptions on J. It is interesting to note that Ruymgaart (1973) realized (2.4), more precisely the relation between C_n (in lieu of $\overline{C}_n$) and R_n and wrote (Ruymgaart (1973, page 7)): *"It would be interesting to see whether results comparable to those of Chapters 2 – 4 can be obtained with [(2.4)] as a starting point. One would have to study the weak convergence of the suitably standardized processes [C_n] and the rate of growth of these processes near the boundary of the unit square."*

This result is useful in many cases. For instance, Genest *et al.* (1995) consider a family of copula densities c_α indexed by α, and propose to estimate α by solving $L_n(\alpha) = 0$, where

$$L_n(\alpha) \equiv \frac{1}{n}\sum_{i=1}^{n} \partial_\alpha \log c_\alpha(F_n(X_i), G_n(Y_i)) = \int_{[0,1]^2} \partial_\alpha \log c_\alpha(x,y)\, d\overline{C}_n(x,y).$$

Their analysis requires that $L_n(\alpha)$ is asymptotically normal, which follows from the discussion above. In the Appendix, we propose regularity conditions on J ensuring a uniform law of large numbers for R_n.

3. Weak convergence of smoothed empirical copula processes

The smoothed empirical distribution function $\widehat{H}_n(x,y)$ is defined by

$$\widehat{H}_n(x,y) = \frac{1}{n}\sum_{i=1}^n K_n(x - X_i, y - Y_i).$$

Here $K_n(x,y) = K(a_n^{-1}x, a_n^{-1}y)$, and

$$K(x,y) = \int_{-\infty}^x \int_{-\infty}^y k(u,v)\,du\,dv, \tag{3.1}$$

for some bivariate kernel function $k : \mathbb{R}^2 \mapsto \mathbb{R}$, with $\int k(x,y)\,dxdy = 1$, and a sequence of bandwidths $a_n \downarrow 0$ as $n \to \infty$. For notational convenience, we have chosen the same bandwidth sequence for each margin. This assumption can be dropped easily.

For small enough bandwidths a_n, the empirical cdf. H_n and the smoothed empirical cdf. $\widehat{H}_n$ are almost indistinguishable:

Lemma 7. *Assume that F and G are Lipschitz, $a_n \to 0$,*

$$\int_{\mathbb{R}^2} (|x| + |y|)\,dK(x,y) < \infty \text{ and } \sup_{x,y} \sqrt{n}\left|\mathbb{E}\widehat{H}_n(x,y) - H(x,y)\right| \to 0,$$

then

$$\sqrt{n}\sup_{x,y}|\widehat{H}_n(x,y) - H_n(x,y)| \xrightarrow{P} 0,$$

and in particular, the smoothed empirical process $\left\{\sqrt{n}(\widehat{H}_n - H)(x,y),\ x,y \in \mathbb{R}\right\}$ converges weakly to a tight Brownian bridge in $D(\mathbb{R}^2)$.

Proof. According to Van der Vaart (1994), we only have to check that

$$\sup_{s,t} \int \left[\int \mathbb{I}_{\{x+\varepsilon \le s, y+\eta \le t\}}\,dK_n(\varepsilon,\eta) - \mathbb{I}_{\{x\le s, y\le t\}}\right]^2 dH(x,y) \to 0.$$

This is the case, since

$$\begin{aligned}
&\sup_{s,t} \int \left[\int \left(\mathbb{I}_{\{x+\varepsilon \le s, y+\eta \le t\}} - \mathbb{I}_{\{x\le s, y\le t\}}\right) dK_n(\varepsilon,\eta)\right]^2 dH(x,y) \\
&\le \sup_{s,t} \int\int \left(\mathbb{I}_{\{x+\varepsilon \le s, y+\eta \le t\}} - \mathbb{I}_{\{x\le s, y\le t\}}\right)^2 dK_n(\varepsilon,\eta)\,dH(x,y) \\
&\text{by Jensen's inequality} \\
&\le \sup_{s,t} \int\int \left(\mathbb{I}_{\{x\in(s-\varepsilon,s]\}} + \mathbb{I}_{\{y\in(t-\eta,t]\}}\right) dH(x,y)\,dK_n(\varepsilon,\eta) \\
&\text{using Fubini} \\
&= \sup_{s,t} \int [F(s) - F(s - a_n x)] + [G(t) - G(t - a_n y)]\,dK(x,y)
\end{aligned}$$

which tends to zero as $a_n \to 0$ by our assumptions on F, G and K. $\square$

The bias term can be handled by means of some smoothness assumptions on H and regularity of K:

Lemma 8. *Assume that H has a bounded p-th derivative, $\lim_{n\to\infty} \sqrt{n}a_n^p = 0$,*

$$\int_{\mathbb{R}^2} x^k y^l k(x,y)dx\,dy = 0,\ 1 \le k+l < p,$$

and $\int |x|^k|y|^l|k(x,y)|\,dx\,dy < \infty$, $1 \le k+l \le p$. Then we have

$$\sup_{x,y} \sqrt{n}\left|\mathbb{E}\widehat{H}_n(x,y) - H(x,y)\right| = \sqrt{n}a_n^p.$$

Proof. The claim follows readily after a simple Taylor expansion. $\square$

Next we study the weak convergence of the smoothed empirical copula process

$$\widehat{Z}_n(x,y) = \sqrt{n}(\widehat{C}_n - C)(x,y),\ 0 \le x, y \le 1,$$

based on the smoothed empirical copula function

$$\widehat{C}_n(x,y) = \widehat{H}_n\left(\widehat{F}_n^-(x), \widehat{G}_n^-(y)\right).$$

The following lemma establishes asymptotic tightness of the smoothed empirical process.

