Nonparametric inference for copulas and measures of dependence under length-biased sampling and informative censoring

Yassir Rabhi[1] and Taoufik Bouezmarni[2]

State University of New York[1] and University of Sherbrooke[2]

Abstract: Length-biased data are often encountered in cross-sectional surveys and prevalent-cohort studies on disease durations. Under length-biased sampling subjects with longer disease durations have greater chance to be observed. As a result, covariate values linked to the longer survivors are favoured by the sampling mechanism. When the sampled durations are also subject to right censoring, the censoring is informative. Modelling dependence structure without adjusting for these issues leads to biased results. In this paper, we consider copulas for modelling dependence when the collected data are length-biased and account for both informative censoring and covariate bias that are naturally linked to length-biased sampling. We address nonparametric estimation of the bivariate distribution, copula function and its density, and Kendall and Spearman measures for right-censored length-biased data. The proposed estimator for the bivariate cdf is a Hadamard-differentiable functional of two MLEs (Kaplan-Meier and empirical cdf) and inherits their efficiency. Based on this estimator, we devise two estimators for copula function and a local-polynomial estimator for copula density that accounts for boundary bias. The limiting processes of the estimators are established by deriving their i.i.d. representations. As a by-product, we establish the oscillation behavior of the bivariate cdf estimator. In addition, we introduce estimators for Kendall and Spearman measures and study their weak convergence. The proposed method is applied to analyze a set of right-censored length-biased data on survival with dementia, collected as part of a nationwide study in Canada.

Keywords: Hadamard-Differentiable functional, Copulas, length-biased sampling, covariate(s) bias, local-linear kernel estimation, i.i.d. representation of copula, measures of dependence

1 Introduction

Modeling the dependence structure between two or more random variables is crucial in statistical analysis. Many measures of association have been proposed in the literature to capture and quantify such dependence. Copula function and its density, for instance, are well known to provide local and global overview of the dependence structure between multiple variables. Copulas embody the dependence structure that couples a multivariate cdf with its marginal distributions. So far, most of the studies on copulas and measures of association were focused on either complete data or right-censored data, and in both, data are randomly sampled from the population. In length-biased sampling, however, data are not randomly selected from the population of interest, but with probability proportional to the length of the selection-variable (e.g. lifetime), i.e.

$$f_{lb}(y, x_1, \ldots, x_m) = \frac{y}{E[Y^*]} f(y, x_1, \ldots, x_m), \tag{1}$$

where f_{lb} and f are the multivariate densities from the sampled and targeted populations, respectively, and Y^* is the selection-variable. The truncation mechanism in such setting tends to over-select large values and under-select small values of some variables (e.g. lifetime). Equation (1) leads to the relationship

$$\mathfrak{C}_{lb}(v, u_1, \ldots, u_m) = \frac{F_{Y^*}^{-1}(v)}{E[Y^*]} \mathfrak{C}(v, u_1, \ldots, u_m), \tag{2}$$

with $\mathfrak{C}_{lb}$ and $\mathfrak{C}$ the copula densities from the sampled and targeted populations, respectively, and $F_{Y^*}^{-1}$ the inverse cdf of Y^*. In this particular situation, one needs to account for the selection bias in the sample to model dependence between the population variables. If not, one may risk to underestimate (resp. overestimate) the degree of dependence between variables for small values (resp. large values) of the selection-variable (e.g. Y^*), and eventually, this will lead to biased results.

Length-biased sampling is often encountered in cross-sectional surveys and prevalent-cohort studies. The literature on selection bias can be traced as far back as Wicksell (1925) (the corpuscle problem), with seminal contributions by Fisher (1934), Neyman (1955) and Cox (1969). The phenomenon of length-biased sampling appears in different areas of research, see for instance Lynden-Bell (1971) in astronomy, Nowell et al. (1988) in land economics, Zelen (1993) in screening and early detection of disease, Nowell & Stanley (1991) in marketing, Terwilliger et al. (1997) in genetics and linkage mapping, Feuerverger & Hall (2000) in applied physics, Wolfson et al. (2001) in epidemiology and geriatric medicine, de Uña-Álvarez (2004) in labor economy, Kvam (2008) in nano physics and Leiva et al. (2008) in water quality.

Most of the nonparametric studies on copulas were focused on representative samples of complete data or right-censored data. For instance, see Deheuvels (1979), Stute (1984), Fermanian et al. (2004), Sancetta & Satchell (2004), Chen & Huang (2007), Omelka et al. (2009) and Segers (2012) in complete data, and Rabhi & Bouezmarni (2016) in right-censored data. To the best of our knowledge, however, no methodology has been proposed in the literature for nonparametric estimation of copulas and measures of dependence under length-biased sampling and right-censoring. Right-censoring is known to be informative under such sampling. This creates additional challenges in building an estimation method for copulas and measures of association.

In this manuscript, we address nonparametric estimation of the bivariate distribution, copula function and its density, and Kendall and Spearman measures when both variables are left-truncated and one of them is subject to informative censoring. Here, we consider the case $m = 1$ in equation (2), however, we can extend it to the multivariate case $m \geq 2$. The proposed estimator for the bivariate cdf is a Hadamard-differentiable functional of two MLEs; Kaplan-Meier and empirical distributions. This estimator inherits the efficiency of the two MLEs [see van der Vaart (1991)]. Based on the bivariate cdf estimator, we devise two estimators for copula function and a local-polynomial estimator for copula density, that accounts for boundary bias. In addition, we introduce estimators for Kendall and Spearman measures. The limiting processes of the estimators are established by deriving i.i.d. representations for these ones. As by-product, we establish the oscillation behavior of the bivariate distribution estimator.

The rest of this manuscript is organized as follows. In §2, we present estimators for the bivariate distribution and copula function, and provide their asymptotic i.i.d. representations and limiting processes. The local-polynomial estimator of copula density is introduced in §3, and its triangular representation and weak convergence are established. In §4, we propose estimators for Kendall and Spearman measures of association and present their asymptotic distributions. In §5, we illustrate our methods on a prevalent-cohort data on survival with dementia, collected as part of a nationwide study in Canada. The proofs of the main results are given in the appendix.

2 Copula and bivariate distributions

2.1 Data setting and notations

We begin by defining the variables that represent the general population and the data obtained from the cross-sectional (c-s) sampling with follow-up. Let (Y^*, X^*) and T^* be two independent random vectors representing, respectively, the bivariate of interest and the truncation-time from the population (T^* is a univariate vector). At the c-s sampling time, one only observes the data (Y^*, X^*, T^*) given that $Y^* \geq T^*$. The resulting sample is biased, as such, we denote by (Y, X, T) the random vector associated with the observed subjects, which arise from the conditional distribution of (Y^*, X^*, T^*) given that $Y^* \geq T^*$. When the n selected subjects to the study are further followed-up, their residual lifetime $R = Y - T$ is subject to random right-censoring R_c. The observed data, obtained from such sampling, are of the form

$$(T_i, X_i, Z_i, \delta_i),\ i = 1, \ldots, n,$$

where $Z = T + \gamma$, $\gamma = \min(R, R_c)$ and $\delta = \mathrm{I}(R \leq R_c)$ is the censoring indicator. In this work, we only consider the case of uncensored covariate X (e.g. age at disease onset). However, as noted above, X suffers from a bias induced by the c-s sampling design. In the sequel, we assume that R_c is independent of (Y, X, T). This assumption is common in right-censored left-truncated data setting, and is reasonable in most practical situations [see Bergeron et al. (2008) and Ning et al. (2014)]. Let F_{X^*,Y^*}, F_{T^*} (with density f_{T^*}) and G be the distributions of (X^*, Y^*), T^* and R_c, respectively, and denote $F_{X,Z,\delta}(x, y, 1) = P[X \leq x, Z \leq y, \delta = 1]$.

2.2 Estimators

We describe the methodology for estimating the bivariate distribution F_{X^*,Y^*} and the copula function $\mathbb{C}$. Our estimation approach is essentially based on the relationship between the population distribution F_{X^*,Y^*} and the cross-sectional distributions $F_{X,Z,\delta}$ and G, through a functional Φ. The relationship is given by

$$F_{X^*,Y^*}(x, y) = \Phi\big(G, F_{X,Z,\delta}\big)(x, y) = \frac{\iint_{u \leq x, v \leq y} \dfrac{1}{w(v)} \,\mathrm{d}F_{X,Z,\delta}(u, v, 1)}{\iint_{u,v>0} \dfrac{1}{w(v)} \,\mathrm{d}F_{X,Z,\delta}(u, v, 1)}, \tag{3}$$

where w is the weight function

$$w(y) = \int_0^y \big[1 - G(t)\big]\,\mathrm{d}t.$$

Equation (3) is obtained by establishing, first, the equation

$$F_{X,Z,\delta}(x,y,1) = \iint_{u\le x, v\le y} \frac{v - \int_{r\le v}(v-r)\,\mathrm{d}G(r)}{\int_{t\ge 0} t\,\mathrm{d}F_{Y^*}(t)}\,\mathrm{d}F_{X^*,Y^*}(u,v), \tag{4}$$

and noticing by partial integration that $v - \int_0^v (v-r)\,\mathrm{d}G(r) = w(r)$. Deriving the left and right sides of (4), one obtains

$$\mathrm{d}F_{X^*,Y^*}(u,v) = \frac{1}{w(v)}\,\mathrm{d}F_{X,Z,\delta}(u,v,1) \times \int_{t\ge 0} t\,\mathrm{d}F_{Y^*}(t), \tag{5}$$

and by integrating the two sides of (5) on $\mathbb{R}_+^2$, one finds

$$\int_{t\ge 0} t\,\mathrm{d}F_{Y^*}(t) = 1\Big/\int_{u,v\ge 0} w^{-1}(v)\,\mathrm{d}F_{X,Z,\delta}(u,v,1). \tag{6}$$

The relationship in (3) is the combined result of (5) and (6). An estimator for F_{X^*,Y^*} can be defined by replacing in (3) the arguments of Φ, $F_{X,Z,\delta}$ and G, by the empirical estimator $\widehat{F}_{X,Z,\delta}(x,y,1) = \sum_{i=1}^n \delta_i\,\mathrm{I}(X_i \le x, Z_i \le y)\Big/n$ and the Kaplan-Meier estimator $\widehat{G}$. The estimator of $F = F_{X^*,Y^*}$ is given by

$$\widehat{F}(x,y) = \Phi\big(\widehat{G}, \widehat{F}_{X,Z,\delta}\big) = \sum_{i=1}^n \frac{\delta_i/\widehat{w}(Z_i)}{\sum_{j=1}^n \delta_j/\widehat{w}(Z_j)}\,\mathrm{I}(X_i \le x,\, Z_i \le y), \tag{7}$$

where $\widehat{w}(y) = \int_0^y \big[1-\widehat{G}(t)\big]\,\mathrm{d}t$. The weights $\big[\delta_i/\widehat{w}(Z_i)\big]/\sum_{j=1}^n \big[\delta_j/\widehat{w}(Z_j)\big]$ in (7) account for both the truncation and censoring mechanisms.

