

Some algebraic properties of Archimedean Copula functions and their applications in the statistical estimation of the survival function

Abdushukurov Akhmedovich and Muradov Rustamjon

Department of Probability Theory and Mathematical Statistics, National University of Uzbekistan and Institute of Mathematics, Tashkent, Uzbekistan

Received: 31 December 2015, Accepted: 8 February 2016

Published online: 19 June 2016.

Abstract: In this paper we study some algebraic properties of Archimedean copulas and simple smoothed estimator of distribution function under random right censored observations in the presence of covariate. Where the dependence between a life time and a censoring variable may expressed by a given Archimedean copula. We prove an almost sure asymptotic representation which provides a key tool for obtaining weak convergence result for estimator.

Keywords: Archimedean copulas, smoothed estimator, asymptotic representation, weak convergence.

1 Introduction

What is copula? From one point a view, copula is a function which joins or "couples" a multivariate distribution function(d.f.) to its one-dimensional marginal d.f.-s. Alternatively, copula is multivariate d.f. whose one-dimensional margins are uniform on the interval $(0, 1)$. The word "copula" was first used in a mathematical or statistical sense by Sklar (1959, see, [9]). But the functions themselves predate the use of the term, appearing in the work of Hoeffding, Fréchet, Dall'Aglio, and many others(see, [10]). Over the past forty years or so, copulas have played an important role in several areas of statistics.

Why are copulas of interest to researchers of probability and statistics? As Fisher (1997, see, [3]) notes in the Encyclopedia of Statistical Sciences, "Copulas are of interest to statisticians for two main reasons: First, as a way of studying scale-free measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions, ...".

The concept of copulas in relation to multivariate distributions and dependence can be motivated in many different ways. A copula can be defined from both axiomatic and statistical points of view. Although, the emphasis in this exposition will focus on the statistical perspective, for the sake of completeness and those interested in more technical definitions. Axiomatically, a copula can be defined as follows.

Definition 1. (Axiomatic) A two-dimensional copula C is a mapping from $I^2 = [0, 1] \times [0, 1]$ to $I = [0, 1]$ which satisfies the following three conditions:

- (1) $C(u, 0) = C(0, u) = 0$ for every $u \in [0, 1]$;
- (2) $C(u, 1) = C(1, u) = u$ for every $u \in [0, 1]$;
- (3) $C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$ for every $u_1, v_1, u_2, v_2 \in [0, 1]$ satisfying $u_1 \leq u_2, v_1 \leq v_2$.

Copulas can be defined informally as follows: Let X and Y be continuous random variables (r.v.-s) with d.f.-s $F(x) = P(X \leq x)$ and $G(y) = P(Y \leq y)$, and joint distribution function $H(x, y) = P(X \leq x, Y \leq y)$. For every (x, y) in $[-\infty, \infty]^2$ consider the point in I^3 ($I = [0, 1]$) with coordinates $(F(x), G(y), H(x, y))$. This mapping from I^2 to I is a copula. Copulas are also known as dependence functions or uniform representations. Statistically, a copula can be defined as follows.

Definition 2. (Statistical) A copula $C(u, v) : [0, 1]^2 \rightarrow [0, 1]$ is a bivariate distribution function with uniform marginals.

A first example of copulas is the product copula $C(u, v) = uv$, which characterizes independent r.v.-s when the d.f.-s are continuous.

The importance of copulas in statistics is described in Sklar's Theorem. A formal presentation is provided next.

Theorem 1. (A. Sklar) Let H be a joint d.f. with margins F and G . Then there exists a copula C such that for all x, y in R ,

$$H(x, y) = C(F(x), G(y)). \quad (1)$$

If F and G are continuous, then C is unique; otherwise, C is uniquely determined on $\text{Ran}(F) \times \text{Ran}(G)$. Conversely, if C is a copula and F and G are d.f.-s, then the function H defined by (1) is a joint d.f. with margins F and G . Thus copulas link joint d.f.-s to their one-dimensional margins.

The representation (1) suggests that if the copula C were known, then substituting continuous marginal estimators for F and G would yield a plug-in estimate of their associated joint d.f. H . Moreover, in light of Sklar's result with arrive at the following functional definition of a copula.

Definition 3. (Functional) Given a bivariate d.f. H with marginals F and G , the function defined as

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)),$$

for $(u, v) \in [0, 1]^2$, where $F^{-1}(u)$ and $G^{-1}(v)$ are the inverse functions of F and G respectively, is the copula corresponding to H .

2 Some algebraic properties of Archimedean copulas

In this paper, we discuss an important class of copulas known as Archimedean copulas. These copulas find a wide range of applications for a number of reasons:

- (1) the ease with which they can be constructed;
- (2) the great variety of families of copulas which belong to this class;
- (3) the many nice properties possessed by the members of this class.