Lemma 9. *Let $C(x,y)$ have continuous partial derivatives and assume that the assumptions of Lemma 7 hold. Then the process $\{\widehat{Z}_n(x,y) :\ (x,y) \in [0,1]^2\}$ is stochastically equicontinuous, that is, for all $\eta > 0$*

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \mathbb{P}\left\{\sup_{|x-x'|\le\delta, |y-y'|\le\delta} \left|\widehat{Z}_n(x,y) - \widehat{Z}_n(x',y')\right| > \eta\right\} = 0.$$

Proof. Observe that

$$\mathbb{P}\left\{\sup_{|u-u'|\le\delta, |v-v'|\le\delta} |\widehat{Z}_n(u,v) - \widehat{Z}_n(u',v')| > \eta\right\} \le I + II,$$

with

$$
\begin{aligned}
I &= \mathbb{P}\Bigg\{ \sup_{|u-u'|\le\delta, |v-v'|\le\delta} \Big| \sqrt{n}\left(\widehat{C}_n(\ \widehat{F}_n(\widehat{F}_n^- u), \widehat{G}_n(\widehat{G}_n^- v)\ ) - C(\ F(\widehat{F}_n^- u), G(\widehat{G}_n^- v)\ )\ \right) \\
&\qquad -\sqrt{n}\left(\ \widehat{C}_n(\ \widehat{F}_n(\widehat{F}_n^- u'), \widehat{G}_n(\widehat{G}_n^- v')\ ) - C(\ F(\widehat{F}_n^- u', G(\widehat{G}_n^- v')\ )\ \right)\Big| > \frac{\eta}{2}\Bigg\} \\
II &= \mathbb{P}\Bigg\{ \sup_{|u-u'|\le\delta, |v-v'|\le\delta} \Big| \sqrt{n}\left(\ C(\ \widehat{F}_n(\widehat{F}_n^- u), \widehat{G}_n(\widehat{G}_n^- v)\ ) - C(\ F(\widehat{F}_n^- u), G(\widehat{G}_n^- u)\ )\ \right) \\
&\qquad -\sqrt{n}\left(\ C(\ \widehat{F}_n(\widehat{F}_n^- u'), \widehat{G}_n(\widehat{G}_n^- v')\ ) - C(\ F(\widehat{F}_n^- u'), G(\widehat{G}_n^- v')\ )\ \right)\Big| > \frac{\eta}{2}\Bigg\}.
\end{aligned}
$$

We will deal with the two terms I and II separately. The first term can be handled by noticing that

$$
\begin{aligned}
\widehat{C}_n(x,y) &= \widehat{H}_n(\widehat{F}_n^-(x), \widehat{G}_n^-(y)) \\
&= H_n(\widehat{F}_n^-(x), \widehat{G}_n^-(y)) + o_P(n^{-1/2}) \\
&\qquad \text{by Lemma 7} \\
&= H_n^*\left(F(\widehat{F}_n^-(x)), G(\widehat{G}_n^-(y))\right) + o_P(n^{-1/2})
\end{aligned}
$$

where $H_n^*(u,v)$ is the empirical cdf. based on $(F(X_1), G(Y_1)), \cdots, (F(X_n), G(Y_n))$. Therefore, the first probability can be bounded by

$$
\begin{aligned}
I &\le \mathbb{P}\Bigg\{ \sup_{(x,x'):\ |\widehat{F}_n(x)-\widehat{F}_n(x')|\le\delta,\ (y,y'):\ |\widehat{G}_n(y)-\widehat{G}_n(y')|\le\delta} \Big|\sqrt{n}\,(H_n^*-H^*)\,(F(x),G(y)) \\
&\qquad -\sqrt{n}\,(H_n^*-H^*)\,(F(x'),G(y'))\Big| > \frac{\eta}{4}\Bigg\} + o(1) \\
&\le \mathbb{P}\Bigg\{ \sup_{(x,x'):\ |F(x)-F(x')|\le 3\delta,\ (y,y'):\ |G(y)-G(y')|\le 3\delta} \Big|\sqrt{n}\,(H_n^*-H^*)\,(F(x),G(y)) \\
&\qquad -\sqrt{n}\,(H_n^*-H^*)\,(F(x'),G(y'))\Big| > \frac{\eta}{4}\Bigg\} + \\
&\qquad +\mathbb{P}\left\{\sup_x |\widehat{F}_n(x)-F(x)| > \delta\right\} + \mathbb{P}\left\{\sup_y |\widehat{G}_n(y)-G(y)| > \delta\right\} + o(1) \\
&\le \mathbb{P}\left\{ \sup_{(u,u'):\ |u-u'|\le 3\delta,\ (v,v'):\ |v-v'|\le 3\delta} \Big|\sqrt{n}\,(H_n^*-H^*)\,(u,v) - \sqrt{n}\,(H_n^*-H^*)\,(u',v')\Big| > \frac{\eta}{4}\right\} + \\
&\qquad +\mathbb{P}\left\{\sup_x |\widehat{F}_n(x)-F(x)| > \delta\right\} + \mathbb{P}\left\{\sup_y |\widehat{G}_n(y)-G(y)| > \delta\right\} + o(1)
\end{aligned}
$$

which tends to 0 as $n\to\infty$ and $\delta\downarrow 0$ from the weak convergence of the process $n^{1/2}(H_n^*-H^*)$. The second term II can be made arbitrarily small by invoking that C has continuous partial derivatives so that