One key step is to explore the efficiency of $\widehat{F}$, by studying the theoretical aspect of the functional Φ. Let $D[a,b]$ be the Banach space of all cadlag functions defined on an interval $[a,b] \subset \overline{\mathbb{R}}_+$, equipped with the uniform norm, and $BV([a,b]\times[c,d])$ the set of all cadlag functions of total bounded variation defined on $[a,b]\times[c,d] \subset \overline{\mathbb{R}}_+^2$. Given $A \in D[a,b]$ and $B \in BV([a,b]\times[c,d])$, consider the maps $\omega(A)(y) = \int_{[0,y]}(1-A)\,\mathrm{d}t$,

$$\varphi(A,B)(x,y) = \int_{[0,x]\times[0,y]} \frac{1}{A}\,\mathrm{d}B, \quad \psi(A,B) = \int_{\mathbb{R}_+^2} \frac{1}{A}\,\mathrm{d}B$$

and

$$\Phi : (A, B) \mapsto (A_* = \omega(A), B) \mapsto \frac{\varphi(A_*, B)(x, y)}{\psi(A_*, B)}.$$

The functional Φ is Hadamard-differentiable on the domain $\mathcal{C}_1 = \{(A, B) : \int |\mathrm{d}B| \leq M, \omega(A) \geq \epsilon\}$, for $M, \epsilon > 0$, at every point (A, B) such that $1/\omega(A)$ is of bounded variation [see lemma 1 and van der Vaart & Wellner (1997)]. The estimator $\widehat{F}$ is a Hadamard-differentiable functional of two MLEs, $\widehat{G}$ and $\widehat{F}_{X,Z,\delta}$, and inherits their efficiency [see theorem 4.1 in van der Vaart (1991)].

We may estimate the copula function $\mathbb{C}(u, v) = F\Big(F_1^{-1}(u), F_2^{-1}(v)\Big)$, $u, v \in [0, 1]$, from right-censored length-biased data, by

$$\widehat{\mathbb{C}}_1(u, v) = \widehat{F}\left(\widehat{F}_1^{-1}(u), \widehat{F}_2^{-1}(v)\right) = \sum_{i=1}^{n} \frac{\delta_i / \widehat{w}(Z_i)}{\sum_{j=1}^{n} \delta_j / \widehat{w}(Z_j)} \mathrm{I}\left(X_i \leq \widehat{F}_1^{-1}(u), Z_i \leq \widehat{F}_2^{-1}(v)\right), \qquad (8)$$

where $\widehat{F}_1(x) = \widehat{F}(x, \infty)$ and $\widehat{F}_2(y) = \widehat{F}(\infty, y)$ are the empirical counterparts of the marginal distributions F_1 and F_2 of F. Note that F_1^{-1}, F_2^{-1}, $\widehat{F}_1^{-1}$ and $\widehat{F}_2^{-1}$ represent the respective inverse functions of F_1, F_2, $\widehat{F}_1$ and $\widehat{F}_2$. Another estimator for $\mathbb{C}$ is

$$\widehat{\mathbb{C}}_2(u, v) = \sum_{i=1}^{n} \frac{\delta_i / \widehat{w}(Z_i)}{\sum_{j=1}^{n} \delta_j / \widehat{w}(Z_j)} \mathrm{I}\left(\widehat{F}_1(X_i) \leq u,\, \widehat{F}_2(Z_i) \leq v\right), \qquad (9)$$

which is asymptotically equivalent to $\widehat{\mathbb{C}}_1$. Notice that $\widehat{\mathbb{C}}_2(u, v) = \widehat{F}\left(\widehat{F}_1^{-1}\big(u^+\big)^-, \widehat{F}_2^{-1}\big(v^+\big)^-\right)$, where $\widehat{F}(x^-, y^-) = \lim_{\substack{(u,v)\to(x,y)\\ u<x, v<y}} \widehat{F}(u, v)$ and $\widehat{F}_k^{-1}\big(u^+\big) = \lim_{\substack{t\to u\\ t>u}} \widehat{F}_k^{-1}(t)$ $(k = 1, 2)$.

Let $u_{\scriptscriptstyle L}$ denote the upper bound of the support of $L(y) = P[Z \leq y]$. To avoid identifiability problem in $[u_{\scriptscriptstyle L}, \infty)$ due to right censoring, we note that $\widehat{F}$, $\widehat{F}_2$ and $\widehat{\mathbb{C}}$ are respectively defined on the sets $\mathcal{A} = [0, +\infty) \times [0, u_{\scriptscriptstyle L})$, $[0, u_{\scriptscriptstyle L})$ and $\mathcal{B} = [0, 1] \times F_2\big([0, u_{\scriptscriptstyle L})\big)$. Let $D(\mathcal{A})$ and $D(\mathcal{B})$ denote the respective Banach spaces of all cadlag functions defined on $\mathcal{A}$ and $\mathcal{B}$.

2.3 Asymptotic properties

We begin by introducing an i.i.d. representation of $\widehat{F}$ in Theorem 1. This result leads to the derivation of the representations of $\widehat{F}_1^{-1}$ and $\widehat{F}_2^{-1}$ in Lemma 3, and that of the copula estimator $\widehat{\mathbb{C}}_1$ in Theorem 2. Let $\mu = \int_{u,v\geq 0} w^{-1}(v)\, \mathrm{d}F_{X,Z,\delta}(u,v,1)$, $L_0^G(t) = P[\gamma \leq t, \delta = 0]$, $L^G(t) = P[\gamma \leq t]$, and denote

$$\chi_i''(x,y) = \frac{\delta_i}{\mu\, w(Z_i)} \Big[\mathrm{I}(X_i \leq x, Z_i \leq y) - F(x,y) \Big],$$

$$\chi_i'(x,y) = \int_{(u,v)\in\mathcal{A}} \int_{t\leq v} \Big[\mathrm{I}(u \leq x, v \leq y) - F(x,y) \Big]\, \eta_i(t)\, \mathrm{d}t\, \frac{\mathrm{d}F(u,v)}{w(v)},$$

with

$$\eta_i(t) = \overline{G}(t) \left[\frac{\mathrm{I}(\gamma_i \leq t, \delta_i = 0)}{\overline{L^G}(\gamma_i)} - \int_0^{t\wedge\gamma_i} \frac{\mathrm{d}L_0^G(s)}{\overline{L^G}^2(s)} \right].$$

The latter is the i.i.d. random term of the representation of the Kaplan-Meier estimator $\widehat{G}$ [see Lo et al. (1989)]. The assumptions used in the next results are given in the appendix.

Theorem 1 *Denote $\chi_i^F = \chi_i' + \chi_i''$. Under Assumption B1, the bivariate cdf estimator $\widehat{F}$ admits for $(x,y) \in \mathcal{A}$ the representation*

$$\widehat{F}(x,y) - F(x,y) = \frac{1}{n} \sum_{i=1}^{n} \chi_i^F(x,y) + r_n^F(x,y), \tag{10}$$

where $\sup_{\mathbb{R}^+\times[0,\tau]} \left| r_n^F(x,y) \right| = \mathcal{O}\left(n^{-1}\log n\right)$ a.s. for every $\tau < u_L$. Thus, $n^{1/2}\Big[\widehat{F} - F\Big]$ converges weakly to a Gaussian process $\mathcal{F}$ in $D(\mathcal{A})$.

The proof of Theorem 1 is given in the appendix. The covariance process of $\mathcal{F}$ is given by

$$\Sigma_{\mathcal{F}}(x,y,x_0,y_0) = \Sigma_1(x,y,x_0,y_0) + \Sigma_2(x,y,x_0,y_0) + \Sigma_2(x_0,y_0,x,y) + \Sigma_3(x,y,x_0,y_0),$$

where,

$$\Sigma_1(x,y,x_0,y_0) = \int_{(u,v)\in\mathcal{A}} \frac{\Big[\mathrm{I}(u\le x, v\le y) - F(x,y)\Big]\Big[\mathrm{I}(u\le x_0, v\le y_0) - F(x_0,y_0)\Big]}{\mu\, w(v)}\,\mathrm{d}F(u,v),$$

$$\Sigma_2(x,y,x_0,y_0) = \int_{\substack{(u,v)\in\mathcal{A}\\(u_0,v_0)\in\mathcal{A}}} \frac{\Big[\mathrm{I}(u\le x, v\le y) - F(x,y)\Big]\Big[\mathrm{I}(u_0\le x_0, v_0\le y_0) - F(x_0,y_0)\Big]}{\mu\, w(v)\, w(v_0)}$$
$$\times\left\{\int_{t\le v}\int_{s\le t\wedge r} \overline{G}(t)\,\frac{\mathrm{d}L_0^G(s)}{\overline{L^G}^2(s)}\,\mathrm{d}t\right\}\mathrm{d}F_{X,Z,\gamma,\delta}(u_0,v_0,r,1)\,\mathrm{d}F(u,v)$$

and

$$\Sigma_3(x,y,x_0,y_0) = \int_{\substack{(u,v)\in\mathcal{A}\\(u_0,v_0)\in\mathcal{A}}} \Big[\mathrm{I}(u\le x, v\le y) - F(x,y)\Big]\Big[\mathrm{I}(u_0\le x_0, v_0\le y_0) - F(x_0,y_0)\Big]$$
$$\times\,\frac{\sigma_G^*(v,v_0)}{w(v)\,w(v_0)}\,\mathrm{d}F(u_0,v_0)\,\mathrm{d}F(u,v),$$

with $\sigma_G^*(v,v_0) = \int_0^v\int_0^{v_0}\sigma_G(t,s)\,\mathrm{d}s\,\mathrm{d}t$ and $\sigma_G(t,s)$ the covariance function of the limiting process of the Kaplan-Meier estimator $\widehat{G}$. The next result establishes the oscillation behavior of $\widehat{F}$. This result is required to derive the i.i.d. representations of $\widehat{\mathbb{C}}_1$ and $\widehat{\mathbb{C}}_2$.