As mentioned in the introduction, Archimedean copulas originally appeared not in statistics, but rather in the study of probabilistic metric spaces, where they were studied as part of the development of a probabilistic version of the triangle inequality. For an account of this history, see Schweizer (1991, see, [10]).

Let φ be a continuous, strictly decreasing function from $[0, 1]$ to $[0, \infty]$ such that $\varphi(1) = 0$.

Definition 4. The pseudo-inverse of φ is the function $\varphi^{[-1]}$ with $Dom\varphi^{[-1]} = [0, \infty]$ and $Ran(f)$ given by

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 < t < \varphi(0), \\ 0, & \varphi(0) \leq t \leq \infty. \end{cases} \tag{2}$$

Note that $\varphi^{[-1]}$ is continuous and non increasing on $[0, \infty]$, and strictly decreasing on $[0, \varphi(0)]$. Furthermore,

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t, & 0 < t < \varphi(0), \\ 0, & \varphi(0) \leq t \leq \infty, \end{cases} = \min(t, \varphi(0)).$$

If $\varphi(0) = \infty$, then $\varphi^{[-1]} = \varphi^{-1}$ (see Fig.1).

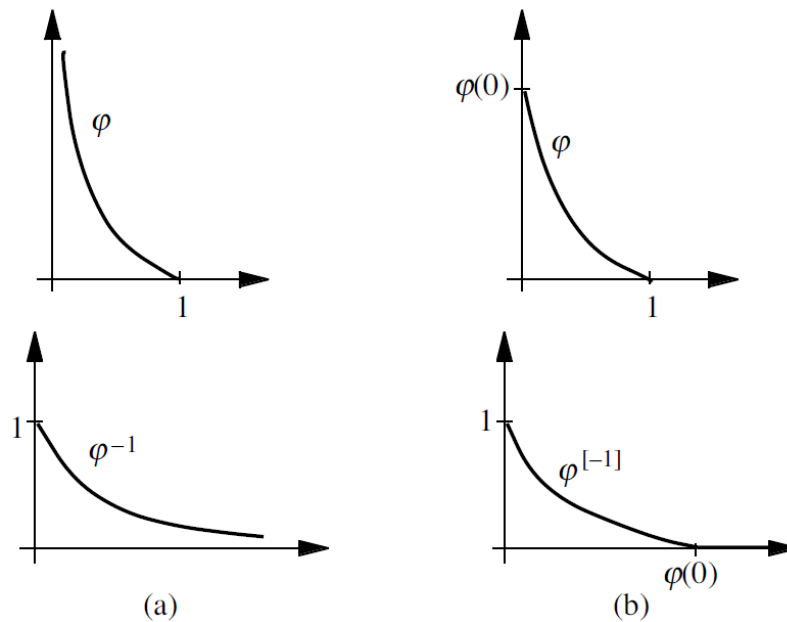


Fig. 1: Strict (a) and non-strict (b) generators and inverses.

Definition 5. Copulas of the form

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \tag{3}$$

are called Archimedean copulas, where the function φ is called a generator of the copula $\varphi(1) = 0$.

Theorem 2. [7]. Let φ be a continuous, strictly decreasing function from I to $[0, \infty]$ such that $\varphi(1) = 0$, and let $\varphi^{[-1]}$ be the pseudo-inverse of φ defined by (2). Then the function C given by (3) is a copula if and only if φ is convex.

We conclude this work with two theorems concerning some algebraic properties of Archimedean copulas.

Theorem 3. [7]. Let C be an Archimedean copula with generator φ . Then:

- (1) C is symmetric, i.e., $C(u, v) = C(v, u)$ for all u, v in I ;

- (2) C is associative, i.e., $C(C(u, v), w) = C(u, C(v, w))$ for all u, v, w in I ;
- (3) If $c > 0$ is any constant, then $c\varphi$ is also a generator of C .