$$
\begin{aligned}
&\left[C\left(u,v\right)-C\left(F\widehat{F}^{-}u,G\widehat{G}^{-}v\right)\right]-\left[C\left(u',v'\right)-C\left(F\widehat{F}^{-}u',G\widehat{G}^{-}v'\right)\right]\\
&=-C'_{u,v}\left((\widehat{F}-F)\widehat{F}^{-}u,(\widehat{G}-G)\widehat{G}^{-}v\right)+C'_{u',v'}\left((\widehat{F}-F)\widehat{F}^{-}u',(\widehat{G}-G)\widehat{G}^{-}v'\right)\\
&\quad+\mathcal{O}(\|\widehat{F}-F\|_{\infty}+\|\widehat{G}-G\|_{\infty})\\
&=C'_{u,v}\left((\widehat{F}-F)\widehat{F}^{-}u'-(\widehat{F}-F)\widehat{F}^{-}u,(\widehat{G}-G)\widehat{G}^{-}v'-(\widehat{G}-G)\widehat{G}^{-}v\right)\\
&\quad+\left[C'_{u',v'}-C'_{u,v}\right]\left((\widehat{F}-F)\widehat{F}^{-}u',(\widehat{G}-G)\widehat{G}^{-}v'\right)+\mathcal{O}(\|\widehat{F}-F\|_{\infty}+\|\widehat{G}-G\|_{\infty})\\
&=C'_{u,v}\left((\widehat{F}-F)\widehat{F}^{-}u'-(\widehat{F}-F)\widehat{F}^{-}u,(\widehat{G}-G)\widehat{G}^{-}v'-(\widehat{G}-G)\widehat{G}^{-}v\right)+\\
&\quad\mathcal{O}(\|\widehat{F}-F\|_{\infty}+\|\widehat{G}-G\|_{\infty})
\end{aligned}
$$

for $u\to u', v\to v'$. Next, observe that

$$
\begin{aligned}
&\mathbb{P}\left\{\sup_{|u-u'|\le\delta}\sqrt{n}\left|(\widehat{F}-F)\widehat{F}^{-}u'-(\widehat{F}-F)\widehat{F}^{-}u\right|>\eta\right\}\\
&=\mathbb{P}\left\{\sup_{x,x':\ |\widehat{F}x-\widehat{F}x'|\le\delta}\sqrt{n}\left|(\widehat{F}-F)x'-(\widehat{F}-F)x\right|>\eta\right\}\\
&\le\mathbb{P}\left\{\sup_{x,x':\ |Fx-Fx'|\le3\delta}\sqrt{n}\left|(\widehat{F}-F)x'-(\widehat{F}-F)x\right|>\eta\right\}+\mathbb{P}\left\{\|F-\widehat{F}\|_{\infty}>\delta\right\}\\
&\le\mathbb{P}\left\{\sup_{x,x':\ |Fx-Fx'|\le3\delta}\sqrt{n}\left|(F_n-F)x'-(F_n-F)x\right|>\frac{\eta}{2}\right\}+\\
&\quad+\mathbb{P}\left\{\|F-\widehat{F}\|_{\infty}>\delta\right\}+\mathbb{P}\left\{\sqrt{n}\|\widehat{F}_n-F_n\|_{\infty}>\frac{\eta}{4}\right\}\\
&=\mathbb{P}\left\{\sup_{|u-u'|\le3\delta}\sqrt{n}\left|(F_n^{*}-F^{*})u'-(F_n^{*}-F^{*})u\right|>\frac{\eta}{2}\right\}+\\
&\quad+\mathbb{P}\left\{\|F-\widehat{F}\|_{\infty}>\delta\right\}+\mathbb{P}\left\{\sqrt{n}\|\widehat{F}_n-F_n\|_{\infty}>\frac{\eta}{4}\right\}\\
&\to 0,\ \text{ as } n\to\infty,\delta\downarrow 0
\end{aligned}
$$

by the weak convergence of the uniform empirical process and Lemma 7. Similarly, we can show that

$$
\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathbb{P}\left\{\sup_{|v-v'|\le\delta}\left|(\widehat{G}-G)\widehat{G}^{-}v'-(\widehat{G}-G)\widehat{G}^{-}v\right|>\eta\right\}\to 0.
$$

Hence, II is asymptotically negligible as well, and the proof is complete. □

We have obtained the following result :

Theorem 10. *Under the assumptions of Lemma 7 and provided C has continuous partial derivatives, the smoothed empirical copula process $\{\widehat{Z}_n(u,v),\ 0 \le u,v \le 1\}$ converges weakly to the Gaussian process $\{\mathbb{G}_C(u,v),\ 0 \le u,v \le 1\}$ in $\ell^\infty([0,1]^2)$.*

Proof. In view of Lemma 9, we only have to show the finite dimensional convergence of the process $\{\widehat{Z}_n(u,v),\ 0 \le u,v \le 1\}$. Take $x,y \in \mathbb{R}$ arbitrarily, and set $u = F(x), v = G(y)$ and $\widehat{u} = \widehat{F}(x)$ and $\widehat{v} = \widehat{G}(y)$. Note that $\widehat{u} \xrightarrow{P} u$ and $\widehat{v} \xrightarrow{P} v$ by Lemma 7, and argue that

$$
\begin{aligned}
&\widehat{Z}_n(u,v)\\
&= \widehat{Z}_n(\widehat{u},\widehat{v}) + o_P(1)\\
&\qquad \text{since } \widehat{Z}_n \text{ is stochastically equicontinuous and } \widehat{u} \xrightarrow{P} u \text{ and } \widehat{v} \xrightarrow{P} v\\
&= \sqrt{n}(\widehat{H}_n - H)(x,y) + \sqrt{n}\left[C(u,v) - \widehat{C}(\widehat{u},\widehat{v})\right] + o_P(1)\\
&= \sqrt{n}(\widehat{H}_n - H)(x,y) + \sqrt{n}\left[(F - \widehat{F})(x)\partial_1 C(u,v) + (G - \widehat{G})(y)\partial_2 C(u,v)\right] + o_P(1)\\
&\qquad \text{since } C \text{ has continuous partial derivatives}\\
&= \sqrt{n}(H_n - H)(x,y) + \sqrt{n}\left[(F - F_n)(x)\partial_1 C(u,v) + (G - G_n)(y)\partial_2 C(u,v)\right] + o_P(1)\\
&\qquad \text{by Lemma 7}\\
&= Z_n(u,v) + o_P(1).
\end{aligned}
$$

