Nonparametric inference for copula density function under random censoring

Yassir Rabhi[1] **and Taoufik Bouezmarni**[2]

Abstract: This paper proposes a nonparametric estimation for the copula and copula density functions for right censored data. An empirical version for the bivariate distribution function is studied and its i.i.d. representation is provided. Based on this estimator an i.i.d. representation is deduced for the copula function. An estimator for the copula density, based on local-linear kernel smoothing, is proposed and its triangular i.i.d. approximation is derived. We distinguish the boundary and the interior cases and we provide an estimator for each case.

Key words: Copula density; local-linear kernel estimation; right censored data; i.i.d. representation of the copula

1 Introduction

Modelling the dependence structure is crucial in statistics analysis. The use of copula function to model this structure becomes more popular. In fact, thanks to Sklar (1959), the copula provides the dependence part of the multivariate distribution and allows a flexible way for modelling the dependence structure. For more details, see Joe (1997) and Nelsen (2005). In this work, we consider two random variables, one is supposed to be right censored and the other is completely observed. We are interested in estimating the copula density function for censored data using nonparametric kernel smoothing method. Compared to parametric and semiparametric approaches, the nonparametric method provides more flexibility and supposes a minimum conditions.

For complete data, the estimation of the copula functions is well investigated. Parametric and semiparametric methods, proposed by Genest & Rivest (1993) and Genest et al. (1995), are widely used and studied. However, theses methods suffer from the misspecification problem. Many goodness-of-fit procedures are proposed to select the adequate copula among parametric models, see Genest et al. (2009). To overcome the misspecification problem, Deheuvels (1979) introduced the first nonparametric estimator, the empirical copula. The asymptotic properties of the latter estimator were investigated by Stute (1984), Fermanian et al. (2004), Tsukahara (2005) and Segers (2012). This estimator is frequently used for testing the independence between random variables, see Genest & Rémillard (2004), and in goodness-of-fit tests, see Genest et al. (2009). Smoothed versions of the empirical copula were proposed by Fermanian et al. (2004), Chen & Huang (2007) and Omelka et al. (2009). The last two papers, consider approaches

[1]Université de Sherbrooke, Department of Mathematics. e-mail: yassir.rabhi@usherbrooke.ca,
[2]Université de Sherbrooke, Department of Mathematics. e-mail: taoufik.bouezmarni@usherbrooke.ca

that overcome the boundary bias, because the copula function is defined on a compact support. Gijbels & Mielniczuk (1990) proposed an estimator for the bivariate copula density by using kernel smoothing, in particular, the reflection method. For d dimension, and based on Bernstein polynomials, an estimator was introduced by Sancetta & Satchell (2004).

For right censored data, the parametric and semiparametric estimation were investigated by Shih & Louis (1995) and Joe (1996). To select the adequate copula, goodness-of-fit tests for censored copula models have been investigated by Wang & Wells (2000) and Lakhal-Chaieb (2010), among others.

The nonparametric estimation of the distribution function was studied, among others, by Dabrowska (1988), Stute (1996) and Van der Laan (1996). Recently, Gribkova & Lopez (2015) presented an overview of existing methods for the estimation of the bivariate distribution under three censoring scenarios. They introduced three nonparametric estimators for the copula function, under each scenario, which are empirical estimator versions for right censored data. Also, they have proposed smoothed versions for their estimators, based on kernel functions, and deduced an estimator for the copula density. To solve the boundary problem, they suggest to use the transformed method. The latter authors studied the weak convergence of their copula estimators, however, they have only established the uniform convergence for the copula density estimator, but not its asymptotic distribution.

In this paper, we propose a local-polynomial estimator for the copula density and an empirical estimator for the copula function. The latter is slightly different from the copula cdf estimator in Gribkova & Lopez (2015). To study the weak convergence of the copula density estimator, we provide first an i.i.d. representation of the bivariate cdf estimator $\widehat{F}$ [proposed in Stute (1996)] with a remainder-term of order $\mathcal{O}_{a.s.}\left(n^{-3/4}(\log n)^{\alpha_1}\right)$, which is faster than the one in the representation of Stute (1996). This result, with such a remainder-term rate, is required when we smooth the copula cdf estimator. Then, we obtained an i.i.d. representation for the copula function estimator. As by-product, we establish a result on the oscillation behavior of the bivariate process $\widehat{F}$. To overcome the boundary bias near 0 and 1, we use the local-linear kernel functions for the estimation of the copula density. We give the expressions of the estimator for the interior and boundary regions of $[0,1]$. A triangular i.i.d. representation is established for the copula density estimator in the interior and boundary regions.

The rest of this manuscript paper is organized as follows. In Section 2, we present estimators for the bivariate distribution and the copula function, of a random vector composed by a right-censored variable and a complete random variable. We provide the asymptotic i.i.d. representation for both estimators. Also, a new nonparametric estimator for the copula density function and its asymptotic i.i.d. representation is investigated. The proofs of the theoretical results are given in the appendix.