For convenience, let Ω denote the set of continuous strictly decreasing convex functions φ from I to $[0, \infty]$ with $\varphi(1) = 0$. By now the reader is surely wondering about the meaning of the term "Archimedean" for these copulas. Recall the Archimedean axiom for the positive real numbers: If a, b are positive real numbers, then there exists an integer n such that $na > b$. An Archimedean copula behaves like a binary operation on the interval I , in that the copula C assigns to each pair u, v in I a number $C(u, v)$ in I . From Theorem 3, we see that the "operation" C is commutative and associative, and preserves order, i.e., $u_1 \leq u_2$ and $v_1 \leq v_2$ implies $C(u_1, v_1) \leq C(u_2, v_2)$. Algebraists call (I, C) an ordered Abelian semigroup. For any u in I , we can define the C -powers u_C^n of u recursively: $u_C^1 = u$, and $u_C^{n+1} = C(u, u_C^n)$, note that u_C^2 belongs to the diagonal section $\delta_C(u)$ of C . The version of the Archimedean axiom for (I, C) is, for any two numbers u, v in $(0, 1)$, there exists a positive integer n such that $u_C^n < v$. The next theorem shows that Archimedean copulas satisfy this version of the Archimedean axiom and hence merit their name. The term "Archimedean" for these copulas was introduced in Ling (1965, see, [7]).

Theorem 4. [7]. Let C be an Archimedean copula generated by φ in Ω . Then for any u, v in I , there exists a positive integer n such that $u_C^n < v$.

3 The random right censoring model

In such research areas as bio-medicine, engineering, insurance, social sciences and many areas researchers are interested in positive variables, which are expressed as a time until a certain event. For example, in medicine the survival time of individual, while in industrial trials, time until breakdown of a machine are non-negative random variables of interest. But in such practical situations, the observed data may be incomplete, that is censored. This is the case, for example, in medicine when the event of interest-death due to a given cause and the censoring event is death due to other cause. In industrial study, it may occur that some piece of equipment is taken away (that is censored) because it shows some sign of future failure. Moreover, the r.v.-s of interest (lifetimes, failure times) and censoring r.v.-s usually can be influenced by other variable, often called prognostic factor or covariate. In medicine, dose of a drug and in engineering some environmental conditions (temperature, pressure) are influenced to the observed variables. The basic problem consist in estimation of distribution of lifetime by such censored dependent data. The aim of paper is considering this problem in the case of right random censoring model in the presence of covariable.

Let's consider the case when the support of covariate C is the interval $[0, 1]$ and we describe our results on fixed design points $0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1$ at which we consider responses (survival or failure times) X_1, \dots, X_n and censoring times Y_1, \dots, Y_n of identical objects, which are under study. These responses are independent and nonnegative r.v.-s with conditional distribution function (d.f.) at x_i , $F_{x_i}(t) = P(X_i \leq t/C_i = x_i)$. They are subjected to random right censoring, that is for X_i there is a censoring variable Y_i with conditional d.f. $G_{x_i}(t) = P(Y_i \leq t/C_i = x_i)$ and at n -th stage of experiment the observed data is

$$S^{(n)} = \{(Z_i, \delta_i, C_i), 1 \leq i \leq n\},$$

where $Z_i = \min(X_i, Y_i)$, $\delta_i = I(X_i \leq Y_i)$ with $I(A)$ denoting the indicator of event A . Note that in sample $S^{(n)}$ r.v. X_i is observed only when $\delta_i = 1$. Commonly, in survival analysis [4] to assume independence between the r.v.-s X_i and Y_i conditional on the covariate C_i . But, in some practical situations, this assumption does not hold. Therefore, in this article

we consider a dependence model in which dependence structure is described through copula function. So let

$$S_x(t_1, t_2) = P(X_x > t_1, Y_x > t_2), t_1, t_2 \geq 0,$$

the joint survival function of the response X_x and the censoring variable Y_x at x . Then the marginal survival functions are $S_x^X(t) = 1 - F_x(t) = S_x(t, 0)$ and $S_x^Y(t) = 1 - G_x(t) = S_x(0, t), t \geq 0$. We suppose that the marginal d.f.-s F_x and G_x are continuous. Then according to the Theorem of Sklar, the joint survival function $S_x(t_1, t_2)$ can be expressed as

$$S_x(t_1, t_2) = C_x(S_x^X(t_1), S_x^Y(t_2)), t_1, t_2 \geq 0, \tag{4}$$

where $C_x(u, v)$ is a known copula function depending on x, S_x^X and S_x^Y in a general way.

4 Construction of estimator

Assume that at the fixed design value $x \in (0, 1)$, C_x in (4) is Archimedean copula, i.e.

$$S_x(t_1, t_2) = \varphi_x^{[-1]}(\varphi_x(S_x^X(t_1)) + \varphi_x(S_x^Y(t_2))), t_1, t_2 \geq 0, \tag{5}$$

where, for each $x, \varphi_x : [0, 1] \rightarrow [0, +\infty]$ is a known continuous, convex, strictly decreasing function with $\varphi_x(1) = 0$. $\varphi_x^{[-1]}$ is a pseudo-inverse of φ_x (see, Nelsen [7]) and given by

$$\varphi_x^{[-1]}(s) = \begin{cases} \varphi_x^{-1}(s), & 0 \leq s \leq \varphi_x(0), \\ 0, & \varphi_x(0) \leq s \leq \infty, \end{cases}$$