The required finite dimensional convergence of the process follows from Theorem 4. □

Note that we could not prove the last result in the same way as for the empirical copula process Z_n. Indeed, the transformation of Lemma 3 works no longer for smoothed empirical cdf.'s. In contrast, we can repeat the same arguments leading to Theorem 10 to prove Theorem 4.

Smoothing the empirical copula function C_n itself provides another way to build smooth estimates of C. Set

$$\widetilde{C}_n(x,y) = \int K\left(\frac{x-u}{a_n}, \frac{y-v}{a_n}\right) dC_n(u,v),$$

where K is the previous integrated kernel (3.1), and $a_n \searrow 0$. The associated copula process is now

$$\widetilde{Z}_n(x,y) \equiv \sqrt{n}(\widetilde{C}_n - C)(x,y).$$

We have the following result:

Theorem 11. *If k is bounded and compactly supported and if $na_n^2 \to 0$, then*

$$\|\sqrt{n}(\widetilde{C}_n - C) - \sqrt{n}(C_n - C)\|_\infty = o_P(1).$$

In particular, provided C has continuous partial derivatives, the process $\sqrt{n}(\widetilde{C}_n - C)$ tends weakly to the Gaussian process $\mathbb{G}_C$ in $\ell^\infty([0,1]^2)$.

Proof. Observe that

$$\begin{aligned}
\widetilde{Z}_n(x,y) &= \sqrt{n}(\widetilde{C}_n - C)(x,y) \\
&= \sqrt{n}\int (C_n - C)(x - a_n u, y - a_n v)k(u,v)\,du\,dv + \\
&\qquad + \sqrt{n}\int (C(x - a_n u, y - a_n v) - C(x,y))k(u,v)\,du\,dv,
\end{aligned}$$

and

$$\begin{aligned}
&\widetilde{Z}_n(x,y) - Z_n(x,y) \\
&= \sqrt{n}\int \left[(C_n - C)(x - a_n u, y - a_n v) - (C_n - C)(x,y)\right] k(u,v)\,du\,dv + \\
&\quad + \sqrt{n}\int (C(x - a_n u, y - a_n v) - C(x,y))k(u,v)\,du\,dv.
\end{aligned}$$

Since C is Lipschitz continuous, and k is bounded and compactly supported, the second term on the right is of order $\mathcal{O}(\sqrt{n}a_n)$. Moreover, invoking the equicontinuity of the process Z_n, the first term on the right tends to zero in probability since the kernel k is compactly supported. □

4. Semi-parametric estimation procedures

The marginal distributions are often better known than the dependence function. For instance, X and Y are the asset values of two firms, being evaluated by their equity prices and their short and long term debts. Usually, these processes are assumed to be diffusions, which implies that the margins are Gaussian random variables. Therefore, it is common to model the margins in a parametric way. In contrast, the dependence structure between two firms does not seem to be obvious *a priori*. No well stated empirical findings nor theoretical models allow us to formulate a natural parametric family for the copulas. That is why it is relevant to assume that F (resp. G) belongs to some parametric family $\mathcal{F} = \{F_\theta, \theta \in \Theta\}$ (resp. $\mathcal{G} = \{G_\pi, \pi \in \Pi\}$), whilst to leave the copula C unspecified.

The parameters θ and π are estimated by $\widehat{\theta}$ and $\widehat{\pi}$, respectively, at a first stage. Let θ_0 and π_0 be the "true values" of the parameters. We then can construct the process

$$\sqrt{n}(\widehat{C}_{\widehat{\theta},\widehat{\pi}} - C) \equiv \sqrt{n}(H_n(F^-_{\widehat{\theta}}, G^-_{\widehat{\pi}}) - H(F^-, G^-)).$$

More generally, we could study the case of misspecified models: even if the marginal distributions are misspecified, *viz.*, if the true marginal cdf.'s do not belong to $\mathcal{F}$ and $\mathcal{G}$, the process could still converge. For instance, if the parameters are estimated by pseudo-maximum likelihood (Gouriéroux *et al.*. (1984)), $\hat{\theta}$ and $\hat{\pi}$ converge to some pseudo-values θ_0^* and π_0^* under suitable regularity assumptions. In such a case, the relevant process is no longer the one defined above, but rather

$$\sqrt{n}(H_n(F_{\hat{\theta}}^-, G_{\hat{\pi}}^-) - H(F_{\theta_0^*}^-, G_{\pi_0^*}^-)).$$

Our result could be extended easily in this direction. Nonetheless, since $H(F_{\theta_0^*}^-, G_{\pi_0^*}^-))$ is no longer a copula function, we exclude this case.

Theorem 12. *Assume that C has continuous partial derivatives,*

$$\|F_{\widehat{\theta}} - F\|_\infty + \|G_{\widehat{\pi}} - G\|_\infty \xrightarrow{P} 0 \tag{4.1}$$

and that the vector

$$\left(\sqrt{n}(F_{\widehat{\theta}} - F)(x), \sqrt{n}(G_{\widehat{\pi}} - G)(y), \sqrt{n}(H_n - H)(x,y) \right) \tag{4.2}$$

converges jointly to a Gaussian limit for almost all $0 \le x, y \le 1$. Then

$$\sqrt{n}\left(\widehat{C}_{\widehat{\theta},\widehat{\pi}} - C\right)(u,v),\ 0 \le u, v \le 1$$

converges weakly to a Gaussian process in $\ell^\infty([0,1]^2)$.