2 Estimators and main results

Consider the lifetime $Y > 0$ and the covariate $X > 0$ having the joint distribution $F = F_{X,Y}$, with marginal cdf $F_1(x) = F(x, \infty)$ and $F_2(y) = F(\infty, y)$. Suppose that Y is subject to right-censoring by the random variable C, with distribution G, and that one observes $Z = \min(Y; C)$ and $\delta = \mathrm{I}(Y \leq C)$, the failure indicator. We assume that Y and C are independent, $P[\delta = 1|X, Y] = P[\delta = 1|Y]$ and X is an uncensored variable. The observed i.i.d. data is of the form (Z_i, δ_i, X_i), $i = 1, \ldots, n$. Under this setting, an estimator for the copula function

$$\mathbb{C}(u, v) = F\left(F_1^{-1}(u), F_2^{-1}(v)\right), \; 0 \leq u, v \leq 1, \tag{1}$$

is given by

$$\widehat{\mathbb{C}}(u, v) = \widehat{F}\left(\widehat{F}_1^{-1}(u), \widehat{F}_2^{-1}(v)\right) = \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_i}{1 - \widehat{G}(Z_i^-)} \,\mathrm{I}\left(X_i \leq \widehat{F}_1^{-1}(u), Z_i \leq \widehat{F}_2^{-1}(v)\right), \tag{2}$$

where $\widehat{F}(x, y)$ is defined by

$$\widehat{F}(x, y) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{1 - \widehat{G}(Z_i^-)} \,\mathrm{I}\left(X_i \leq x, Z_i \leq y, \delta_i = 1\right), \tag{3}$$

$\widehat{F}_1(x)$ is the empirical counterpart of $F_1(x)$, $\widehat{F}_2(y) = \widehat{F}(\infty, y)$ and $\widehat{G}$ is the Kaplan-Meier estimator of G. The biavariate cdf estimator (3) [see Stute (1996)] can be derived from the relationship

$$\mathrm{d}F_{X,Y}(x, y) = \frac{1}{1 - G(y^-)}\,\mathrm{d}F_{X,Z,\delta}(x, y, 1).$$

As reported in Gribkova & Lopez (2015), $\widehat{F}_2$ is the Kaplan-Meier estimator of F_2. The latter authors studied the asymptotic properties of a slightly different version of the copula estimator (2), though without giving any i.i.d. representation for their estimator. In this manuscript, we provide an i.i.d. representation for $\widehat{\mathbb{C}}(u, v)$ (Theorem 1 below), which will help us to study the limit distribution of our nonparametric estimator of the copula density. To begin, we introduce the following representation of $\widehat{F}(x, y)$.

Proposition 1 *Let $L(z) = P[Z \leq z]$, $L_0(z) = P[Z \leq z, \delta = 0]$ and denote*

$$\chi_i(x, y) = \frac{\mathrm{I}(X_i \leq x, Z_i \leq y, \delta_i = 1)}{\overline{G}(Z_i)} - \iint_{\substack{u \leq x \\ v \leq y}} \left[1 + \frac{\mathrm{I}(Z_i \leq v, \delta_i = 0)}{\overline{L}(Z_i)} - \int_0^{v \wedge Z_i} \frac{\mathrm{d}L_0(t)}{\overline{L}^2(t)}\right] \mathrm{d}F(u, v). \tag{4}$$

The biavariate cdf estimator $\widehat{F}(x, y)$ admits the approximation

$$\widehat{F}(x, y) - F(x, y) = \frac{1}{n}\sum_{i=1}^{n} \chi_i(x, y) + r_n(x, y), \tag{5}$$

where $\sup_{0 \leq x, y \leq U} |r_n(x, y)| = \mathcal{O}\left(n^{-3/4}(\log n)^{\alpha_1}\right)$ a.s. ($\alpha_1 \geq 1$) and $\overline{Q} = 1 - Q$, for a distribution Q.

Note that representations of $\widehat{F}(x,y)$ has been studied previously in the literature [e.g. Stute (1996)], though with a slower rate of the remainder term than ours. The rate of r_n in (5) is needed for the study of the asymptotic distribution of the copula density estimator, defined below.

Theorem 1
1. *Suppose G is a Lipshitz function on $[0,u_L]$, where u_L is the upper bound of L. Let $\{a_n\}$ be a sequence of positive values such that $a_n = \mathcal{O}\left(n^{-1/2}(\log n)^{\alpha_2}\right)$ ($\alpha_2 \geq 1/2$). Then with probability 1,*

$$\sup_{x,y\leq U} \sup_{\substack{|x-x_0|\leq a_n \\ |y-y_0|\leq a_n}} \left| \left[\widehat{F}(x,y) - F(x,y)\right] - \left[\widehat{F}(x_0,y_0) - F(x_0,y_0)\right] \right| = \mathcal{O}\left(n^{-3/4}(\log n)^{\alpha_3}\right), \qquad (6)$$

for any $U \leq u_L$, where $\alpha_3 \geq 1$.
2. *Suppose that the second partial derivatives of F are bounded. Let $a,b \in (0,1)$ be such that F_k ($k=1,2$) is twice differentiable in $[F_k^{-1}(a)-\epsilon, F_k^{-1}(b)+\epsilon]$, for some $\epsilon > 0$, with the first derivative f_k bounded away from zero and the second derivative bounded in absolute value. Let $L_1(z) = P[Z \leq z, \delta = 1]$ and denote*

$$\eta_i(u) = \frac{u - \mathrm{I}(X_i \leq F_1^{-1}(u))}{f_1(F_1^{-1}(u))}$$

and

$$\xi_i(v) = \frac{\overline{F}_2(F_2^{-1}(v))}{f_2(F_2^{-1}(v))}\left[\frac{\mathrm{I}(Z_i \leq F_2^{-1}(v), \delta_i = 1)}{\overline{L}(Z_i)} - \int_0^{F_2^{-1}(v)\wedge Z_i} \frac{\mathrm{d}L_1(t)}{\overline{L}^2(t)}\right].$$