We assume that copula generator function φ_x is strict, i.e. $\varphi_x(0) = \infty$ and hence $\varphi_x^{[-1]} = \varphi_x^{-1}$. From (5), it follows that

$$\begin{aligned} P(Z_x > t) &= 1 - H_x(t) = \overline{H_x(t)} = S_x^Z(t) = S_x(t, t) = \\ &= \varphi_x^{-1}(\varphi_x(S_x^X(t)) + \varphi_x(S_x^Y(t))), t \geq 0, \end{aligned} \tag{6}$$

Let $H_x^{(1)}(t) = P(Z_x \leq t, \delta_x = 1)$ be a subdistribution function and $\Lambda_x(t)$ is crude hazard function of r.v. X_x subjecting to censoring by Y_x ,

$$\Lambda_x(dt) = \frac{P(X_x \in dt, X_x \leq Y_x)}{P(X_x \geq t, Y_x \geq t)} = \frac{H_x^{(1)}(dt)}{S_x^Z(t-)}. \tag{7}$$

From (7) one can obtain following expression of survival function S_x^X :

$$\begin{aligned} S_x^X(t) &= \varphi_x^{-1}[-\int_0^t S_x^Z(u-) \varphi_x'(S_x^Z(u)) d\Lambda_x(u)] = \\ &= \varphi_x^{-1}[-\int_0^t \varphi_x'(S_x^Z(u)) dH_x^{(1)}(u)], t \geq 0, \end{aligned} \tag{8}$$

(see, for example, [2,8]). In order to constructing the estimator of S_x^X according to representation (8), we introduce some smoothed estimators of $S_x^Z, H_x^{(1)}$ and regularity conditions for them. Similarly to Breakers and Veraverbeke [2], we will

also use the Gasser-Müller weights

$$\omega_{ni}(x, h_n) = \frac{1}{q_n(x, h_n)} \int_{x_{i-1}}^{x_i} \frac{1}{h_n} \pi\left(\frac{x-z}{h_n}\right) dz, \quad i = 1, \dots, n, \quad (9)$$

with

$$q_n(x, h_n) = \int_0^{x_n} \frac{1}{h_n} \pi\left(\frac{x-z}{h_n}\right) dz,$$

where $x_0 = 0$, π is a known probability density function(kernel) and $\{h_n, n \geq 1\}$ is a sequence of positive constants, tending to zero as $n \rightarrow \infty$, called bandwidth sequence. Let's introduce the weighted estimators of H_x, S_x^Z and $H_x^{(1)}$ respectively as

$$H_{xh}(t) = \sum_{i=1}^n \omega_{ni}(x, h_n) I(Z_i \leq t),$$

$$S_{xh}^Z(t) = 1 - H_{xh}(t), \quad (10)$$

$$H_{xh}^{(1)}(t) = \sum_{i=1}^n \omega_{ni}(x, h_n) I(Z_i \leq t, \delta_i = 1).$$

Then pluggin in (8) estimators (10) we get corresponding estimator of $S_x^X(t)$ as

$$S_{xh}^X(t) = 1 - F_{xh}(t) = \varphi_x^{-1} \left[- \int_0^t \varphi'_x(S_{xh}^Z(u)) dH_{xh}^{(1)}(u) \right], \quad t \geq 0, \quad (11)$$

Remark that in the case of no covariate, estimator (11) reduces to estimator first obtained by Zeng and Klein [12]. In the case of the independent copula $\varphi(y) = -\log y$, Zeng and Klein estimate reduces to a exponential-hazard estimate (see, [1,11]). Also it is well-known that under independent censoring case Kaplan-Meier's product-limit estimator and exponential-hazard estimators are asymptotical equivalent. Therefore, we will show that estimator (11) and copula-graphic estimator of Breakers and Veraverbeke have the same asymptotic behaviours.

5 Asymptotic results

For the design points x_1, \dots, x_n , denote

$$\underline{\Delta}_n = \min_{1 \leq i \leq n} (x_i - x_{i-1}), \quad \overline{\Delta}_n = \max_{1 \leq i \leq n} (x_i - x_{i-1}).$$

For the kernel π , let

$$\|\pi\|_2^2 = \int_{-\infty}^{\infty} \pi^2(u) du, \quad m_v(\pi) = \int_{-\infty}^{\infty} u^v \pi(u) du, \quad v = 1, 2,$$

$$\|\pi\|_{\infty} = \sup_{u \in R} \pi(u).$$

Moreover, we use next assumptions on the design and on the kernel function:

(A1) As $n \rightarrow \infty$, $x_n \rightarrow 1$, $\overline{\Delta}_n = O(\frac{1}{n})$, $\overline{\Delta}_n - \underline{\Delta}_n = o(\frac{1}{n})$.