Proof. Observe that since

$$\widehat{C}_{\widehat{\theta},\widehat{\pi}}(x,y) = H_n(F_{\widehat{\theta}}^{-1}x, G_{\widehat{\pi}}^{-1}y) = H_n^*(FF_{\widehat{\theta}}^- x, GG_{\widehat{\pi}}^- y),$$

we can proceed exactly as in the proof of Lemma 9 and establish stochastic equicontinuity of $\sqrt{n}(\widehat{C}_{\widehat{\theta},\widehat{\pi}} - C)$. At this point we need to invoke assumption (4.1).

Next let $x, y \in \mathbb{R}$ be arbitrary, and observe that for all $u = F(x)$, $\widehat{u} \equiv F_{\widehat{\theta}}(x)$, $v = G(y)$ and $\widehat{v} \equiv G_{\widehat{\pi}}(y)$, we find following the same reasoning as in the proof of Theorem 10

$$\begin{aligned}
&\sqrt{n}(\widehat{C}_{\widehat{\theta},\widehat{\pi}} - C)(u,v) \\
&= \sqrt{n}(H_n - H)(x,y) + \sqrt{n}\left[(F - F_{\widehat{\theta}})(x)\partial_1 C(u,v) + (G - G_{\widehat{\pi}})(y)\partial_2 C(u,v)\right] + o_P(1).
\end{aligned}$$

The finite dimensional convergence of the process is an immediate consequence of (4.2). $\square$

The delta method could be used to show that both $\sqrt{n}(F_{\widehat{\theta}} - F)$ and $\sqrt{n}(G_{\widehat{\pi}} - G)$ converge to a Gaussian limit. Clearly if the mapping $\theta \mapsto F_\theta$ is continuously differentiable for all x, asymptotic normality of $\sqrt{n}(\widehat{\theta} - \theta_0)$ carries over to $\sqrt{n}(F_{\widehat{\theta}} - F)(x)$. Alternatively, we can assume *Hellinger* differentiability (see, *e.g.*, LeCam 1986, Chapter 17.3) of the corresponding densities instead, and as a byproduct obtain (4.1) as well.

Lemma 13. *Let $\{p_\theta : \theta \in \Theta \subseteq \mathbb{R}^k\}$ be a family of probability densities on $\mathbb{R}$, which are Hellinger differentiable at $\theta_0 \in \Theta$, that is, there exists a function $\Delta(\cdot) \in L_2$ such that*

$$p_\theta^{1/2}(x) = p_{\theta_0}^{1/2}(x) + (\theta - \theta_0)^t \Delta(x) + r_{\theta,\theta_0}(x),$$

where $\int r^2_{\theta,\theta_0}(x)\, dx = o(\|\theta - \theta_0\|^2)$ as $\theta \to \theta_0$. Let $\mathcal{H}$ be a class of uniformly bounded measurable functions on $\mathbb{R}$. Then

(i) *the map $\theta \mapsto P_\theta h \equiv \int h(x) p_\theta(x)\, dx$ from Θ to $\ell^\infty(\mathcal{H})$ is differentiable at θ_0 with derivative $2\int h(x) p_{\theta_0}^{1/2}(x)\Delta(x)\, dx$.*

(ii) *for any sequence $\sqrt{n}(\hat\theta - \theta_0) \to \mathcal{N}(0, \Sigma)$ weakly with $\Sigma > 0$, the process $\left\{\sqrt{n}(P_{\hat\theta} h - P_{\theta_0} h),\ h \in \mathcal{H}\right\}$ converges weakly to a tight Gaussian process in $\ell^\infty(\mathcal{H})$.*

Proof. Since $\mathcal{H}$ is uniformly bounded, we assume without loss of generality that $\|h\|_\infty \le 1$. Observe that the remainder can be written as

$$P_\theta h - P_{\theta_0} h - 2(\theta - \theta_0)^t \int h p_{\theta_0}^{1/2} \Delta = 2 \int h p_{\theta_0}^{1/2} r_{\theta,\theta_0} + \int h \left[p_{\theta_0}^{1/2} - p_\theta^{1/2}\right]^2.$$

The remainder is small – it is of order $o(\|\theta - \theta_0\|)$ uniformly in $h \in \mathcal{H}$ – since by the Cauchy-Schwarz inequality

$$\sup_{h \in \mathcal{H}} \left| \int h p_{\theta_0}^{1/2} r_{\theta,\theta_0} \right| \le \left(\int r^2_{\theta,\theta_0} \cdot \int p_{\theta_0} \right)^{1/2} = o(\|\theta - \theta_0\|),$$

and

$$\sup_{h \in \mathcal{H}} \left| \int h \left[p_{\theta_0}^{1/2} - p_\theta^{1/2}\right]^2 \right| \le \int \left[p_{\theta_0}^{1/2} - p_\theta^{1/2}\right]^2 = \mathcal{O}(\|\theta - \theta_0\|^2).$$

The second assertion follows from the delta method. □

Applying the preceding Lemma 13 with $\mathcal{H} = \left\{\mathbb{I}_{(-\infty,x]}(\cdot),\ x \in \mathbb{R}\right\}$, and combined with Theorem 12 immediately yields

Corollary 14. *Assume that F_θ and G_θ have Hellinger differentiable densities at θ_0 and π_0, respectively, that C has continuous partial derivatives, and that the vector*

$$\left(\sqrt{n}(\widehat\theta - \theta_0), \sqrt{n}(\widehat\pi - \pi_0), \sqrt{n}(H_n - H)(x, y) \right) \tag{4.3}$$

converges jointly to a Gaussian limit for almost all $0 \le x, y \le 1$. Then

$$\sqrt{n}\left(\widehat{C}_{\widehat\theta,\widehat\pi} - C\right)(u, v),\ 0 \le u, v \le 1$$

converges weakly to a tight Gaussian process in $\ell^\infty([0,1]^2)$.