Proposition 1 (Oscillation behavior of $\widehat{F}$)

Let $\{a_n\}$ be a sequence of positive values such that $a_n = \mathcal{O}\left(n^{-1/2}(\log n)^{\alpha_1}\right)$, with $\alpha_1 \ge 1/2$, and denote $\mathcal{A}_\tau = \mathbb{R}^+\times[0,\tau]$, where $\tau < u_L$. Suppose Assumption B1 holds, then,

$$\sup_{\substack{(x,y)\in\mathcal{A}_\tau\\(x_0,y_0)\in\mathcal{A}_\tau}}\ \sup_{\substack{|x-x_0|\le a_n\\|y-y_0|\le a_n}} \left|\big[\widehat{F}(x,y)-F(x,y)\big] - \big[\widehat{F}(x_0,y_0)-F(x_0,y_0)\big]\right| = \mathcal{O}_{a.s.}\left(n^{-3/4}(\log n)^{\alpha_2}\right)$$

for every $\tau < u_L$, where $\alpha_2 \ge 1$.

The proof is detailed in the appendix. Next, we introduce the i.i.d. representation for $\widehat{\mathbb{C}}_1$, which leads to a representation for the related copula estimator $\widehat{\mathbb{C}}_2$. This will helps to study the limit distribution of our nonparametric estimator of the copula density. Let $\partial_1\mathbb{C}$ and $\partial_2\mathbb{C}$ be the partial derivatives of $\mathbb{C}$ with respect to the first and second arguments of $\mathbb{C}$, respectively, and denote $\xi_i^{F_1}(u) = \chi_i^F\big(F_1^{-1}(u), \infty\big)$ and $\xi_i^{F_2}(v) = \chi_i^F\big(\infty, F_2^{-1}(v)\big)$.

Theorem 2 *Let* $(u_*, v_*) = \big(F_1^{-1}(u), F_2^{-1}(v)\big)$ *and* $\mathcal{B}_\tau = [0,1] \times F_2\big([0,\tau]\big)$*, where* $\tau < u_L$*. Under Assumptions B1 and B2(i,ii,iii), the copula estimator* $\widehat{\mathbb{C}}_1$ *admits for* $(u,v) \in \mathcal{B}$ *the representation*

$$\widehat{\mathbb{C}}_1(u,v) - \mathbb{C}(u,v) = \frac{1}{n}\sum_{i=1}^{n}\Big\{\chi_i^F(u_*, v_*) + \xi_i^{F_1}(u)\,\partial_1\mathbb{C}(u,v) + \xi_i^{F_2}(v)\,\partial_2\mathbb{C}(u,v)\Big\} + r_n^{\mathbb{C}}(u,v) \quad (11)$$

where $\sup_{\mathcal{B}_\tau}\big|r_n^{\mathbb{C}}(u,v)\big| = \mathcal{O}\left(n^{-3/4}(\log n)^{\alpha_*}\right)$ *a.s. for every* $\tau < u_L$*, with* $\alpha_* \geq 1$*. Therefore,* $n^{1/2}\Big[\widehat{\mathbb{C}}_1 - \mathbb{C}\Big]$ *converges weakly to a Gaussian process* $\mathbb{C}_L$ *in* $D(\mathcal{B})$*, with asymptotic variance*

$$\Sigma_{\mathbb{C}_L}(u,v) = E\Big[\big(\chi_1^F(u_*, v_*) + \xi_1^{F_1}(u)\,\partial_1\mathbb{C}(u,v) + \xi_1^{F_2}(v)\,\partial_2\mathbb{C}(u,v)\big)^2\Big]$$

The proof of Theorems 2 is detailed in the appendix. Note that the i.i.d. representation of $\widehat{\mathbb{C}}_1$ comes from three sources, the i.i.d. representations of the bivariate estimator $\widehat{F}$ and the empirical quantile estimators $\widehat{F}_1^{-1}$ and $\widehat{F}_2^{-1}$ (see Lemma 3). Having established Theorem 2, one may analogously derive a representation for $\widehat{\mathbb{C}}_2(u,v) = \widehat{F}\left(\widehat{F}_1^{-1}(u^+)^-, \widehat{F}_2^{-1}(v^+)^-\right)$, given by,

$$\widehat{\mathbb{C}}_2(u,v) - \mathbb{C}(u,v) = \frac{1}{n}\sum_{i=1}^{n}\Big\{\chi_i^F(u_*^-, v_*^-) + \xi_i^{F_1}(u^+)\,\partial_1\mathbb{C}(u,v) + \xi_i^{F_2}(v^+)\,\partial_2\mathbb{C}(u,v)\Big\} + r_n^*(u,v), \quad (12)$$

where $\chi_i^F(u_*^-, v_*^-) = \lim_{\substack{(x,y)\to(u_*,v_*)\\ x<u_*, y<v_*}} \chi_i^F(u_*, v_*)$ and $\sup_{\mathcal{B}_\tau}|r_n^*(u,v)| = \mathcal{O}\left(n^{-3/4}(\log n)^{\alpha^*}\right)$ a.s. for any $\tau < u_L$ ($\alpha^* \geq 1$). Note that the covariance process of $\mathbb{C}_L$ can be deduced from the covariance function $\Sigma_{\mathcal{F}}$ of the limit process $\mathcal{F}$.

3 Local-polynomial estimator for copula density

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

$$K_{x,h}(u) = K(u)\,\frac{a_2(x,h) - a_1(x,h)u}{a_0(x,h)a_2(x,h) - a_1^2(x,h)}\,\mathrm{I}(x \in \mathcal{A}_i), \quad (i = 1,2,3), \tag{13}$$

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. Their purpose of using $K_{x,h}$ is to boost the rate of the estimator bias from $\mathcal{O}(h)$ to $\mathcal{O}(h^2)$ near the compact support boundaries. Here, 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 $K_{x,h}$

K1: (i) $\int_{\frac{u-1}{h}}^{\frac{u}{h}} K_{u,h}^2(t)\,\mathrm{d}t < \infty$, $(u = x, y)$.

(ii) $\int_{\frac{u-1}{h}}^{\frac{u}{h}} t^2\,|K_{u,h}(t)|\,\mathrm{d}t < \infty$, $(u = x, y)$.

To estimate the copula density $\mathfrak{C}$, we consider the copula estimator $\widehat{\mathbb{C}}_2$. An estimator for $\mathfrak{C}$ is given by

$$\widehat{\mathfrak{C}}(x,y) = \frac{1}{h^2}\sum_{i=1}^{n} \frac{\delta_i/\widehat{w}(Z_i)}{\sum_{j=1}^{n}\delta_j/\widehat{w}(Z_j)}\,K_{x,h}\left(\frac{x - \widehat{F}_1(X_i)}{h}\right)\,K_{y,h}\left(\frac{y - \widehat{F}_2(Z_i)}{h}\right). \tag{14}$$

We note that $\mathfrak{C}$ is defined on the set $\mathcal{B}_* = \mathcal{B}\backslash\{(0,0),(1,0)\}$. The copula estimator $\widehat{\mathbb{C}}_2$ allows for explicit expressions for the estimators of $\widehat{\mathfrak{C}}$, and Kendall and Spearman measures of association defined below. 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 3

Suppose Assumptions B1, B2 and K1 hold, and denote $(u_*, v_*) = \big(F_1^{-1}(u), F_2^{-1}(v)\big)$ *and*

$$\chi_i^{\mathbb{C}}(u,v) = \chi_i^F(u_*^-, v_*^-) + \xi_i^{F_1}(u^+)\,\partial_1\mathbb{C}(u,v) + \xi_i^{F_2}(v^+)\,\partial_2\mathbb{C}(u,v)$$

The copula density estimator $\widehat{\mathfrak{C}}(x,y)$ *admits for* $(x,y) \in \mathcal{B}_*$ *the representation*

$$\begin{aligned}\widehat{\mathfrak{C}}(x,y) - \mathfrak{C}(x,y) =& \frac{1}{nh^2}\sum_{i=1}^{n}\int_{[-1,1]^2}\Big\{\chi_i^{\mathbb{C}}(x-uh, y-vh) - \chi_i^{\mathbb{C}}(x-uh,1)\mathrm{I}(y-vh \le 1) \\ & - \chi_i^{\mathbb{C}}(1, y-vh)\mathrm{I}(x-uh \le 1)\Big\}\,\mathrm{d}K_{x,h}(u)\,\mathrm{d}K_{y,h}(v) + r_n^{\mathfrak{C}}(u,v), \end{aligned} \tag{15}$$

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

The proof is given in the appendix. The representation (15) follows from the i.i.d. representation of $\widehat{\mathbb{C}}_2$ in (12).

Corollary 1

Suppose Assumptions B1,B2 and K1 hold and nh^6, $(\log n)^{4\alpha^*}/nh^4 \to 0$ *as* $n \to \infty$ *and* $h \to 0$. *Then, for* $(x,y) \in \mathcal{B}_*$, $n^{1/2}h\left[\widehat{\mathfrak{C}}(x,y) - \mathfrak{C}(x,y)\right]$ *converges in distribution to a zero-mean bivariate normal distribution.*

The proof follows from the triangular representation (15) by using the Lindeberg-Feller CLT theorem.