The copula estimator $\widehat{\mathbb{C}}(u,v)$, for $0 < u, v < 1$, have the representation

$$\begin{aligned}\widehat{\mathbb{C}}(u,v) - \mathbb{C}(u,v) =& \frac{1}{n}\sum_{i=1}^{n}\left\{\chi_i(F_1^{-1}(u),F_2^{-1}(v)) + \eta_i(u)\frac{\partial F}{\partial x}(F_1^{-1}(u),F_2^{-1}(v)) - \xi_i(v)\frac{\partial F}{\partial y}(F_1^{-1}(u),F_2^{-1}(v))\right\} \\ &+ r_n^*(u,v), \qquad (7)\end{aligned}$$

where $\sup_{u,v}|r_n^*(u,v)| = \mathcal{O}\left(n^{-3/4}(\log n)^{\alpha^*}\right)$ a.s. ($\alpha^* \geq 1$), and $\partial F/\partial x$ and $\partial F/\partial y$ denote, respectively, the partial derivatives with respect to the first and second arguments of F.

Note that the i.i.d. representation of $\widehat{\mathbb{C}}(u,v)$ come from three sources, the i.i.d. representations of the bivariate estimator $\widehat{F}(u,v)$ (in Proposition 1), the empirical quantile estimator $\widehat{F}_1^{-1}(u)$ and the Kaplan-Meier quantile estimator $\widehat{F}_2^{-1}(v)$.

Following, we define a nonparametric estimator of the copula density based on local-linear kernel smoothing. Denote $\mathcal{A}_1 = [0,h]$, $\mathcal{A}_2 = [h, 1-h]$ and $\mathcal{A}_3 = [1-h, 1]$. Let K be a symmetric density function supported on $(-1,1)$, $h = h_n$ a bandwidth sequence tending to 0 and

$$K_{x,h}(u) = K(u)\,\frac{a_2(x,h) - a_1(x,h)u}{a_0(x,h)a_2(x,h) - a_1^2(x,h)}\,\mathrm{I}(x \in \mathcal{A}_i), \quad (i=1,2,3), \qquad (8)$$

where

$$a_\ell(x,h) = \int_{(x-1)/h}^{x/h} t^\ell K(t)\,\mathrm{d}t$$

for $\ell = 0,1,2$. Notice that $K_{x,h} = K$ when $x \in \mathcal{A}_2$, $\int_{-1}^{1} K_{x,h}(u)\,\mathrm{d}u = 1$ and $\int_{-1}^{1} u\,K_{x,h}(u)\,\mathrm{d}u = 0$. The kernel function $K_{x,h}$, which represents a local linear version of K, was introduced by Lejeune & Sarda (1992) and Jones (1993) in the context of univariate density estimation. In this paper, we use $K_{x,h}$ for the estimation of the copula density $\mathfrak{C}(x,y)$ in order to remove the boundary biases near 0 and 1, i.e. when $x,y \in [0,h] \cup [1-h,1]$. In the sequel, we require the following conditions on $K_{x,h}$

$$\text{(K1)} \quad \left\{ \begin{array}{l} \displaystyle\iint_{[-1,1]^2} u^2\,|K_{x,h}(u)\,K_{y,h}(v)|\,\mathrm{d}u\,\mathrm{d}v < \infty\,, \\ \displaystyle\iint_{[-1,1]^2} uv\,|K_{x,h}(u)\,K_{y,h}(v)|\,\mathrm{d}u\,\mathrm{d}v < \infty. \end{array} \right.$$

A natural estimator of the copula density $\mathfrak{C}(x,y)$, for $x,y \in [h,1-h]$, would be

$$\widehat{\mathfrak{C}}_I(x,y) = \iint_{[-1,1]^2} h^{-2} K\left(\frac{x-u}{h}\right) K\left(\frac{y-v}{h}\right) \mathrm{d}\widehat{\mathbb{C}}(u,v), \tag{9}$$

where $\widehat{\mathbb{C}}$ is defined in (2), and the index I stands for the interior set $[h,1-h]$ of $[0,1]$. Note, however, that under continuous and increasing marginal cdf F_1 and F_2, the copula function $\mathbb{C}$ can also be estimated by

$$\widehat{\mathbb{C}}^*(u,v) = \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_i}{1-\widehat{G}(Z_i^-)}\,\mathrm{I}\left(\widehat{F}_1(X_i) \le u,\, \widehat{F}_2(Z_i) \le v\right).$$

We may then estimate the copula density, for $x,y \in [h,1-h]$, by

$$\widehat{\mathfrak{C}}_I^*(x,y) = \frac{1}{nh^2}\sum_{i=1}^{n} \frac{\delta_i}{1-\widehat{G}(Z_i^-)}\,K\left(\frac{x-\widehat{F}_1(X_i)}{h}\right) K\left(\frac{y-\widehat{F}_2(Z_i)}{h}\right).$$

A general estimator of $\mathfrak{C}(x,y)$, for $x \in \mathcal{A}_i$ and $y \in \mathcal{A}_j$ $(i,j = 1,2,3)$, is given by

$$\widehat{\mathfrak{C}}(x,y) = \iint_{[-1,1]^2} h^{-2} K_{x,h}\left(\frac{x-u}{h}\right) K_{y,h}\left(\frac{y-v}{h}\right) \mathrm{d}\widehat{\mathbb{C}}(u,v), \tag{10}$$

and can be approximated by

$$\widehat{\mathfrak{C}}^*(x,y) = \frac{1}{nh^2}\sum_{i=1}^{n} \frac{\delta_i}{1-\widehat{G}(Z_i^-)}\,K_{x,h}\left(\frac{x-\widehat{F}_1(X_i)}{h}\right) K_{y,h}\left(\frac{y-\widehat{F}_2(Z_i)}{h}\right).$$

In the next result, we establish a triangular i.i.d. representation for $\widehat{\mathfrak{C}}(x,y)$, leading to a bivariate-normal limit distribution for this estimator.