(A2) π is a probability density function with compact support $[-M, M]$ for some $M > 0$, with $m_1(\pi) = 0$ and $|\pi(u) - \pi(u')| \leq C(\pi)|u - u'|$, where $C(\pi)$ is some constant. Let $T_{H_x} = \inf\{t \geq 0 : H_x(t) = 1\}$. Then $T_{H_x} = \min(T_{F_x}, T_{G_x})$. For our results we need some smoothness conditions on functions $H_x(t)$ and $H_x^{(1)}(t)$. We formulate them for a general (sub)distribution function $N_x(t)$, $0 \leq x \leq 1, t \in R$ and for a fixed $T > 0$.

(A3) $\frac{\partial}{\partial x} N_x(t) = \dot{N}_x(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.

- (A4) $\frac{\partial}{\partial t} N_x(t) = N'_x(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.
- (A5) $\frac{\partial^2}{\partial x^2} N_x(t) = \dot{N}_x(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.
- (A6) $\frac{\partial^2}{\partial t^2} N_x(t) = N''_x(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.
- (A7) $\frac{\partial^2}{\partial x \partial t} N_x(t) = \dot{N}'_x(t)$ exists and is continuous in $(x, t) \in [0, 1] \times [0, T]$.
- (A8) $\frac{\partial \varphi_x(u)}{\partial u} = \varphi'_x(u)$ and $\frac{\partial^2 \varphi_x(u)}{\partial u^2} = \varphi''_x(u)$ are Lipschitz in the x -direction with a bounded Lipschitz constant and $\frac{\partial^3 \varphi_x(u)}{\partial u^3} = \varphi'''_x(u) \leq 0$ exists and is continuous in $(x, u) \in [0, 1] \times (0, 1]$.

It is clear that for existence of right hand side of representation (8) we must require the conditions (A 4) for functions $H_x(t)$ and $H_x^{(1)}(t)$ in $[0, 1] \times [0, T]$ with $T < T_{H_x}$ and existence of $\varphi'_x(u)$ on $[0, 1] \times (0, 1]$.

We derive an almost sure representation result with rate.

Theorem 5. Assume (A 1), (A 2), $H_x(t)$ and $H_x^{(1)}(t)$ satisfy (A 5)-(A 7) in $[0, T]$ with $T < T_{H_x}$, φ_x satisfies (A 8) and $h_n \rightarrow \infty$, $\frac{\log n}{nh_n} \rightarrow 0$, $\frac{nh_n^5}{\log n} = O(1)$. Then, as $n \rightarrow \infty$,

$$F_{xh}(t) - F_x(t) = \sum_{i=1}^n \omega_{ni}(x, h_n) \Psi_{tx}(Z_i, \delta_i) + r_n(t),$$

where

$$\Psi_{tx}(Z_i, \delta_i) = \frac{-1}{\varphi'_x(S_x^X(t))} \left[\int_0^t \varphi''_x(S_x^Z(u)) (I(Z_i \leq u) - H_x(u)) dH_x^{(1)}(u) - \varphi'_x(S_x^Z(t)) (I(Z_i \leq t, \delta_i = 1) - H_x^{(1)}(t)) - \int_0^t \varphi''_x(S_x^Z(u)) (I(Z_i \leq u, \delta_i = 1) - H_x^{(1)}(u)) dH_x(u) \right]$$

and

$$\sup_{0 \leq t \leq T} |r_n(t)| \stackrel{a.s.}{=} O\left(\left(\frac{\log n}{nh_n}\right)^{3/4}\right).$$

The weak convergence of the empirical process $(nh_n)^{1/2} \{F_{xh}(\cdot) - F_x(\cdot)\}$ in the space $\ell^\infty[0, T]$ of uniformly bounded functions on $[0, T]$, endowed with the uniform topology is the contents of the next theorem.