We end this section by examining an important case where $\widehat{\theta}$ and $\widehat{\pi}$ are *Z-estimators* solving *generalized score equations*

$$\eta_n(\theta) \equiv \int \psi(x,\theta)\, dF_n(x) = 0 \quad \text{and} \quad \widetilde{\eta}_n(\pi) \equiv \int \widetilde{\psi}(y,\pi)\, dG_n(y) = 0, \tag{4.4}$$

where $\psi(x,\cdot)$ $(\widetilde{\psi}(y,\cdot))$ are continuous functions in θ F-almost surely (G-almost surely), and $\psi(\cdot,\theta) \in L_2(F)$ for all θ $(\widetilde{\psi}(\cdot,\pi) \in L_2(G)$ for all $\pi)$. Under sufficient regularity of ψ the Z-functional $T : \mathbb{C}(\Theta) \to \Theta$ which selects a zero – *e.g.*, $T(\eta_n) = \widehat{\theta}$ – is Hadamard differentiable at $\eta = \int \psi\, dF$, *cf.* Rieder (1994). Here $\mathbb{C}(\Theta)$ is the space of all bounded, continuous, real-valued functions on the parameter space $\Theta \subset \mathbb{R}^k$, equipped with the sup-norm. We introduce the functions

$$\eta(\theta) = \int \psi(x,\theta)\, dF(x) \quad \text{and} \quad \widetilde{\eta}(\pi) = \int \widetilde{\psi}(y,\pi)\, dG(y).$$

Theorem 15. *Assume that*

(i) *η $(\widetilde{\eta})$ is a bounded, continuous function, locally homeomorphic at θ_0 (π_0), and has bounded continuous partial derivatives at θ_0 (π_0) and $\eta(\theta_0) = 0$ $(\widetilde{\eta}(\pi_0) = 0)$;*
(ii) *$\{\psi(\cdot,\theta),\ \theta \in \Theta\}$ is an F-Donsker class and $\{\widetilde{\psi}(\cdot,\pi),\ \pi \in \Pi\}$ is a G-Donsker class;*
(iii) *F_θ (G_θ) has a Hellinger differentiable density at $\theta_0 \in \Theta \subseteq \mathbb{R}^K$ $(\pi_0 \in \Pi \subseteq \mathbb{R}^L)$;*
(iiii) *C has continuous partial derivatives.*

Then

$$\sqrt{n}\left(\widehat{C}_{\widehat{\theta},\widehat{\pi}} - C\right)(u,v),\ \ 0 \le u,v \le 1$$

converges weakly to a tight Gaussian process in $\ell^\infty([0,1]^2)$.

Proof. In view of Corollary 14 and assumptions (iii) and (iiii), we only need to verify the finite dimensional convergence (4.2). Invoking assumption (i), Theorem 1.4.2 in Rieder (1994, page 10) guarantees that there exists a neighborhood V of η in $\mathbb{C}(\Theta)$ and a functional $T : V \to \Theta$ such that $f(T(f)) = 0$ for all $f \in V$. Observe that $\eta \in V$ by assumption (i), and η_n is with probability tending to one in such a neighborhood by assumption (ii). Moreover, Rieder's result implies that every such functional T is Hadamard differentiable at η with derivative $T'_\eta(f) = -\left[\eta'(\theta_0)\right]^{-1} f(\theta_0)$. The same holds for $\widetilde{\eta}$, and denote the corresponding functional which assigns a zero by $\widetilde{T}$, that is, $\widetilde{T}(\widetilde{\eta}) = \pi_0$. Note that assumption (ii) implies that $(\ \sqrt{n}(\eta_n - \eta), \sqrt{n}(\widetilde{\eta}_n - \widetilde{\eta}), \sqrt{n}(H_n - H)\ )$ converges weakly to a Gaussian process in $\mathbb{C}(\Theta) \times \mathbb{C}(\Pi) \times D(\mathbb{R}^2)$, whence

$$\sqrt{n}\left(\ \widehat{\theta} - \theta_0, \widehat{\pi} - \pi_0, H_n - H\ \right) = \sqrt{n}\left(\ T(\eta_n) - T(\eta), \widetilde{T}(\widetilde{\eta}_n) - \widetilde{T}(\widetilde{\eta}), H_n - H\ \right)$$

converges weakly to a tight Gaussian process in $\mathbb{R}^K \times \mathbb{R}^L \times D(\mathbb{R}^2)$ after an application of the functional delta-method, which is actually a stronger assertion than the required (4.2). $\square$

Appendix

Let $\overline{C}_n$ be the càdlàg version of C_n defined in (2.2). A consequence of Theorem 4 is that $\overline{C}_n(x,y) \xrightarrow{P} C(x,y)$, uniformly in $0 \le x, y \le 1$. We generalize this convergence by describing classes of functions $\alpha : [0,1]^2 \to \mathbb{R}$ for which

$$\sup_{\alpha} \left| \int_{[0,1]^2} \alpha(x,y)\, d(\overline{C}_n - C)(x,y) \right| \xrightarrow{\text{a.s.}} 0. \tag{4.5}$$

We will formulate the result in terms of bracketing numbers $N_B(\delta, L_1(C), \mathcal{A})$. Recall that for any $\delta > 0$, a δ-bracket in $L_1(C)$ is a pair of functions $[\alpha_L, \alpha_U]$ such that $\alpha_L \le \alpha \le \alpha_U$ and $\int |\alpha_U - \alpha_L|\, dC \le \delta$, and the δ-bracketing number $N_B(\delta, L_1(C), \mathcal{A})$ is the minimal number of $\delta-$brackets needed to cover $\mathcal{A}$.