Theorem 2

Suppose that the assumptions in Theorem 1 and (K1) hold, and the first and second partial derivatives of $\mathfrak{C}$ are bounded. Let $\chi_i^*(u,v)$ be the i.i.d. random term in (7) given by

$$\chi_i^*(u,v) = \chi_i(F_1^{-1}(u), F_2^{-1}(v)) + \eta_i(u)\frac{\partial F}{\partial x}(F_1^{-1}(u), F_2^{-1}(v)) - \xi_i(v)\frac{\partial F}{\partial y}(F_1^{-1}(u), F_2^{-1}(v)).$$

The copula density estimator $\widehat{\mathfrak{C}}(x,y)$ admits the representations:

1. *If $x, y \in \mathcal{A}_1 \cup \mathcal{A}_2$, with $x, y > 0$,*

$$\widehat{\mathfrak{C}}(x,y) - \mathfrak{C}(x,y) = \frac{1}{nh^2}\sum_{i=1}^{n}\iint_{[-1,1]^2} \chi_i^*(x-uh, y-vh)\,\mathrm{d}K_{x,h}(u)\,\mathrm{d}K_{y,h}(v) + r_n^1(u,v), \quad (11)$$

where $\sup_{u,v}|r_n^1(u,v)| = \mathcal{O}\left(n^{-3/4}h^{-2}(\log n)^{\alpha^*} + h^2\right)$ a.s. ($\alpha^* \geq 1$). Hence, by the Lindeberg-Feller theorem, and under the assumptions that nh^6, $(\log n)^{4\alpha^*}/nh^4 \to 0$ (as $n \to \infty$ and $h \to 0$), the process $n^{1/2}h\left[\widehat{\mathfrak{C}}(x,y) - \mathfrak{C}(x,y)\right]$ converges weakly to a bivariate normal distribution with mean 0 and covariance matrix Σ_1.

2. *If $x \in \mathcal{A}_3$ and/or $y \in \mathcal{A}_3$, with $0 < x, y < 1$,*

$$\widehat{\mathfrak{C}}(x,y) - \mathfrak{C}(x,y) = \frac{1}{nh^2}\sum_{i=1}^{n}\iint_{[-1,1]^2}\Big\{\chi_i^*(x-uh, y-vh) - \chi_i^*(x-uh,1)\mathrm{I}(y-vh \leq 1) - \chi_i^*(1, y-vh)\mathrm{I}(x-uh \leq 1)\Big\}\,\mathrm{d}K_{x,h}(u)\,\mathrm{d}K_{y,h}(v) + r_n^2(u,v), \quad (12)$$

with $\sup_{u,v}|r_n^2(u,v)| = \mathcal{O}\left(n^{-3/4}h^{-2}(\log n)^{\alpha^*} + h^2\right)$ a.s.

Remark 1

1. *The triangular i.i.d. representation (12) actually generalizes the one in (11), by noticing that $\int_{-1}^{1}\mathrm{I}(y-vh \leq 1)\,\mathrm{d}K_{y,h}(v) = \int_{-1}^{1}\mathrm{I}(x-uh \leq 1)\,\mathrm{d}K_{x,h}(u) = 0$ if $x, y \notin \mathcal{A}_3$. The i.i.d. random term in (12) is reduced to*

$$\chi_i^*(x-uh, y-vh) - \chi_i^*(x-uh, 1)\mathrm{I}(y-vh \leq 1),$$

if $x \notin \mathcal{A}_3$ and $y \in \mathcal{A}_3$, and equal to

$$\chi_i^*(x-uh, y-vh) - \chi_i^*(1, y-vh)\mathrm{I}(x-uh \leq 1),$$

if $x \in \mathcal{A}_3$ and $y \notin \mathcal{A}_3$.

2. *Note that as $h \to 0$ the term $a_\ell(x,h)$ in (8) converges to $a_\ell(x,1) = \int_{-1}^{b_1} t^\ell K(t)\,\mathrm{d}t$ if $x \in \mathcal{A}_1$, and converges to $a_\ell(x,2) = \int_{b_2}^{1} t^\ell K(t)\,\mathrm{d}t$ if $x \in \mathcal{A}_3$, where b_1 and b_2 are two positive constants.*

Appendix: Proofs of main results

Proof of Proposition 1.
Let $H(u,v,0) = P[X \le u, Z \le v, Y - C \le 0] = F_{X,Z,\delta}(u,v,1)$ and $\widehat{H}$ the empirical counterpart of H. First, Notice that

$$\widehat{F}(x,y) = \iint_{u \le x, v \le y} \frac{1}{\widehat{\overline{G}}(v)}\, \mathrm{d}\widehat{H}(u,v,0),$$

and the difference $\widehat{F}(x,y) - F(x,y)$ can be written as

$$\begin{aligned}\widehat{F}(x,y) - F(x,y) = &\int_{v \le y} \Big[\widehat{\overline{G}}^{-1}(v) - \overline{G}^{-1}(v)\Big]\, \mathrm{d}\big[\widehat{H}(x,v,0) - H(x,v,0)\big] - F(x,y)\\ &+ \int_{v \le y} \overline{G}^{-1}(v)\, \mathrm{d}\widehat{H}(x,v,0) + \int_{v \le y} \Big[\widehat{\overline{G}}^{-1}(v) - \overline{G}^{-1}(v)\Big]\, \mathrm{d}H(x,v,0).\end{aligned}$$