Remark 1

One practical issue of interest in real-data applications is the choice of the bandwidth h.

To select this parameter, we may minimize with respect to h the integrated squared error $\mathrm{ISE}(h) = \int_{\mathcal{B}_*} \left[\widehat{\mathfrak{C}}(x, y; h) - \mathfrak{C}(x, y)\right]^2 \mathrm{d}x\, \mathrm{d}y$*. This is equivalent to choose h that minimizes*

$$\mathrm{ISE}_*(h) = \int_{\mathcal{B}_*} \widehat{\mathfrak{C}}(x, y; h)^2 \, \mathrm{d}x \, \mathrm{d}y - 2 \int_{\mathcal{B}} \widehat{\mathfrak{C}}(x, y; h) \, \mathrm{d}\mathbb{C}(x, y).$$

The unknown copula function $\mathbb{C}$, in the second term on the R.H.S. of the latter equality, can be replaced by the estimator $\widehat{\mathbb{C}}_2$. The data driven bandwidth is then

$$\widehat{h}_{opt} = \arg\min_h \left\{ \int_{\mathcal{B}_*} \widehat{\mathfrak{C}}_{-i}(x, y; h)^2 \mathrm{d}x\mathrm{d}y - 2 \sum_{i=1}^{n} \frac{\delta_i / \widehat{w}(Z_i)}{\sum_{j=1}^{n} \delta_j / \widehat{w}(Z_j)} \widehat{\mathfrak{C}}_{-i}\big(\widehat{F}_1(X_i), \widehat{F}_2(Z_i); h\big) \right\}, \tag{16}$$

where $\widehat{\mathfrak{C}}_{-i}$ is a leave-one-out estimate of $\mathfrak{C}$ given by

$$\widehat{\mathfrak{C}}_{-i}(x, y; h) = \frac{1}{h^2} \sum_{\substack{\ell=1 \\ \ell \neq i}}^{n} \frac{\delta_\ell / \widehat{w}(Z_\ell)}{\sum_{j=1}^{n} \delta_j / \widehat{w}(Z_j)} K_{x,h} \left(\frac{x - \widehat{F}_1(X_\ell)}{h} \right) K_{y,h} \left(\frac{y - \widehat{F}_2(Z_\ell)}{h} \right).$$

4 Kendall and Spearman measures of association

In this section, we discuss two measures of dependence, or concordance, known as the Kendall's tau and Spearman's rho. For such quantities, we introduce two estimators that are adapted for right-censored length-biased data. Kendall's tau measures the difference between the probability of concordance and discordance between two random variables X^* and Y^*, and is defined by

$$\tau_{X^*,Y^*} = P\Big[(X^* - X^*_0)(Y^* - Y^*_0) > 0\Big] - P\Big[(X^* - X^*_0)(Y^* - Y^*_0) < 0\Big],$$

where (X^*_0, Y^*_0) is an independent copy of (X^*, Y^*). Since the tail region information on the survival function of Y^* may not be identifiable in $[u_L, \infty)$ due to right censoring (as indicated in §2.2), we estimate here a truncated version of Kendall's tau, given by

$$\tau_{X^*,Y^*} = 4 \int_{\mathcal{B}} \mathbb{C}(u, v) \, \mathrm{d}\mathbb{C}(u, v) - 1.$$

An estimator of τ_{X^*,Y^*} under right-censored length-biased data is

$$\widehat{\tau}_{X^*,Y^*} = 4\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\delta_i\,\delta_j/\big[\widehat{w}(Z_i)\,\widehat{w}(Z_j)\big]}{\Big[\sum_{l=1}^{n}\delta_l/\widehat{w}(Z_l)\Big]^2}\,\mathrm{I}\Big(X_j \le X_i, Z_j \le Z_i\Big) - 1, \tag{17}$$

The weight $\Big[\delta_i\delta_j/\big(\widehat{w}(Z_i)\,\widehat{w}(Z_j)\big)\Big]/\Big[\sum_{l=1}^{n}\delta_l/\widehat{w}(Z_l)\Big]^2$ accounts for the truncation and the censoring mechanisms, and replace the uniform weight $1/n$ in the empirical version of $\widehat{\tau}_{X^*,Y^*}$ for complete data. Wang & Wells (2000) discussed the limitations of estimating Kendall's tau under right-censoring. In the next result, we establish the asymptotic distribution of the Kendall's tau estimator. The proof is detailed in the appendix.

Theorem 4 *Suppose Assumptions B1 and B2(i,ii,iii) hold. We have* $\sqrt{n}\big[\widehat{\tau}_{X,Y} - \tau_{X,Y}\big]$ *converges weakly to the normal variable* Z_τ*, given by*

$$Z_\tau = 4\left\{\int_{\mathcal{B}} \mathbb{C}(u,v)\,\mathrm{d}\mathbb{C}_L(u,v) + \int_{\mathcal{B}} \mathbb{C}_L(u,v)\,\mathrm{d}\mathbb{C}(u,v)\right\},$$

where $\mathbb{C}_L$ *is the limiting process of* $\sqrt{n}\big[\widehat{\mathbb{C}}_1(u,v) - \mathbb{C}(u,v)\big]$.

Spearman's rho dependence measure for the random vector (X,Y) is defined as

$$\rho_{X,Y} = 3\left\{P\Big[(X - X_0)(Y - Y_1) > 0\Big] - P\Big[(X - X_0)(Y - Y_1) < 0\Big]\right\},$$

where (X_0, Y_0), (X_1, Y_1) and (X, Y) are independent and identically distributed random vectors. As indicated above, to avoid identifiability problem in $[u_L, \infty)$ caused by right censoring, we estimate a truncated version of Spearman's rho;

$$\rho_{X,Y} = 12\int_{\mathcal{B}} uv\,\mathrm{d}\mathbb{C}(u,v) - 3.$$

An estimator of ρ_{X^*,Y^*}, for right-censored length-biased data, is given by

$$\widehat{\rho}_{X^*,Y^*} = 12\sum_{i=1}^{n} \frac{\delta_i/\widehat{w}(Z_i)}{\sum_{l=1}^{n}\delta_l/\widehat{w}(Z_l)}\,\widehat{F}_1(X_i)\,\widehat{F}_2(Z_i) - 3. \tag{18}$$

The limit distribution of $\widehat{\rho}_{X,Y}$ is derived in the following theorem. The proof is given in the appendix.

Theorem 5 *Under Assumptions B1 and B2(i,ii,iii),* $\sqrt{n}\big[\widehat{\rho}_{X,Y} - \rho_{X,Y}\big]$ *converges weakly to the Gaussian variable*

$$Z_\rho = 12 \int_{\mathcal{B}} u\, v\, \mathrm{d}\mathbb{C}_L(u,v).$$

where $\mathbb{C}_L$ *is the limiting process of* $\sqrt{n}\big[\widehat{\mathbb{C}}_1(u,v) - \mathbb{C}(u,v)\big]$.

5 Survival with dementia

We apply the method described in §2, §3 and §4 to a set of right-censored length-biased data collected on elderly Canadians with dementia. In 1991/1992 a nationwide cross-sectional survey was conducted in five regions of Canada among 9008 community-residing persons and 1255 institutionalized persons aged 65 and older. The CSHA-1 (Canadian Study of Health and Aging 1) identified 1132 persons with dementia who were followed for a period of 5 years until 1996/1997. The primary purpose of the CSHA-1 was the study of the risk factors for dementia and to determine its prevalence in the Canadian population. Wolfson et al. (2001) and Asgharian et al. (2002) reported that those patients with missing date of onset or with survival ≥ 20 years, who unlikely had dementia, need to be excluded. We then considered a sample of $n = 807$ patients in our statistical analysis, among whom 627 died and 180 were censored during the follow-up. The variable Y^* (lifetime) is defined as the time elapsed from the onset of dementia to death, the covariate X^* is the age at onset-of-dementia (AAO) and the left-truncation variable T^* is the time from disease onset to study recruitment.

The purpose of the present example is to study the dependence structure between lifetime Y^* and age at onset-of-dementia X^*. First, we used the nonparametric method

Table 1: Kendall's tau and Spearman's rho estimates for lifetime vs age at onset-of-dementia

Groups	all patients	AAO $\leq$ 75	75 $<$ AAO $\leq$ 85	85 $<$ AAO
$\widehat{\tau}_{X^*,Y^*}$	-0.256	-0.183	-0.022	-0.125
$\widehat{\rho}_{X^*,Y^*}$	-0.366	-0.249	-0.015	-0.170

of Wang (1991) to estimate the truncation distribution. Figure 1 displays this estimator and indicates that a uniform truncation distribution is a reasonable assumption. Addona & Wolfson (2006, p. 277) developed goodness-of-fit tests and found that the uniform assumption is valid for this data. In Table 1, the estimated values of Kendall and Spearman measures show a moderate dependence between Y^* and X^* for $n = 807$ patients with dementia. However, when we divided those individuals into three groups of age at onset-of-dementia (AAO), the dependence becomes weaker for the groups $75 < AAO \leq 85$ and $AAO > 85$. The plots in Figure 3 concur with this remark. The curves of the copula densities estimators for the two groups $75 < AAO \leq 85$ and $AAO > 85$ are relatively flat to the level of the plane $z = 1$ (grey). Notice a sharp peak in the neighborhood of $(x, y) = (0, 1)$ in the plots of the total group of patients and the group $AAO \leq 75$. This can be interpreted as those patients who experienced dementia in an early age (small values of X^*) are more likely to live longer (large values of Y^*). We note that we used the kernel function $K(x) = 0.75(1 - x^2)\,\mathrm{I}_{[-1,1]}(x)$ and the bandwidth h_n is selected via formula (16) in Remark 1.