Theorem 16. *Let $\mathcal{A}$ be a class of functions $\alpha : [0,1]^2 \to \mathbb{R}$, and assume that for every $\delta > 0$*

$$N_B(\delta, L_1(C), \mathcal{A}) < \infty \text{ for every } \delta > 0.$$

Then the uniform law of large numbers (4.5) holds true.

Proof. We first rewrite $\int \alpha(x,y)\, d(\overline{C}_n - C)(x,y)$. For this matter, observe that, using the same notation as in Lemma 3,

$$\int_{[0,1]^2} \alpha(x,y)\, d\overline{C}_n(x,y) = \int_{\mathbb{R}^2} \alpha(F_n x, G_n y)\, dH_n(x,y) = \int_{[0,1]^2} \alpha(F_n^* x, G_n^* y)\, dH_n^*(x,y)$$

and, similarly,

$$\int_{[0,1]^2} \alpha(x,y)\, dC(x,y) = \int_{\mathbb{R}^2} \alpha(Fx, Gy)\, dH(x,y) = \int_{[0,1]^2} \alpha(x,y)\, dH^*(x,y).$$

Hence

$$\begin{aligned} &\int_{[0,1]^2} \alpha(x,y)\, d(\overline{C}_n - C)(x,y) \\ &= \int_{[0,1]^2} \alpha(x,y)\, d(H_n^* - H^*)(x,y) + \int_{[0,1]^2} [\alpha(F_n^* x, G_n^* y) - \alpha(F^* x, G^* y)]\, dH_n^*(x,y). \end{aligned}$$

Since by assumption $\mathcal{A}$ has finite bracketing numbers, it is an H^*-Glivenko-Cantelli class (recall $C = H^*$ by Lemma 3), so

$$\sup_{\alpha \in \mathcal{A}} \left| \int_{[0,1]^2} \alpha(x,y)\, d(H_n^* - H^*)(x,y) \right| \xrightarrow{\text{a.s.}} 0.$$

Fix $\delta > 0$, and let $\{(\alpha_l, \alpha_U)\}$ be a finite set of δ-brackets covering $\mathcal{A}$ in $L_1(H^*)$. Note that the envelope $A \equiv \sup_{\alpha\in\mathcal{A}} |\alpha| \in L_1(H^*)$, and we can take without loss of generality the functions (α_L, α_U) to be continuous as the set of bounded continuous functions on $[0,1]^2$ is dense in $L_1(H^*, [0,1]^2)$. Now observe that

$$
\begin{aligned}
&\sup_{\alpha\in\mathcal{A}} \left| \int_{[0,1]^2} [\alpha(F_n^* x, G_n^* y) - \alpha(F^* x, G^* y)]\ dH_n^*(x,y) \right| \\
&\le \max_{(\alpha_L,\alpha_U)} \left| \int_{[0,1]^2} \alpha_U(F_n^* x, G_n^* y) - \alpha_L(F^* x, G^* y)\, dH_n^*(x,y) \right| \\
&\le \max_{(\alpha_L,\alpha_U)} \int_{[0,1]^2} |\alpha_U(F_n^* x, G_n^* y) - \alpha_U(F^* x, G^* y)|\, dH_n^*(x,y) + \\
&\quad + \max_{(\alpha_L,\alpha_U)} \left| \int_{[0,1]^2} \alpha_U(F^* x, G^* y) - \alpha_L(F^* x, G^* y)\, dH_n^*(x,y) \right| \\
&\le \max_{(\alpha_L,\alpha_U)} \sup_{0\le x,y\le 1} |\alpha_U(F_n^* x, G_n^* y) - \alpha_U(F^* x, G^* y)| + \\
&\quad + \max_{(\alpha_L,\alpha_U)} \int_{[0,1]^2} |\alpha_U(F^* x, G^* y) - \alpha_L(F^* x, G^* y)|\, dH^*(x,y) + \\
&\quad + \max_{(\alpha_L,\alpha_U)} \left| \int_{[0,1]^2} \alpha_U(F^* x, G^* y) - \alpha_L(F^* x, G^* y)\, d(H_n^* - H^*)(x,y) \right|
\end{aligned}
$$

Note that we only have finitely many (α_L, α_U) by assumption. Use the Glivenko-Cantelli theorem and the fact that each α_U is uniformly continuous on $[0,1]^2$ to show that the first term tends to zero. The second term is bounded by δ by the construction of the δ-cover. The third term is small by the strong law of large numbers and since each (α_L, α_U) is H^*-integrable. This concludes our proof. $\square$

The previous result can be used in the context of semiparametric estimation of dependence parameters. Assume that θ is a finite dimensional parameter and

$$M(\theta) \equiv \int_{[0,1]^2} \alpha_\theta(x,y)\, dC(x,y)$$

has a unique, well-separated maximum at θ_0. For instance, consider a collection of copula's C_θ with densities c_θ, and let $\alpha_\theta \equiv \log c_\theta$ (*cf.* Genest *et al.* (1995)). Based on independent observations $(X_1, Y_n), \cdots, (X_n, Y_n)$, estimate θ_0 by $\widehat{\theta}$, which maximizes

$$M_n(\theta) \equiv \frac{1}{n} \sum_{i=1}^{n} \alpha_\theta\left(F_n(X_i), G_n(Y_i)\right) = \int_{[0,1]^2} \alpha_\theta(x,y)\, d\overline{C}_n(x,y)$$

over θ. Under the conditions of the preceding theorem on the class $\mathcal{A}$,

$$\sup_{\alpha \in \mathcal{A}} |\int_{[0,1]^2} \alpha_\theta(x,y)\, d(\overline{C}_n - C)(x,y)| \xrightarrow{\text{a.s.}} 0,$$

and consequently $\widehat{\theta} \xrightarrow{P} \theta_0$. This reasoning is common to establish consistency, see Van der Vaart (1998).