Theorem 6. Assume (A 1), (A 2), $H_x(t)$ and $H_x^{(1)}(t)$ satisfy (A 5)-(A 7) in $[0, T]$ with $T < T_{H_x}$, and that φ_x satisfies (A 8). (I) If $nh_n^5 \rightarrow 0$ and $\frac{(\log n)^3}{nh_n} \rightarrow 0$, then, as $n \rightarrow \infty$,

$$(nh_n)^{1/2} \{F_{xh}(\cdot) - F_x(\cdot)\} \implies \mathbf{W}_x(\cdot) \text{ in } \ell^\infty[0, T].$$

(II) If $h_n = Cn^{-1/5}$ for some $C > 0$, then, as $n \rightarrow \infty$,

$$(nh_n)^{1/2} \{F_{xh}(\cdot) - F_x(\cdot)\} \implies \mathbf{W}_x^*(\cdot) \text{ in } \ell^\infty[0, T],$$

where $\mathbf{W}_x(\cdot)$ and $\mathbf{W}_x^*(\cdot)$ are Gaussian processes with means

$$E\mathbf{W}_x(t) = 0, \quad E\mathbf{W}_x^*(t) = a_x(t),$$

and same covariance

$$\text{Cov}(\mathbf{W}_x(t), \mathbf{W}_x^*(s)) = \text{Cov}(\mathbf{W}_x^*(t), \mathbf{W}_x^*(s)) = \Gamma_x(t, s),$$

with

$$a_x(t) = \frac{-C^{5/2} m_2(\pi)}{2\varphi'_x(S_x^X(t))} \int_0^t [\varphi''_x(S_x^Z(u)) \dot{H}_x(u) dH_x^{(1)}(u) - \varphi'_x(S_x^Z(u)) d\dot{H}_x^{(1)}(u)],$$

and

$$\begin{aligned} \Gamma_x(t, s) &= \frac{\|\pi\|_2^2}{\varphi'_x(S_x^X(t))\varphi'_x(S_x^X(s))} \left\{ \int_0^{\min(t,s)} (\varphi'_x(S_x^Z(z)))^2 dH_x^{(1)}(z) \right. \\ &+ \int_0^{\min(t,s)} [\varphi''_x(S_x^Z(w))S_x^Z(w) + \varphi'_x(S_x^Z(w))] \int_0^w \varphi''_x(S_x^Z(y)) dH_x^{(1)}(y) dH_x^{(1)}(w) \\ &+ \int_0^{\min(t,s)} \varphi''_x(S_x^Z(w)) \int_w^{\max(t,s)} (\varphi''_x(S_x^Z(y))S_x^Z(y) + \varphi'_x(S_x^Z(y))) dH_x^{(1)}(y) dH_x^{(1)}(w) \\ &\left. - \int_0^t [\varphi''_x(S_x^Z(y))S_x^Z(y) + \varphi'_x(S_x^Z(y))] dH_x^{(1)}(y) \int_0^s [\varphi''_x(S_x^Z(w))S_x^Z(w) + \varphi'_x(S_x^Z(w))] dH_x^{(1)}(w) \right\}. \end{aligned}$$

6 Proofs of Theorems 5 and 6

In order to proving the theorems 5 and 6 we need some auxiliary results for empiricals H_{xh} and $H_{xh}^{(1)}$. While the Lemma 1 below (i.e. Lemma A4 from [11]) about the rates of strong uniform consistency of weighted empiricals is formulated only for H_{xh} , it is still true also for $H_{xh}^{(1)}$ and proved exactly with the same way.

Lemma 1. [11].

(I) Assume (A 1), (A 2), $H_x(t)$ satisfies (A 3), $h_n \rightarrow 0$, $nh_n \rightarrow \infty$, $\frac{nh_n^3}{\log n} = O(1)$. Then, as $n \rightarrow \infty$,

$$\sup_{0 \leq t \leq T} |H_{xh}(t) - H_x(t)| \stackrel{a.s.}{=} O\left(\left(\frac{\log n}{nh_n}\right)^{1/2}\right).$$

(I) Assume (A 1), (A 2), $H_x(t)$ satisfies (A 3) and (A 5), $h_n \rightarrow 0$, $\frac{nh_n^5}{\log n} = O(1)$. Then, as $n \rightarrow \infty$,

$$\sup_{0 \leq t \leq T} |H_{xh}(t) - H_x(t)| \stackrel{a.s.}{=} O\left(\left(\frac{\log n}{nh_n}\right)^{1/2}\right).$$

The next Lemma 2 (Lemma 2 in [2]) provides the convergence rate of Theorem 5.