Using the uniform convergence result of the Kaplan-Meier estimator $\widehat{G}$, this difference is equal to

$$\begin{aligned}\widehat{F}(x,y) - F(x,y) = &\int_{v \le y} \frac{\widehat{G}(v) - G(v)}{\overline{G}^2(v)}\, \mathrm{d}\big[\widehat{H}(x,v,0) - H(x,v,0)\big] - F(x,y)\\ &+ \int_{v \le y} \overline{G}^{-1}(v)\, \mathrm{d}\widehat{H}(x,v,0) + \int_{v \le y} \frac{\widehat{G}(v) - G(v)}{\overline{G}^2(v)}\, \mathrm{d}H(x,v,0) + r_{1,n}(x,y),\end{aligned} \tag{13}$$

where $\sup_{x,y} |r_{1,n}(x,y)| = \mathcal{O}_{a.s.}\big(n^{-1}\log\log n\big)$. Let

$$I_n(x,y) = \int_{v \le y} \frac{\widehat{G}(v) - G(v)}{\overline{G}^2(v)}\, \mathrm{d}\big[\widehat{H}(x,v,0) - H(x,v,0)\big].$$

In the following, we show that the rate of $\sup_{x,y} |I_n(x,y)|$ is of order $\mathcal{O}_{a.s.}\big(n^{-3/4}(\log n)^{\alpha_1}\big)$ $(\alpha_1 \ge 1)$. Using partial integration, we have

$$|I(x,y)| \le \|\widehat{G} - G\|.\|\widehat{H} - H\|.\big(\|G^{-2}\| + \|G^{-2} - 1\|\big) + \left|\int_{v \le y} \frac{\widehat{H}(x,v,0) - H(x,v,0)}{\overline{G}^2(v)}\, \mathrm{d}\big[\widehat{G}(v) - G(v)\big]\right|. \tag{14}$$

The first term on the R.H.S. of (14) is of order $\mathcal{O}_{a.s.}\big(n^{-1}\log\log n\big)$, by the uniform convergence of $\widehat{G}$ and $\widehat{H}$. For the second term, divide $[0,y]$ into m sub-intervals $[0,y_1]$, $[y_1,y_2], \dots, [y_{m-1},y_m]$ of equal length $\ell = a_0 n^{-1/2}(\log n)^q$ ($q \ge 1/2$ and $a_0 > 0$ is some constant), so m is of order

$\mathcal{O}\left(n^{-1/2}(\log n)^{-q}\right)$. We have

$$\begin{aligned}
&\left|\int_0^y \left[\widehat{H}(x,v,0) - H(x,v,0)\right] \frac{\mathrm{d}\left[\widehat{G}(v)-G(v)\right]}{\overline{G}^2(v)}\right| \\
&\qquad \le \sum_{i=0}^{m-1}\left|\int_{y_i}^{y_{i+1}} \left[\widehat{H}(x,v,0) - H(x,v,0)\right] \frac{\mathrm{d}\left[\widehat{G}(v)-G(v)\right]}{\overline{G}^2(v)}\right| \\
&\qquad \le \sum_{i=0}^{m-1} \|\widehat{H}-H\|.\|G^{-2}\| \int_{y_i}^{y_{i+1}} \left|\mathrm{d}\left[\widehat{G}(v)-G(v)\right]\right| \\
&\qquad \le \|\widehat{H}-H\|.\|G^{-2}\| \sum_{i=0}^{m-1} \sup_{u,v\in[y_i,y_{i+1}]} \left|\left[\widehat{G}(v)-G(v)\right] - \left[\widehat{G}(u)-G(u)\right]\right|. \qquad (15)
\end{aligned}$$

The sup-norm term, inside the summation, on the R.H.S. of (15) is of order $\mathcal{O}_{a.s.}\left(n^{-3/4}(\log n)^{(1+q)/2}\right)$, as $n\to\infty$, by the oscillation result in Meng et al. (1991) (see proposition 1, page 6). Since $\|\widehat{H}-H\|$ and m are of order $\mathcal{O}_{a.s.}\left(n^{-1/2}\left(\log\log n\right)^{1/2}\right)$ and $\mathcal{O}\left(n^{-1/2}(\log n)^{-q}\right)$, respectively, the term on the R.H.S. of (15) is of order $\mathcal{O}_{a.s.}\left(n^{-3/4}(\log n)^{\alpha_1}\right)$ $(\alpha_1 \ge 1)$. Hence,

$$\sup_{x,y} |I_n(x,y)| = \mathcal{O}_{a.s.}\left(n^{-3/4}(\log n)^{\alpha_1}\right),$$

and therefore,

$$\widehat{F}(x,y)-F(x,y) = \int_{v\le y} \left[\widehat{G}(v)-G(v)\right] \frac{\mathrm{d}H(x,v,0)}{\overline{G}^2(v)} + \int_{v\le y} \overline{G}^{-1}(v)\,\mathrm{d}\widehat{H}(x,v,0) - F(x,y) + r_{2,n}(x,y), \qquad (16)$$

where $\sup_{x,y}|r_{2,n}(x,y)| = \mathcal{O}_{a.s.}\left(n^{-3/4}(\log n)^{\alpha_1}\right)$. By using the i.i.d. representation of $\widehat{G}(v)-G(v)$ in Lo et al. (1989), we complete the proof. ■