References

[1] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg., Bull. C1 Sci. 5ième sér.*, **65**, 274 – 292.

[2] Deheuvels, P. (1981a). A Kolmogorov-Smirnov Type Test for independence and multivariate samples. *Rev. Roum. Math. Pures et Appl.*, **26**, No 2, 213 – 226.

[3] Deheuvels, P. (1981b). A nonparametric test for independence. *Pub. Inst. Stat. Univ. Paris.*, **26**, No 2, 29 – 50.

[4] Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000). *Copulas for finance. A reading guide and some applications.* GRO, Crédit Lyonnais, Paris.

[5] Capéraà, P., Fougères, A.-L., and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84, 567-577.

[6] Embrechts, P. McNeil, A., Straumann, D (2002) Correlation and dependence in risk management: properties and pitfalls. *In: Risk Management: Value at Risk and Beyond, M.A.H. Dempster, ed.*, Cambridge University Press, Cambridge, pp. 176-223

[7] Genest, C. and MacKay, R. J. (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *La revue canadienne de statistique*, 14, 145-159.

[8] Genest, C. and MacKay, R.J. (1986). The joy of copulas: Bivariate distributions with uniform marginals. *The American Statistician*, 40, 280-283.

[9] Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, 88, 1034-1043.

[10] Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82, 543-552.

[11] Genest, C. and Werker, B. J. M. (2002). Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models. *Proceedings of the Conference on Distributions with Given Marginals and Statistical Modelling (Barcelona), C. M. Cuadras, J. Fortiana et J. A. Rodríguez Lallena (eds.).*

[12] Gouriéroux, C., Monfort, A. and Trognon, A. (1984). Pseudo-maximum likelihood methods : Theory. *Econometrica*, **52**, 681 – 700.

[13] Joe, H. (1997). *Multivariate models and dependence concepts*, Chapman and Hall.

[14] Klaassen, C.A.J. and Wellner, J.A. (1997). Efficient estimation in the bivariate normal copula model: normal margins are least favorable. *Bernouilli*, **3(1)**, 55–77

[15] LeCam, L. (1986). *Asymptotic methods in statistical decision theory.* Springer.

[16] Nelsen, R. B. (1999). *An introduction to copulas*, Lecture Notes in Statistics, **139**, Springer.

[17] RADULOVIĆ, D. AND WEGKAMP, M. (2000). Weak convergence of smoothed empirical processes : beyond Donsker classes. In *High Dimensional Probability II, E. Giné, D. Mason and J. Wellner Editors*, Birkhäuser.

[18] RIEDER, H. (1994). *Robust Asymptotic Statistics*, Springer.

[19] RÜSCHENDORF (1976). Asymptotic normality of multivariate rank order statistics. *Annals of Statisitcs*, **4**, 912 – 923.

[20] RUYMGAART, F.H. (1973). *Asymptotic Theory for Rank Tests for Independence*, MC-tract 43, Mathematisch Instituut, Amsterdam.

[21] RUYMGAART, F.H. (1974). Asymptotic normality of nonparametric tests for independence. *Annals of Statistics*, **2**, 892 – 910.

[22] RUYMGAART, F.H., SHORACK, G.R. AND VAN ZWET, W.R. (1972). Asymptotic normality of nonparametric tests for independence. *Annals of Mathematical Statistics* **43**, 1122 – 1135.

[23] SCHWEIZER, B. AND SKLAR, A. (1974). Operations on distribution functions not derivable from operations on random variables. *Studia Math.*, **52**, 43 − 52.

[24] SCHWEIZER, B. (1991). Thirty years of copulas. *Advances in Probability Distributions with Given Marginals (G. Dall'Aglio, S. Kotz and G. Salinetti (eds)*, Kluwer, Dordrecht.

[25] SEN, P. K. AND SINGER, J. M. (1993). *Large sample methods in Statistics*, Chapman & Hall.

[26] SCHÖNBUCHER, P. AND SCHUBERT, D. (2000). Copula dependent default risk in intensity models. *Preprint Department of Statistics, Bonn University.*

[27] SHORACK, G. AND WELLNER, J. A. (1986). *Empirical processes with applications to Statistics*, Wiley.

[28] SKLAR, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* , **8**, 229 − 231.

[29] VAART, A. W. VAN DER (1994). Weak convergence of smoothed empirical processes. *Scand. J. Statist.* **21**, 501 − 504.

[30] VAART, A. W. VAN DER (1998). *Asymptotic Statistics*, Cambridge UP.

[31] VAART, A. W. VAN DER, AND WELLNER, J. A. (1996). *Weak convergence and empirical processes*, Springer.

ENSAE AND CREST, 3 AVENUE PIERRE LAROUSSE, MALAKOFF CEDEX, FRANCE

E-mail address: fermania@ensae.fr

YALE UNIVERSITY, DEPARTMENT OF STATISTICS, P.O. BOX 208290, NEW HAVEN, CT 06520-8290, UNITED STATES OF AMERICA

E-mail address: dragan.radulovic@yale.edu

E-mail address: marten.wegkamp@yale.edu