Lemma 2. [2]. Under the conditions of theorem 5, as $n \rightarrow \infty$,

$$\sup_{0 \leq t \leq T} \left| - \int_0^t [\varphi'_x(S_{xh}^Z(u)) - \varphi'_x(S_x^Z(u))] d(H_{xh}^{(1)}(u) - H_x^{(1)}(u)) \right| \stackrel{a.s.}{=} O\left(\left(\frac{\log n}{nh_n}\right)^{3/4}\right).$$

Proof. (Proof of Theorem 5) Applying a second order Taylor expansion, we have

$$\begin{aligned} F_{xh}(t) - F_x(t) &= -(S_{xh}^X(t) - S_x^X(t)) \\ &= -\left\{ \varphi_x^{-1} \left[- \int_0^t \varphi'_x(S_{xh}^Z(u)) dH_{xh}^{(1)}(u) \right] - \varphi_x^{-1} \left[- \int_0^t \varphi'_x(S_x^Z(u)) dH_x^{(1)}(u) \right] \right\} \\ &= -\frac{1}{\varphi'_x(S_x^X(t))} \left\{ - \int_0^t \varphi'_x(S_{xh}^Z(u)) dH_{xh}^{(1)}(u) + \int_0^t \varphi'_x(S_x^Z(u)) dH_x^{(1)}(u) \right\} \\ &+ \frac{\varphi''_x(\varphi_x^{-1}(\theta_{xh}(t)))}{2[\varphi'_x(\varphi_x^{-1}(\theta_{xh}(t)))]^2} \left\{ - \int_0^t \varphi'_x(S_{xh}^Z(u)) dH_{xh}^{(1)}(u) + \int_0^t \varphi'_x(S_x^Z(u)) dH_x^{(1)}(u) \right\}^2 \\ &= A_n(t) + B_n(t), \end{aligned} \tag{12}$$

where $\theta_{xh}(t)$ between $[-\int_0^t \phi'_x(S_{xh}^Z(u))dH_{xh}^{(1)}(u)]$ and $[-\int_0^t \phi'_x(S_x^Z(u))dH_x^{(1)}(u)]$. In (12) the first summand we rewrite as

$$A_n(t) = -\frac{1}{\phi'_x(S_x^Z(t))} [Q_{n1}(t) + Q_{n2}(t) + Q_{n3}(t)], \tag{13}$$

where

$$Q_{n1}(t) = -\int_0^t [\phi'_x(S_{xh}^Z(u)) - \phi'_x(S_x^Z(u))]dH_x^{(1)}(u),$$

$$Q_{n2}(t) = -\int_0^t \phi'_x(S_{xh}^Z(u))d(H_{xh}^{(1)}(u) - H_x^{(1)}(u)),$$

and

$$Q_{n3}(t) = -\int_0^t [\phi'_x(S_{xh}^Z(u)) - \phi'_x(S_x^Z(u))]d(H_{xh}^{(1)}(u) - H_x^{(1)}(u)).$$

From Lemma 2, we get

$$\sup_{0 \leq t \leq T} |Q_{n3}(t)| \stackrel{\text{a.s.}}{=} O\left(\left(\frac{\log n}{nh_n}\right)^{3/4}\right). \tag{14}$$

Furthermore, for $0 \leq t \leq T < T_{H_x}$, also by Taylor expansion,

$$Q_{n1}(t) = \int_0^t \phi''_x(S_x^Z(u))(H_{xh}(u) - H_x(u))dH_x^{(1)}(u) - \int_0^t \frac{1}{2} \phi''_x(\eta_{xh}(u))(H_{xh}(u) - H_x(u))^2dH_x^{(1)}(u)$$

$$+ \int_0^t \phi''_x(S_x^Z(u))(H_{xh}(u) - H_x(u))dH_x^{(1)}(u) + q_n(t), \tag{15}$$

where $\eta_{xh}(u) \in [\min(H_{xh}(u), H_x(u)), \max(H_{xh}(u), H_x(u))]$ and from Lemma 1,

$$\sup_{0 \leq t \leq T} |q_n(t)| \stackrel{\text{a.s.}}{=} O\left(\frac{\log n}{nh_n}\right). \tag{16}$$

Integrating by parts, we rewrite $Q_{n2}(t)$ as

$$Q_{n2}(t) = -\phi'_x(S_x^Z(t))(H_{xh}^{(1)}(t) - H_x^{(1)}(t)) + \int_0^t \phi''_x(S_x^Z(u))(H_{xh}^{(1)}(u) - H_x^{(1)}(u))dH_x(u). \tag{17}$$

Therefore, from (10)-(14), and Lemma 1, we have

$$\sup_{0 \leq t \leq T} |A_n(t)| \stackrel{\text{a.s.}}{=} O\left(\left(\frac{\log n}{nh_n}\right)^{1/2}\right). \tag{18}$$

Since,

$$\sup_{0 \leq t \leq T} |B_n(t)| \stackrel{\text{a.s.}}{=} O\left(\left(\sup_{0 \leq t \leq T} |A_n(t)|\right)^2\right), \tag{19}$$

hence, from (15)

$$\sup_{0 \leq t \leq T} |B_n(t)| \stackrel{\text{a.s.}}{=} O\left(\frac{\log n}{nh_n}\right). \tag{20}$$