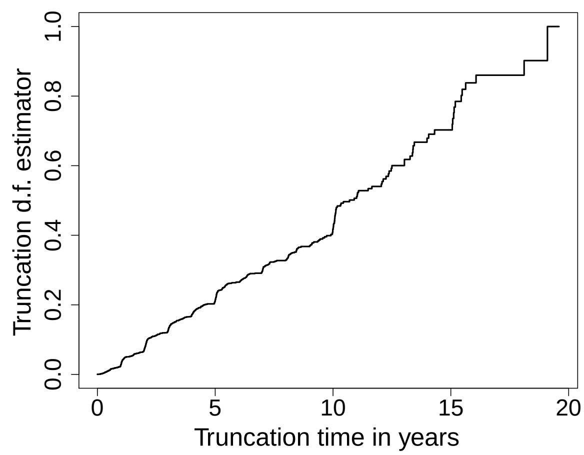

Figure 1: Nonparametric estimator of the truncation cdf for 807 patients with dementia.

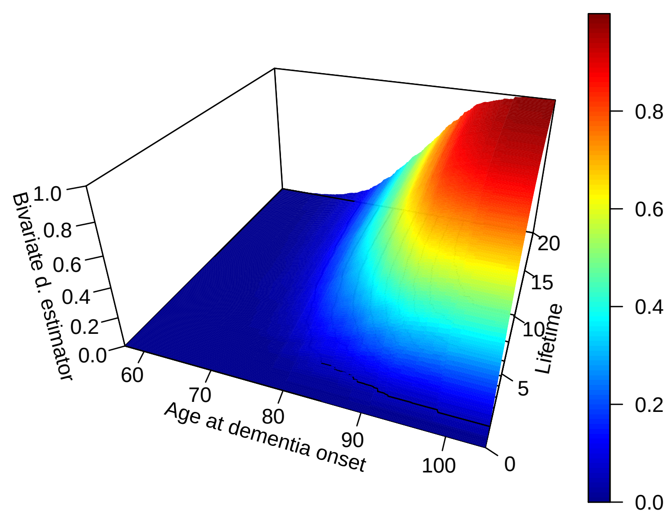

Figure 2: Bivariate cdf estimator of lifetime and age at onset-of-dementia for 807 patients.

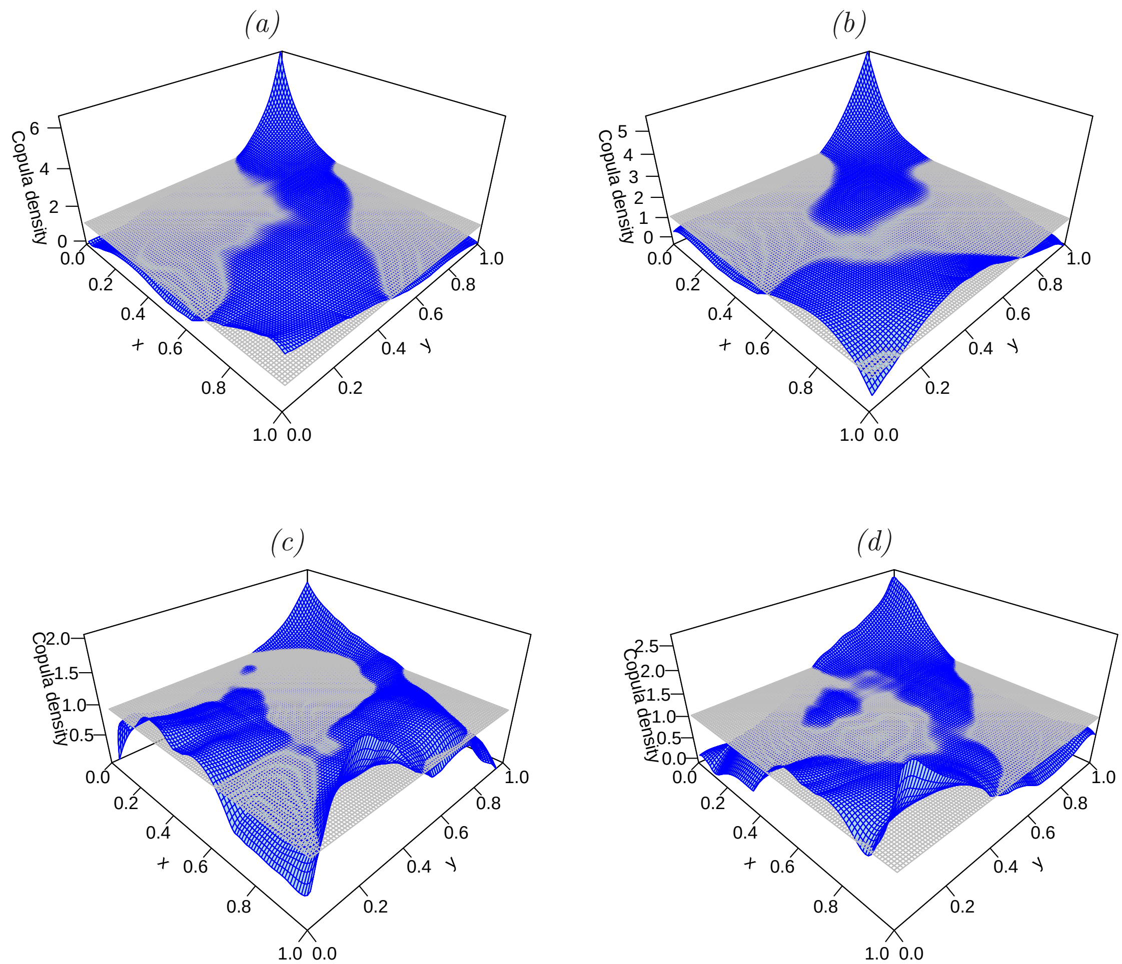

Figure 3: Copula density estimator of lifetime and age at onset-of-dementia (blue) and plane $z = 1$ *(grey): (a) All 807 patients with dementia, (b) patients with* $AAO \leq 75$ *years, (c) patients with* $AAO \in (75, 85]$*, (d) patients with* $AAO > 85$ *years.*

A Appendix

A.1 Notations

We introduce here some notations used in the proofs below. For a general distribution E, l_E and u_E represent the respective lower and upper bounds of the support of E, with $\overline{E} = 1 - E$. Denote $\widehat{\mathbb{F}} = \widehat{F} - F$, $\widehat{\mathbb{G}} = \widehat{G} - G$ and $\widehat{\mathbb{H}} = \widehat{H}_0 - H_0$, where $H_0 = F_{X,Z,\delta}$ and $\widehat{H}_0$ its empirical counterpart. Let l_L be the lower bound of the support of $L(y) = P[Z \leq y]$ and $\mathcal{A}_* = [0, +\infty) \times [l_L, +\infty)$. Denote

$$
\begin{aligned}
F_0(x,y) &= \int_{l_L}^{y} \frac{1}{w(v)} \, \mathrm{d}H_0(x,v), \\
\widehat{F}_0(x,y) &= \int_{l_L}^{y} \frac{1}{\widehat{w}(v)} \, \mathrm{d}\widehat{H}_0(x,v), \\
\widehat{\mu} &= \int_{(u,v)\in\mathcal{A}_*} \frac{1}{\widehat{w}(v)} \, \mathrm{d}\widehat{H}_0(u,v)
\end{aligned}
$$

and notice that $\mu = \int_{(u,v)\in\mathcal{A}_*} \mathrm{d}H_0(u,v)/w(v)$.

A.2 Assumptions

The following regularity conditions are needed to establish the asymptotic results in this paper. Let Q be the cdf of the residual lifetime R. Recall that L is the distribution of Z.

B1: (i) $l_G \wedge l_Q > 0$ and $l_{F_{T^*}} > 0$.

(ii) $u_G \leq u_Q$ with $G(u_G) < 1$.

(iii) $0 < l_L < u_L < \infty$.

The first part of assumption (i) essentially means that there is no immediate failure or censoring at the beginning of the study, while the second part of the assumption means that all subjects recruited to the study are prevalent cases. This condition reflects the setting of the CSHA data and in general is reasonable in most prevalent cohort studies [see Asgharian et al. (2002)]. Assumption (ii) means that the lifetimes of some individuals, who are still alive at the end of the study, will be censored. This is common in the follow-up studies and is due to the limited time of the follow-up. The condition $l_{_L} > 0$ is a direct consequence of (i), while $u_{_L} < \infty$ means that the observed lifetime of individual is finite. The regularity assumption B2 below is required for the asymptotic properties of the estimators of copulas and measures of association.

B2: (i) F_k $(k = 1, 2)$ is twice differentiable in $[F_k^{-1}(a_*) - \epsilon, F_k^{-1}(b_*) + \epsilon]$ for numbers $a_*, b_* \in (0, 1)$ and $\epsilon > 0$.

(ii) $F_k^{(1)} = f_k$ is bounded away from zero and $F_k^{(2)}$ is bounded in absolute value $(k = 1, 2)$.

(iii) The second partial derivatives of F are bounded.

(iv) The first and second partial derivatives of $\mathfrak{C}$ are bounded.

A.3 Lemmas and Proofs

Lemma 1

Under assumption B1, $n^{1/2}\left[\widehat{F} - F\right]$ *converges weakly to a tight process* $\mathcal{F}$ *in* $D(\mathcal{A})$.

Proof.

First, notice that the estimator $\widehat{F}$ depends on the pair $(\widehat{w}, \widehat{H}_0)$ through the composition of the two maps $\varphi(A, B)(x, y) = \int_{[0,x]\times[0,y]} \frac{1}{A}\, \mathrm{d}B$ and $\psi(A, B) = \int_{\mathbb{R}_+^2} \frac{1}{A}\, \mathrm{d}B$ given by

$$\phi : (A, B) \mapsto \frac{\varphi(A, B)(x, y)}{\psi(A, B)}$$

Following similar arguments to that of Veraverbeke et al. (2011) (lemma 1) and van der Vaart & Wellner (1997) (lemma 3.9.17), the maps φ and psi are Hadamard differentiable on the domain $\mathcal{C}_1 = \{(A, B) : \int |\mathrm{d}B| \leq M, A \geq \epsilon\}$, for $M, \epsilon > 0$, at every point (A, B) such that $1/A$ is of bounded variation. Hence, the map ϕ is hadamard differentiable at (w, H_0) tangentially to the set $\mathcal{C}_2 = \{(A, B) \in \mathcal{C}_1 : \psi(A, B) \geq \epsilon > 0\} \subset \overline{\mathbb{R}}^+ \times \overline{\mathbb{R}}^{+\,2}$.