Proof of Theorem 1.
1. Denote $\widehat{\mathbb{F}}(x,y) = \widehat{F}(x,y) - F(x,y)$, $\widehat{\mathbb{G}}(y) = \widehat{G}(y) - G(y)$ and $\widehat{\mathbb{H}}(x,y) = \widehat{H}(x,y) - H(x,y)$, where $H(x,y) = H(x,y,0)$ is defined in Lemma 1' proof and $\widehat{H}$ is its empirical counterpart. Let x_0 and y_0 two positive values such that $|x-x_0|, |y-y_0| \le a_n$, and denote $\underline{x} = (x, x_0)$ and $\underline{y} = (y, y_0)$. We have

$$\begin{aligned}
\widehat{\mathbb{F}}(x,y) - \widehat{\mathbb{F}}(x_0,y_0) &= \int_0^{y_0} \left[\widehat{\overline{G}}^{-1}(v) - \overline{G}^{-1}(v)\right] \mathrm{d}\left[H(x,v)-H(x_0,v)\right] + \int_{y_0}^{y} \left[\widehat{\overline{G}}^{-1}(v) - \overline{G}^{-1}(v)\right] \mathrm{d}H(x,v) \\
&\quad + \int_0^{y_0} \widehat{\overline{G}}^{-1}(v)\,\mathrm{d}\left[\widehat{\mathbb{H}}(x,v) - \widehat{\mathbb{H}}(x_0,v)\right] + \int_{y_0}^{y} \widehat{\overline{G}}^{-1}(v)\,\mathrm{d}\widehat{\mathbb{H}}(x,v) \\
&= \int_0^{y_0} \frac{\widehat{\mathbb{G}}(v)}{\overline{G}^2(v)}\,\mathrm{d}\left[H(x,v)-H(x_0,v)\right] + \int_{y_0}^{y} \frac{\widehat{\mathbb{G}}(v)}{\overline{G}^2(v)}\,\mathrm{d}H(x,v) \\
&\quad + \int_0^{y_0} \widehat{\overline{G}}^{-1}(v)\,\mathrm{d}\left[\widehat{\mathbb{H}}(x,v) - \widehat{\mathbb{H}}(x_0,v)\right] + \int_{y_0}^{y} \widehat{\overline{G}}^{-1}(v)\,\mathrm{d}\widehat{\mathbb{H}}(x,v) + r_n'(\underline{x},\underline{y}), \qquad (17)
\end{aligned}$$

where $\|r'_n\| = \mathcal{O}_{a.s.}\left(n^{-1}\log\log n\right)$, by using the uniform convergence of $\widehat{G}$. Let $I_n^1(\underline{x},\underline{y})$, $I_n^2(\underline{x},\underline{y})$ and $I_n^3(\underline{x},\underline{y})$ be, respectively, the sum of the first two terms, the third term and the fourth term in (17). We want to find the rates of the sup-norm of $I_n^k(\underline{x},\underline{y})$, for $k = 1, 2, 3$. First, we have

$$|I_n^1(\underline{x},\underline{y})| \leq \|\overline{G}^{-2}\|.\|\widehat{G} - G\|.\left(\int_0^{y_0}\left|\frac{\partial H}{\partial v}(x,v) - \frac{\partial H}{\partial v}(x_0,v)\right| \mathrm{d}v + |H(x,y) - H(x,y_0)|\right),$$

and by using Taylor expansion of first order for $|x - x_0|, |y - y_0| \leq a_n$, under bounded first and second partial derivatives of H, and the uniform convergence of $\widehat{G}$,

$$\sup_{\substack{|x-x_0|\leq a_n \\ |y-y_0|\leq a_n}} |I_n^1(\underline{x},\underline{y})| = \mathcal{O}_{a.s.}\left(a_n n^{-1/2}\big(\log\log n\big)^{1/2}\right).$$

For the rates of $I_n^2(\underline{x},\underline{y})$ and $I_n^3(\underline{x},\underline{y})$, notice that by using partial integration

$$|I_n^2(\underline{x},\underline{y})| \leq \|\widehat{\overline{G}}^{-1}\|.\|\widehat{\overline{G}}^{-1} - 1\|.\sup_{\substack{|x-x_0|\leq a_n \\ v\leq y_0}}\left|\widehat{\mathbb{H}}(x,v) - \widehat{\mathbb{H}}(x_0,v)\right|,$$

and $I_n^3(\underline{x},\underline{y})$ can be written as

$$\begin{aligned} I_n^3(\underline{x},\underline{y}) = &\left[\widehat{\mathbb{H}}(x,y) - \widehat{\mathbb{H}}(x_0,y_0)\right]\widehat{\overline{G}}^{-1}(y) + \left[\widehat{\mathbb{H}}(x_0,y_0) - \widehat{\mathbb{H}}(x,y_0)\right]\widehat{\overline{G}}^{-1}(y_0) \\ &+ \int_{y_0}^{y}\left[\widehat{\mathbb{H}}(x_0,y_0) - \widehat{\mathbb{H}}(x,v)\right]\mathrm{d}\widehat{\overline{G}}^{-1}(v). \end{aligned}$$