Then, finally from (9)-(17), we obtain that for $0 \leq t \leq T < T_{H_x}$, as $n \rightarrow \infty$,

$$\begin{aligned}
 F_{xh}(t) - F_x(t) \stackrel{\text{a.s.}}{=} & -\frac{1}{\varphi'_x(S_x^Z(t))} \left\{ \int_0^t \varphi''_x(S_x^Z(u))(H_{xh}(u) - H_x(u))dH_x^{(1)}(u) - \right. \\
 & \left. - \varphi'_x(S_x^Z(t))(H_{xh}^{(1)}(t) - H_x^{(1)}(t)) + \int_0^t \varphi''_x(S_x^Z(u))(H_{xh}^{(1)}(u) - H_x^{(1)}(u))dH_x(u) \right. \\
 & \left. + O\left(\left(\frac{\log n}{nh_n}\right)^{3/4}\right) \right\} = \sum_{i=1}^n \omega_{ni}(x, h_n) \Psi_{tx}(Z_i, \delta_i) + O\left(\left(\frac{\log n}{nh_n}\right)^{3/4}\right),
 \end{aligned}$$

which completes the proof of Theorem 5.

It is necessary to note that almost sure representation of Theorem 5 plays a key role on investigating of estimator (11) and, in particular, it provides a basic tool for obtaining weak convergence result of Theorem 6. But the main summand Ψ_{tx} of this representation is the same as in the case of copula-graphic estimator from [2]. Then the proof of Theorem 6 one can accomponing by line of proof of Theorem 2 from [2]. Therefore, the proof of Theorem 6 is omitted. Thus, the estimator (11) and copula-graphic estimator are asymptotic equivalent.

7 Conclusions

It is necessary to note that in the case of no covariates, this idea first was considered by Zeng and Klein [12] and proposed copula-graphic estimator. Rivest and Wells [8] investigated copula-graphic estimator and derived a closed form expression for estimator when the joint survival function (1) is modeled an Archimedean copula. The copula-graphic estimator is then shown to be uniformly consistent and asymptotically normal. Note that the copula-graphic estimator is equivalent to the product-limit estimator of Kaplan and Meier [5] when the survival and censoring times are assumed to be independent. Braekers and Veraverbeke [2] extend copula-graphic estimator to the fixed design regression case and show that estimator has an asymptotic representation and a Gaussian limit. We consider other estimator of d.f. F_x which had a simpler form than copula-graphic estimator and it is also equivalent to the usual exponential-hazard estimator under independent censoring case. We study the large sample properties of estimator proposed and present result of uniform normality with the same limiting Gaussian process as for copula-graphic estimator.

Acknowledgements

This work was supported by the fundamental grant F4-01 Republic of Uzbekistan.

References

- [1] Abdushukurov A.A., Muradov R.S., *Estimation of survival and mean residual life functions from dependent random censored data*, New Trends in Mathematical Sciences, (2014).
- [2] Breakers R., Veraverbeke N., *A copula-graphic estimator for the conditional survival function under dependent censoring*, The Canadian Journal of Statistics, (2005).
- [3] Fisher N.I., *Copulas*, Encyclopedia of Statistical Sciences, John Wiley & Sons, (1997).
- [4] Fleming T.R., Harrington D.P., *Counting Processes and Survival Analysis*, Wiley. New York, (1991).
- [5] Kaplan E.L., Meier P., *Nonparametric estimation from incomplete observations*, Journal of American Statistical Association, (1958).

- [6] Muradov R.S., Abdushukurov A.A., *Estimation of multivariate distributions and its mixtures*, LAMBERT Academic Publishing. LAP, (In Russian)(2011).
- [7] Nelsen R.B., *An Introduction to Copulas*, Springer, New York, (2006).
- [8] Rivest L.P., Wells M.T., *A martingall aproach to the copula-graphic estimator for the survival function under dependent censoring*, Journal of Multivariate Analysis, (2001).
- [9] Sklar A., *Fonctions de repartition à n dimensions et leurs marges*, Publications de l'Institut de Statistique de l'Université de Paris, (1959).
- [10] Schweizer B., *Advances in Probability Distributions With Given Marginals*, Vol. 67 of Mathematics and its applications, Thirty years of copula, Kluwer Academic Publishers, (1991).
- [11] Van.I.Keilegom, Veraverbeke N., *Estimation and the bootstrap with censored data in fixed design nonparametric regression*, Ann. Inst. Statist. Math., (1997).
- [12] Zeng M., Klein J.P., *Estimates of marginal survival for dependent competing risks based on an assumed copula*, Biometrika, (1995).