Now, by lemma 3.9.17 in van der Vaart & Wellner (1997) the map $G \mapsto w$ is hadamard differentiable at G tangentially to the set of continuous functions on $\overline{\mathbb{R}}^+$, hence by delta method, $\sqrt{n}\big[\widehat{w} - w\big]$ converges weakly to the tight process $\mathcal{W} = \int_0^y \mathcal{G}(t)\,\mathrm{d}t$, where $\mathcal{G}$ is the limiting gaussian process of $\sqrt{n}\big[G - \widehat{G}\big]$. Thus, using the empirical central limit theorem, $\sqrt{n}\big(\widehat{w} - w, \widehat{H}_0 - H_0\big)$ converges to the tight zero-mean process $\big(\mathcal{W}, \mathcal{H}\big)$ in $D[0, u_{G}) \times D(\mathcal{B})$. Therefore, by the functional delta method, $\sqrt{n}\big[\widehat{F} - F\big]$ converges to the tight process

$$\phi'_{w,H_0}(\mathcal{W}, \mathcal{H}) = \frac{\varphi'_{w,H_0}(\mathcal{W}, \mathcal{H})(x, y)\,\psi(w, H_0) - \varphi(w, H_0)(x, y)\,\psi'_{w,H_0}(\mathcal{W}, \mathcal{H})}{\psi(w, H_0)^2}$$

in $D(\mathcal{A})$, where

$$\psi'_{w,H_0}(\mathcal{W}, \mathcal{H}) = \int_{[0,\infty)\times[\ell_L,\tau)} \frac{1}{w}\,\mathrm{d}\mathcal{H} - \int_{[0,\infty)\times[\ell_L,\tau)} \frac{\mathcal{W}}{w^2}\,\mathrm{d}H_0$$

and

$$\varphi'_{w,H_0}(\mathcal{W}, \mathcal{H})(x, y) = \int_{[0,x]\times[\ell_L,y]} \frac{1}{w}\,\mathrm{d}\mathcal{H} - \int_{[0,x]\times[\ell_L,y]} \frac{\mathcal{W}}{w^2}\,\mathrm{d}H_0.$$

■

Lemma 2

Under assumption B1, $\|\widehat{F}_k - F_k\| = \mathcal{O}_{a.s.}\left(\sqrt{\log\log(n)/n}\right)$, *for* $k = 1, 2$.

Proof.

The proof follows from the decomposition (20) of $\widehat{F} - F$ in the proof of Theorem 1 by using the facts that $\|\widehat{G} - G\| = \mathcal{O}_{a.s.}\left(\sqrt{\log\log(n)/n}\right)$ and $\|\widehat{H}_0 - H_0\| = \mathcal{O}_{a.s.}\left(\sqrt{\log\log(n)/n}\right)$.

■

Lemma 3

Under Theorem 2's assumptions, $\widehat{F}_1^{-1}$ *and* $\widehat{F}_2^{-1}$ *admit the representations*

$$\widehat{F}_1^{-1}(p) - F_1^{-1}(p) = \frac{1}{n}\sum_{i=1}^{n} \frac{\chi_i^F(F_1^{-1}(p), \infty)}{f_1(F_1^{-1}(p))} + r_1(p),$$

and

$$\widehat{F}_2^{-1}(p) - F_2^{-1}(p) = \frac{1}{n}\sum_{i=1}^{n} \frac{\chi_i^F(\infty, F_2^{-1}(p))}{f_2(F_2^{-1}(p))} + r_2(p),$$

where $r_1(p)$ *and* $r_2(p)$ *are uniformly of order* $\mathcal{O}_{a.s.}(n^{-3/4}\log(n)^{\beta})$*, with* $\beta > 1$*.*

Proof.

Let $[a,b] = [F_1^{-1}(p) - \epsilon, F_1^{-1}(p) + \epsilon]$, for $\epsilon > 0$. By lemma 3.9.23 in van der Vaart & Wellner (1997) the inverse map $\phi_0 : F \mapsto F^{-1}$ is hadamard differentiable at F_1 tangentially to the set of continuous functions $C[a,b]$, with derivative $\phi'_F : A \mapsto -(A/f)o\,F^{-1}$. The map ϕ'_F is linear, hence, is hadamard differentiable at F_1 tangentially to $C[a,b]$. Thus, using second order von Mises expansion of $\phi_0(\widehat{F}_1)$ (under Theorem 2's assumptions),

$$\phi_0(\widehat{F}_1) - \phi_0(F_1) = \phi'_{F_1}\left(\widehat{F}_1 - F_1\right) + \frac{1}{2}\int \varphi_2(x,y)\,\mathrm{d}\big[\widehat{F}_1(x) - F_1(x)\big]\,\mathrm{d}\big[\widehat{F}_1(y) - F_1(y)\big] + Rem_2, \quad (19)$$

where

$$\begin{aligned}\varphi_2(x,y) = {} & \frac{\mathrm{d}^2F^{-1}(p)}{\mathrm{d}p^2}\Big[p - \mathrm{I}\big(x \le F^{-1}(p)\big)\Big]\Big[p - \mathrm{I}\big(y \le F^{-1}(p)\big)\Big] \\ & + \frac{\mathrm{d}F^{-1}(p)}{\mathrm{d}p}\Big[2p - \mathrm{I}\big(x \le F^{-1}(p)\big) - \mathrm{I}\big(y \le F^{-1}(p)\big)\Big]\end{aligned}$$

is the 2^{nd} order influence function and the remainder term Rem_2 is uniformly of order $o_p(n^{-1})$ [see Fernohlz (1983, 2001) and Reeds (1976)]. By using partial integration, Lemma 2 and the oscillation result in Proposition 1 for $\widehat{F}_1$, the second term on the R.H.S. of (19) can be shown that is uniformly of order $\mathcal{O}_{a.s.}\big(n^{-3/4}\log(n)^{\beta}\big)$ ($\beta > 1$), under Theorem 2's assumptions. Hence, using the representation of $\widehat{F}_1$ in Theorem 1,

$$\widehat{F}_1^{-1}(p) - F_1^{-1}(p) = \phi_0(\widehat{F}_1) - \phi_0(F_1) = \frac{1}{n}\sum_{i=1}^{n} \frac{\chi_i^F(F_1^{-1}(p), \infty)}{f_1(F_1^{-1}(p))} + Rem_2^*,$$

where Rem_2^* is uniformly of order $\mathcal{O}_{a.s.}\big(n^{-3/4}\log(n)^{\beta}\big)$, with $\beta > 1$. The proof for the representation of $\widehat{F}_2^{-1}(p) - F_2^{-1}(p)$ is similar. ■

Lemma 4

Under Theorems 1-2's assumptions, $n^{1/2}\Big[\widehat{\mathbb{C}}_1 - \mathbb{C}\Big]$ *converges weakly to a tight process* $\mathbb{C}_L$.

Proof.

By lemma 3.9.28 in van der Vaart & Wellner (1997), the map ϕ_1 defined by $\phi_1(F)(u,v) = \mathbb{C}(u,v)$ is Hadamard differentiable at F tangentially to the set of continuous functions on $\overline{\mathbb{R}}^2$. Following similar arguments to Lemma 1's proof, $\sqrt{n}\big[\widehat{\mathbb{C}}_1 - \mathbb{C}\big]$ converges weakly to a tight process $\mathbb{C}_L$ by Lemmas 1 and 3. ■

Proof of Theorem 1.

First, notice that $F(x,y) = \widehat{F}(x,y) = \chi_i^F = 0$ for $y < \ell_L$. $\widehat{F}$ and F can then be written as

$$\widehat{F}(x,y) = \frac{\int_{\ell_L}^{y} \mathrm{d}\widehat{H}_0(x,v)/\widehat{w}(v)}{\iint_{(u,v)\in\mathcal{A}_*} \mathrm{d}\widehat{H}_0(u,v)/\widehat{w}(v)} \quad \text{and} \quad F(x,y) = \frac{\int_{\ell_L}^{y} \mathrm{d}H_0(x,v)/w(v)}{\iint_{(u,v)\in\mathcal{A}_*} \mathrm{d}H_0(u,v)/w(v)},$$

where $H_0(u,v) = F_{X,Z,\delta}(u,v,1)$, $\widehat{H}_0$ its empirical counterpart, and $\mathcal{A}_* = [0,+\infty)\times[l_L, u_L)$. By using the uniform convergence results of $\widehat{G}$ and the empirical process $\widehat{H}_0$,

$$\begin{aligned}
\widehat{F}(x,y) - F(x,y) = & \frac{1}{\mu}\int_{\ell_L}^{y} \Big[w(v) - \widehat{w}(v)\Big] \frac{\mathrm{d}\widehat{H}_0(x,v)}{w^2(v)} + \frac{1}{\mu}\int_{\ell_L}^{y} \frac{\mathrm{d}\big[\widehat{H}_0(x,v) - H_0(x,v)\big]}{w(v)} \\
& - \frac{F(x,y)}{\mu}\iint_{(u,v)\in\mathcal{A}_*} \Big[w(v) - \widehat{w}(v)\Big] \frac{\mathrm{d}\widehat{H}_0(u,v)}{w^2(v)} \\
& - \frac{F(x,y)}{\mu}\iint_{(u,v)\in\mathcal{A}_*} \frac{\mathrm{d}\big[\widehat{H}_0(u,v) - H_0(u,v)\big]}{w(v)} + r_{1,n}^{(1)}(x,y), \qquad (20)
\end{aligned}$$

hence,

$$\widehat{F}(x,y)-F(x,y)=\frac{1}{\mu}\int_{\ell_L}^{y}\Big[w(v)-\widehat{w}(v)\Big]\frac{\mathrm{d}\widehat{H}_0(x,v)}{w^2(v)}+\frac{1}{\mu}\int_{\ell_L}^{y}\frac{\mathrm{d}\widehat{H}_0(x,v)}{w(v)}$$
$$-\frac{F(x,y)}{\mu}\iint_{(u,v)\in\mathcal{A}_*}\Big[w(v)-\widehat{w}(v)\Big]\frac{\mathrm{d}\widehat{H}_0(u,v)}{w^2(v)}$$
$$-\frac{F(x,y)}{\mu}\iint_{(u,v)\in\mathcal{A}_*}\frac{\mathrm{d}\widehat{H}_0(u,v)}{w(v)}+r_{1,n}^{(1)}(x,y),$$