Hence, by using theorem 2.3 in Stute (1984)

$$\sup_{\substack{|x-x_0|\leq a_n \\ |y-y_0|\leq a_n}} |I_n^k(\underline{x},\underline{y})| = \mathcal{O}_{a.s.}\left(n^{-3/4}\big(\log n\big)^{1/2}\big(\log\log n\big)^{1/4}\right),$$

for $k = 2, 3$, and the result follows.
2. Using part 1 of Theorem 1 and Taylor expansion, the representation of $\widehat{\mathcal{C}}(u,v)$ follows from the i.i.d. representations of $\widehat{F}(x,y)$, in Proposition 1, and that of $\widehat{F}_1^{-1}(x)$ and $\widehat{F}_2^{-1}(x)$ in Bahadur (1966) and Lo & Singh (1985), respectively, ■

Proof of Theorem 2.
The proof is given for representation (12) when $x, y \in \mathcal{A}_3 = [1-h, 1]$. The proof is similar for the other cases of x and y. Let $\mathrm{d}_u\widehat{\mathbb{C}}(u,v)$ and $\mathrm{d}_v\widehat{\mathbb{C}}(u,v)$ denote, respectively, the differentials of $\widehat{\mathbb{C}}(u,v)$ with respect to the first and second arguments. Using partial integration, first, with

respect to u and then with respect to v, we have

$$
\begin{aligned}
\widehat{\mathfrak{C}}(x,y) &= h^{-2}\iint_{[-1,1]^2} K_{x,h}\left(\frac{x-u}{h}\right) K_{y,h}\left(\frac{y-v}{h}\right) \mathrm{d}\widehat{\mathbb{C}}(u,v) \\
&= h^{-2}\int_{y-h}^{1} K_{x,h}\left(\frac{x-1}{h}\right) K_{y,h}\left(\frac{y-v}{h}\right) \mathrm{d}_v\widehat{\mathbb{C}}(1,v) \\
&\quad + h^{-2}\int_{v=y-h}^{1}\int_{u=x-h}^{1} K_{x,h}^{(1)}\left(\frac{x-u}{h}\right) K_{y,h}\left(\frac{y-v}{h}\right) \frac{\mathrm{d}u}{h}\,\mathrm{d}_v\widehat{\mathbb{C}}(u,v) \\
&= h^{-2}\Bigg\{K_{x,h}\left(\frac{x-1}{h}\right) K_{y,h}\left(\frac{y-1}{h}\right) \widehat{\mathbb{C}}(1,1) \\
&\qquad + \int_{y-h}^{1} \widehat{\mathbb{C}}(1,v)K_{x,h}\left(\frac{x-1}{h}\right) K_{y,h}^{(1)}\left(\frac{y-v}{h}\right) \frac{\mathrm{d}v}{h}\Bigg\} \\
&\quad + h^{-2}\int_{x-h}^{1}\Bigg\{\widehat{\mathbb{C}}(u,1)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right) K_{y,h}\left(\frac{y-1}{h}\right) \\
&\qquad\qquad + \int_{y-h}^{1} \widehat{\mathbb{C}}(u,v)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right) K_{y,h}^{(1)}\left(\frac{y-v}{h}\right) \frac{\mathrm{d}v}{h}\Bigg\}\frac{\mathrm{d}u}{h} \\
&= h^{-2}\Bigg\{\int_{y-h}^{1}\int_{x-h}^{1} K_{x,h}^{(1)}\left(\frac{x-u}{h}\right) K_{y,h}^{(1)}\left(\frac{y-v}{h}\right) \frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h} \\
&\qquad - \int_{y-h}^{1}\int_{x-h}^{1} \widehat{\mathbb{C}}(1,v)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right) K_{y,h}^{(1)}\left(\frac{y-v}{h}\right) \frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h} \\
&\qquad - \int_{y-h}^{1}\int_{x-h}^{1} \widehat{\mathbb{C}}(u,1)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right) K_{y,h}^{(1)}\left(\frac{y-v}{h}\right) \frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h} \\
&\qquad + \int_{y-h}^{1}\int_{x-h}^{1} \widehat{\mathbb{C}}(u,v)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right) K_{y,h}^{(1)}\left(\frac{y-v}{h}\right) \frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h}\Bigg\},
\end{aligned}
$$

and by using the substitutions $u^* = (x-u)/h$ and $v^* = (y-v)/h$,

$$
\begin{aligned}
\widehat{\mathfrak{C}}(x,y) = h^{-2}\iint_{[-1,1]^2} \Big[&\widehat{\mathbb{C}}(x-uh,y-vh) - \widehat{\mathbb{C}}(x-uh,1)\mathrm{I}(y-vh\le 1) - \widehat{\mathbb{C}}(1,y-vh)\mathrm{I}(x-uh\le 1) \\
&+ \mathrm{I}(x-uh\le 1, y-vh\le 1)\Big]\mathrm{d}K_{x,h}(u)\mathrm{d}K_{y,h}(v).
\end{aligned}
$$