where $\sup_{\mathcal{A}}|r_{1,n}^{(1)}(x,y)| = \mathcal{O}_{a.s.}(n^{-1}\log\log n)$. Notice that $\mathrm{d}\big[w(v)-\widehat{w}(v)\big] = \Big[\widehat{G}(v)-G(v)\Big]\,\mathrm{d}v$. By partial integration,

$$\int_{v\le y}\big[w(v)-\widehat{w}(v)\big]\frac{\mathrm{d}\big[\widehat{H}_0(x,v)-H_0(x,v)\big]}{w^2(v)}=\int_{\ell_L}^{y}\big[w(v)-\widehat{w}(v)\big]\frac{\mathrm{d}\big[\widehat{H}_0(x,v)-H_0(x,v)\big]}{w^2(v)}$$
$$=\big[w(v)-\widehat{w}(v)\big]\frac{\big[\widehat{H}_0(x,v)-H_0(x,v)\big]}{w^2(v)}\Bigg|_{v=\ell_L}^{v=y}-\int_{\ell_L}^{y}\frac{\big[\widehat{H}_0(x,v)-H_0(x,v)\big]}{w^2(v)}\Big[\widehat{G}(v)-G(v)\Big]\,\mathrm{d}v$$
$$-\int_{\ell_L}^{y}\big[w(v)-\widehat{w}(v)\big]\big[\widehat{H}_0(x,v)-H_0(x,v)\big]\,\mathrm{d}\left(\frac{1}{w^2(v)}\right).$$

Thus, by the uniform convergence of $\widehat{G}$ and $\widehat{H}_0$,

$$\int_{\ell_L}^{y}\big[w(v)-\widehat{w}(v)\big]\frac{\mathrm{d}\widehat{H}_0(x,v)}{w^2(v)}=\int_{\ell_L}^{y}\big[w(v)-\widehat{w}(v)\big]\frac{\mathrm{d}H_0(x,v)}{w^2(v)}+r_{1,n}^{(2)}(x,y),$$

where $\sup_{\mathcal{A}}|r_{1,n}^{(2)}(x,y)| = \mathcal{O}_{a.s.}(n^{-1}\log\log n)$, and analogously,

$$\iint_{(u,v)\in\mathcal{A}_*}\big[w(v)-\widehat{w}(v)\big]\frac{\mathrm{d}\widehat{H}_0(u,v)}{w^2(v)}=\iint_{(u,v)\in\mathcal{A}_*}\big[w(v)-\widehat{w}(v)\big]\frac{\mathrm{d}H_0(u,v)}{w^2(v)}+r_{1,n}^{(3)},$$

with $r_{1,n}^{(3)} = \mathcal{O}_{a.s.}(n^{-1}\log\log n)$. Therefore,

$$\widehat{F}(x,y)-F(x,y)=\frac{1}{\mu}\int_{\ell_L}^{y}\Big[w(v)-\widehat{w}(v)\Big]\frac{\mathrm{d}H_0(x,v)}{w^2(v)}-\frac{F(x,y)}{\mu}\iint_{(u,v)\in\mathcal{A}_*}\Big[w(v)-\widehat{w}(v)\Big]\frac{\mathrm{d}H_0(u,v)}{w^2(v)}$$
$$+\frac{1}{\mu}\int_{\ell_L}^{y}\frac{\mathrm{d}\widehat{H}_0(x,v)}{w(v)}-\frac{F(x,y)}{\mu}\iint_{(u,v)\in\mathcal{A}_*}\frac{\mathrm{d}\widehat{H}_0(u,v)}{w(v)}+r_{1,n}^{(4)}(x,y),$$

i.e.,

$$\widehat{F}(x,y) - F(x,y) = \iint_{(u,v)\in\mathcal{A}_*} \frac{\mathrm{I}(u \le x, v \le y) - F(x,y)}{\mu} \Big[w(v) - \widehat{w}(v)\Big] \frac{\mathrm{d}H_0(u,v)}{w^2(v)}$$
$$+ \iint_{(u,v)\in\mathcal{A}_*} \frac{\mathrm{I}(u \le x, v \le y) - F(x,y)}{\mu} \frac{\mathrm{d}\widehat{H}_0(u,v)}{w(v)} + r^{(4)}_{1,n}(x,y),$$

where $\sup_{\mathcal{A}} |r^{(4)}_{1,n}(x,y)| = \mathcal{O}_{a.s.}\left(n^{-1}\log\log n\right)$. The result follows by using the representation of $\widehat{G}(v) - G(v)$ in Lo et al. (1989) and Lemma 1 above. ■

Proof of Proposition 1.

As indicated above, $F(x,y) = \widehat{F}(x,y) = 0$ for $y < \ell_L$, and if y_0 is such that $|y - y_0| \le a_n$ and n is sufficiently large, then $F(x_0,y_0) = \widehat{F}(x_0,y_0) = 0$.

Recall that $\widehat{F}$ and F can be written as $F(x,y) = F_0(x,y)/\mu$ and $\widehat{F}(x,y) = \widehat{F}_0(x,y)/\widehat{\mu}$, where F_0, $\widehat{F}_0$ and $\widehat{\mu}$ are defined in the Notations section. Let $\widehat{\mathbb{F}}_0 = \widehat{F}_0 - F_0$, 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 $\underline{y} = (y, y_0)$. By using the uniform convergence rate of $\widehat{\mathbb{G}}$ and $\widehat{\mathbb{H}}_0$ and by employing Taylor expansion of first order for $|x - x_0|, |y - y_0| \le a_n$, under bounded first partial derivatives of H_0, we have

$$\widehat{\mathbb{F}}(x,y) - \widehat{\mathbb{F}}(x_0,y_0) = \Big[\widehat{\mathbb{F}}_0(x,y) - \widehat{\mathbb{F}}_0(x_0,y_0)\Big]\mu^{-1} + r^{(2)}_{2,n}(\underline{x},\underline{y}) + r^{(1)}_{2,n}(\underline{x},\underline{y}),$$

where $\sup_{\substack{|x-x_0|,|y-y_0|\le a_n \\ (x,y)\in\mathcal{A}}} \left|r^{(2)}_{2,n}(\underline{x},\underline{y})\right| = \mathcal{O}_{a.s.}\left(a_n n^{-1/2}\big(\log\log n\big)^{1/2}\right)$ and $\sup_{\mathcal{A}} \left|r^{(1)}_{2,n}(\underline{x},\underline{y})\right| = \mathcal{O}_{a.s.}\left(n^{-1}\log\log n\right)$. Now, let's focus on $\widehat{\mathbb{F}}_0(x,y) - \widehat{\mathbb{F}}_0(x_0,y_0)$ on the R.H.S. of the latter equality. We have

$$\widehat{\mathbb{F}}_0(x,y) - \widehat{\mathbb{F}}_0(x_0,y_0) = \int_{\ell_L}^{y_0} \frac{1}{\widehat{w}(v)}\,\mathrm{d}\big[\widehat{\mathbb{H}}(x,v) - \widehat{\mathbb{H}}(x_0,v)\big] + \int_{y_0}^{y} \frac{1}{\widehat{w}(v)}\,\mathrm{d}\widehat{\mathbb{H}}(x,v)$$
$$+ \int_{\ell_L}^{y_0} \left[\frac{1}{\widehat{w}(v)} - \frac{1}{w(v)}\right] \mathrm{d}\big[H_0(x,v) - H_0(x_0,v)\big] + \int_{y_0}^{y} \left[\frac{1}{\widehat{w}(v)} - \frac{1}{w(v)}\right] \mathrm{d}H_0(x,v)$$

$$
\begin{aligned}
&= \int_{\ell_L}^{y_0} \frac{w(v)-\widehat{w}(v)}{w^2(v)}\, \mathrm{d}\big[H_0(x,v)-H_0(x_0,v)\big] + \int_{y_0}^{y} \frac{w(v)-\widehat{w}(v)}{w^2(v)}\, \mathrm{d}H_0(x,v) \\
&\quad + \int_{\ell_L}^{y_0} \frac{1}{\widehat{w}(v)}\, \mathrm{d}\big[\widehat{\mathbb{H}}(x,v)-\widehat{\mathbb{H}}(x_0,v)\big] + \int_{y_0}^{y} \frac{1}{\widehat{w}(v)}\, \mathrm{d}\widehat{\mathbb{H}}(x,v) + r_{2,n}^{(3)}(\underline{x},\underline{y}), \qquad (21)
\end{aligned}
$$

where $\sup_{\substack{|x-x_0|,|y-y_0|\le a_n \\ (x,y)\in\mathcal{A}}} \left|r_{2,n}^{(3)}\right| = \mathcal{O}_{a.s.}\left(n^{-1}\log\log n\right)$, by using the uniform convergence rate of $\widehat{\mathbb{G}}$. 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 (21). 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

$$
|I_n^1(\underline{x},\underline{y})| \le \left\|\frac{1}{w^2}\right\| . \|\widehat{w}-w\| . \left(\int_{\ell_L}^{y_0} \left|\frac{\partial H_0}{\partial v}(x,v) - \frac{\partial H_0}{\partial v}(x_0,v)\right| \mathrm{d}v + |H_0(x,y)-H_0(x,y_0)|\right),
$$

and by using Taylor expansion of first order for $|x-x_0|, |y-y_0| \le a_n$, under bounded first and second partial derivatives of H_0, and the uniform convergence rate of $\widehat{\mathbb{G}}$,

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

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

$$
|I_n^2(\underline{x},\underline{y})| \le 4\left\|\frac{1}{\widehat{w}}\right\| . \sup_{\substack{|x-x_0|\le a_n \\ \ell_L\le v\le y_0}} \left|\widehat{\mathbb{H}}(x,v)-\widehat{\mathbb{H}}(x_0,v)\right|,
$$

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

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

Hence, by using theorem 2.3 in Stute (1984)

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

for $k = 2, 3$, and the result follows. ■

Proof of Theorem 2. Using the oscillation result in Proposition 1 and Taylor expansion, the representation of $\widehat{\mathbb{C}}_1(u, v)$ follows from the i.i.d. representations of $\widehat{F}(x, y)$ (Theorem 1) and that of $\widehat{F}_1^{-1}(x)$ and $\widehat{F}_2^{-1}(x)$ (Lemma 3). The weak convergence of $\sqrt{n}\left[\widehat{\mathbb{C}}_1 - \mathbb{C}\right]$ follows by using Lemma 4. ■

Proof of Theorem 3.

The proof is given for representation (15) when $x, y \in \mathcal{A}_3 = [1-h, 1]$. The proof is similar for the other cases of x and y. Using partial integration, first, with respect to u and then with respect to v, we have

$$\begin{aligned}\widehat{\mathfrak{C}}(x,y) =& h^{-2}\Big\{\int_{y-h}^{1}\int_{x-h}^{1} K_{x,h}^{(1)}\left(\frac{x-u}{h}\right)K_{y,h}^{(1)}\left(\frac{y-v}{h}\right)\frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h} \\ &- \int_{y-h}^{1}\int_{x-h}^{1} \widetilde{\mathbb{C}}(1,v)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right)K_{y,h}^{(1)}\left(\frac{y-v}{h}\right)\frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h} \\ &- \int_{y-h}^{1}\int_{x-h}^{1} \widetilde{\mathbb{C}}(u,1)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right)K_{y,h}^{(1)}\left(\frac{y-v}{h}\right)\frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h} \\ &+ \int_{y-h}^{1}\int_{x-h}^{1} \widetilde{\mathbb{C}}(u,v)K_{x,h}^{(1)}\left(\frac{x-u}{h}\right)K_{y,h}^{(1)}\left(\frac{y-v}{h}\right)\frac{\mathrm{d}u}{h}\frac{\mathrm{d}v}{h}\Big\}.\end{aligned}$$

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

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

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

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

$$- \Big[\widetilde{\mathbb{C}}(1, y - vh) - \mathbb{C}(1, y - vh)\Big]\mathrm{I}(x - uh \leq 1)\Big\}\mathrm{d}K_{x,h}(u)\mathrm{d}K_{y,h}(v)$$

$$+ \iint_{[-1,1]^2} \Big[\mathfrak{C}(x - uh, y - vh) - \mathfrak{C}(x, y)\Big] K_{x,h}(u)\, K_{y,h}(v)\, \mathrm{d}u\, \mathrm{d}v.$$

By employing the i.i.d. representation of $\widehat{\mathbb{C}}_2$ in (12) and Taylor expansion of second order, the result follows by using the fact that $\int_{-1}^{1} u\, K_{x,h}(u)\, \mathrm{d}u = \int_{-1}^{1} v\, K_{y,h}(v)\, \mathrm{d}v = 0.$ ■

Proof of Theorem 4.

Note that by lemma 1 in Veraverbeke et al. (2011) the map $\phi_2 : \mathbb{C} \to 4\int_{\mathcal{B}} \mathbb{C}\, \mathrm{d}\mathbb{C} - 1$ is Hadamard-differentiable at $\mathbb{C}$ tangentially to the set of continuous functions on $\mathcal{B}$, with derivative

$$\phi'_{2,\mathbb{C}}(\xi) = 4\left\{\int \mathbb{C}\, \mathrm{d}\xi + \int \xi\, \mathrm{d}\mathbb{C}\right\}.$$

Thus, by the functional delta method

$$\sqrt{n}\big[\widehat{\tau}_{X,Y} - \tau_{X,Y}\big] = 4\sqrt{n}\big[\phi_2(\widehat{\mathbb{C}}_2) - \phi_2(\mathbb{C})\big] \xrightarrow{d} \phi'_{2,\mathbb{C}}(\mathbb{C}_L).$$

■

Proof of Theorem 5.

Analogously to the proof of lemma 1 in Veraverbeke et al. (2011), the map $\phi_3 : \mathbb{C} \to 12\int_{\mathcal{B}} u\, v\, \mathrm{d}\mathbb{C} - 3$ is Hadamard-differentiable at $\mathbb{C}$ tangentially to the set of continuous functions on $\mathcal{B}$, with derivative

$$\phi'_{3,\mathbb{C}}(\xi) = 12\int u\, v\, \mathrm{d}\xi.$$

Thus, by the functional delta method

$$\sqrt{n}\big[\widehat{\rho}_{X,Y} - \rho_{X,Y}\big] = 12\sqrt{n}\big[\phi_3(\widehat{\mathbb{C}}_2) - \phi_3(\mathbb{C})\big] \xrightarrow{d} \phi'_{3,\mathbb{C}}(\mathbb{C}_L).$$

■

Acknowledgements. This work has been supported by Canssi and Statlab-CRM (Y. Rabhi), and FQRNT and NSERC of Canada (T. Bouezmarni). The authors thank the editor, associate editor and three referees for their constructive comments. The dementia data reported in this article were collected as part of the Canadian Study of Health and Aging. The authors thank Prof. Christina Wolfson for providing the data and Prof. Masoud Asgharian for his comments.

References

Addona, V. and Wolfson, D.B. (2006). A formal test for the stationarity of the incidence rate using data from a prevalent cohort study with follow-up. *Lifetime Data Analysis* **12(3)**, 267–284.

Asgharian, M., M'lan, C. E. and Wolfson, D. B. (2002). Length-biased sampling with right-censoring: an unconditional approach. *J. Am. Statist. Assoc.* **97**, 202–209.

Bergeron, P.-J., Asgharian, M. and Wolfson, D. B. (2008). Covariate bias induced by length-biased sampling of failure times. *J. Am. Statist. Assoc.* **103**, 737–742.

Chen, S.X. and Huang, T. (2007). Nonparametric estimation of copula functions for dependent modeling. *Canadian J. of Statistics* **35**, 265-282.

Cox, D. R. (1969). Some sampling problems in technology. In *New Developments in Survey Sampling*, eds. Johnson and Smith. Wiley, New York.

de Uña-Álvarez, J. (2004). Nonparametric estimation under length-biased sampling and type I censoring: a moment based approach. *Ann. Inst. Statist. Math.* **56**, 667–681.

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.

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

Fernholz, L.T. (1983) von Mises calculus for statistical functionals. *Lecture notes in statistics* **19**. Springer-Verlag, New York.

Fernholz, L.T. (2001) On multivariate higher order von Mises expansion. *Metrika* **53**, 123-140.

Feuerverger, A. and Hall, P. (2000). Methods for density estimation in thick-slice versions of Wicksell problem. *J. Am. Statist. Assoc.* **95**, 535–546.

Fisher, R. A. (1934). The effect of methods of ascertainment upon the estimation of frequencies. *Annals of Human Genetics* **6**, 13–25, Cambridge.

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

Kvam, P. (2008). Length bias in the measurements of carbon nanotubes. *Technometrics* **50**, 462–467.

Leiva, V., Barros, M., Paula, G.A. and Sanhueza, A. (2008). Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. *Environmetrics* **19**, 235–249.

Lejeune, M. and Sarda, P. (1992). Smooth estimators of distribution and density functions. *Comput. Stat. and Dat. Analy.* **14**, 457–471.

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

Lynden-Bell, D. (1971). A method for allowing for known observational selection in small samples applied to 3CR quasars. *Mon. Not. R. Astr. Soc.* **155**, 95–118.

Neyman, J. (1955). Statistics-servant of all sciences. *Science* **122**, 401–406.

Ning, J., Qin, J. and Shen, Y. (2014). Score estimating equations from embedded likelihood functions under accelerated failure time models. *J. Am. Statist. Assoc.* **109**, 1625–1635.

Nowell, C. and Stanley, R. L. (1991). Length-biased sampling in mall intercept surveys. *Journal of Marketing Research* **28**, 475–479.

Nowell, C., Evans, M. A. and McDonald, L. (1988). Length-biased sampling in contingent valuation studies. *Land Economics* **64**, 367–371.

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.

Rabhi, Y. and Bouezmarni, T. (2016). Nonparametric inference for copula density function under random censoring. *Technical Report* 2016-150, Dept. of Mathematics, University of Sherbrooke.

Reeds, J.A. (1976). On the definition of von Mises functionals. *Ph.D. dissertation*, Harvard University, Cambridge (MA, USA).

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

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

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

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

Terwilliger, J.D., Shannon, W.D., Lathrop, G.M., Nolan, J.P., Goldin, L.R., Chase, G.A. and Weeks, D.E. (1997). True and false positive peaks in genomewide scans: applications of length-biased sampling to linkage mapping. *Am. J. Hum. Genet.* **61**, 430–438.

van der Vaart, A.W. (1991). Efficiency and Hadamard differentiability. *Scand. J. Stat.* **18**(1), 63-75.

van der Vaart, A.W. and Wellner, J.A. (1996). Weak convergence and empirical processes with applications to statistics. Springer-Verlag, New York.

Veraverbeke, N., Omelka, M. and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Stat.* **38**, 766-780.

Wang, M. C. (1991). Nonparametric estimation from cross-sectional survival data. *J. Am. Statist. Assoc.* **86**, 130–143.

Wang, W. and Wells, M.T. (2000). Estimation of Kendalls tau under censoring. *Stat. Sinica* **10**, 1199–1215.

Wicksell, S.D. (1925). The corpuscle problem, part I. *Biometrika* **17**, 84–89.

Wolfson, C., Wolfson, D.B., Asgharian M., M'lan, C. E., østbye, T., Rockwood, K. and Hogan D.B. (2001). A re-evaluation of the duration of survival after the onset of dementia. *The New England Journal of Medicine* **344**, 1111–1116.

Zelen, M. (1993). Optimal scheduling of examinations for the early detection of disease. *Biometrika* **80**, 279–293.