The difference $\widehat{\mathfrak{C}}(x,y) - \mathfrak{C}(x,y)$ can be written as

$$
\begin{aligned}
\widehat{\mathfrak{C}}(x,y) - \mathfrak{C}(x,y) &= h^{-2}\iint_{[-1,1]^2}\Bigg\{\Big[\widehat{\mathbb{C}}(x-uh,y-vh) - \mathbb{C}(x-uh,y-vh)\Big] \\
&\qquad - \Big[\widehat{\mathbb{C}}(x-uh,1) - \mathbb{C}(x-uh,1)\Big]\mathrm{I}(y-vh\le 1) \\
&\qquad - \Big[\widehat{\mathbb{C}}(1,y-vh) - \mathbb{C}(1,y-vh)\Big]\mathrm{I}(x-uh\le 1)\Bigg\}\mathrm{d}K_{x,h}(u)\mathrm{d}K_{y,h}(v) \\
&\quad + \iint_{[-1,1]^2}\Big[\mathfrak{C}(x-uh,y-vh) - \mathfrak{C}(x,y)\Big]K_{x,h}(u)\,K_{y,h}(v)\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}
$$

By employing the i.i.d. representation of $\widehat{\mathbb{C}}$ (in Theorem 1) and Taylor expansion of second order, the result follows by using the fact that $\int_{-1}^{1} u\,K_{x,h}(u)\,\mathrm{d}u = \int_{-1}^{1} v\,K_{y,h}(v)\,\mathrm{d}v = 0$. ■

References

[1] Bahadur, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37**, 577-580.

[2] Chen, S. X. and Huang, T. (2007). Nonparametric Estimation of Copula Functions for Dependent Modeling. *Scand. J. Statist.* **35**, 265–282.

[3] Dabrowska, D. M. (1988). KaplanMeier estimate on the plane. *Ann. Statist.* **16**, 1475-1489.

[4] Deheuvels, P. (1979). La Fonction de Dépendance Empirique et ses Propriétés. Un Test non Paramétrique D'independance. *Bullettin de l'académie Royal de Belgique, Classe des Sciences*, 274–292.

[5] Fermanian, J-D., Radulovic, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10**, 847–860.

[6] Genest, C. and Rivest, L. P. (1993). Statistical Inference Procedure for Bivariate Archimedean Copulas. *J. Am. Statist. Assoc.* **88**, 1034–1043.

[7] 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.

[8] Genest, C. and Rmillard, B. (2004). Tests of independence and randomness based on the empirical copula process. *Test* **13**, 335–369.

[9] Genest, C., Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics* **44**, 199–213.

[10] Gijbels, I. and Mielniczuk J. (1990). Estimating The Density of a Copula Function. *Communications in Statistics - Theory and Methods.* **19**, 445–464.

[11] Gribkova, S. and Lopez, O. (2015). Nonparametric copula estimation under bivariate censoring. *Scand. J. Statist.* **42**, 925-946.

[12] Joe, H. (1997). Multivariate Models and Dependence Concepts. *Chapman and Hall, London.*

[13] Joe, H. and Xu, J. (1996). The estimation method of inference functions for margins for multivariate models. *Technical Report, Department of Statistics, University of British Columbia*, **166**.

[14] Jones, M. C. (1993). Simple boundary correction for kernel density estimation. *Statistic Computing* **3**, 135-146.

[15] Kiefer, J. (1967). On Bahadurs representation of sample quantiles. *Ann. Math. Statist.* **38**, 1323-1342.

[16] Lejeune, M. and Sarda, P. (1992). Smooth estimators of distribution and density functions. *Computational Statistics and Data Analysis* **14**, 457–471.

[17] Lakhal-Chaieb, M. L. (2010). Copula inference under censoring. *Biometrika.* **97 (2)**, 505–512.

[18] Lo, S. H. and Singh, K. (1985). The product-limit estimator and the bootstrap: some asymptotic representations. *Probab. Theory Related Fields* **71**, 455–465.

[19] Lo, S. H., Mack, Y. P. and Wang, J. L. (1989). Density and hazard rate estimation for censored data via strong representation of the Kaplan-Meier estimator. *Probab. Theory Related Fields* **80**, 461–473.

[20] Meng, X.-L., Bassiakos, Y. and Lo, S.-H. (1991). Large-sample properties for a general estimator of the treatment effect in the two-sample problem with right censoring. *Ann. Statist.* **19 (4)**, 1786-1812.

[21] Nelsen, R.B. (2006). An Introduction to Copulas. *Springer, New York.*

[22] Omelka, M., Gijbels, I. and Veraverbeke, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. *Ann. Statist.* **37 (5)**, 3023–3058.

[23] Sancetta, A. and Satchell, S. (2004). The Bernstein Copula and its Applications to Modeling and Approximations of Multivariate Distributions. *Econometric Theory* **20**, 535–562.

[24] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **18 (3)**, 764-782.

[25] Shih, J. and Louis, T. (1995). Inferences on the Association Parameter in Copula Models for Bivariate Survival Data. *Biometrics* **51**, 1384–1399.

[26] Sklar, A. (1959). Fonction de Répartition à n Dimensions et leurs Marges. *Publications de l'Institut de Statistique de l'Université de Paris* **8**, 229–231.

[27] Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab.* **12**, 361-379.

[28] Stute, W. (1996). Distributional convergence under random censorship when covariables are present. *Scand. J. Statist.* **23**, 461-471.

[29] Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.* **33**, 357-375.

[30] Van der Laan, M. J. (1996). Efficient estimation in the bivariate censoring model and repairing NPMLE. *Ann. Statist.* **24**, 596-627.

[31] Wang, W. and Wells, M. T. (2000). Model selection and semiparametric inference for bivariate failure-time data. *J. Amer. Statist. Assoc.* **95 (449)**, 62–76.