ADV MATH SCI JOURNAL

Advances in Mathematics: Scientific Journal **10** (2021), no.6, 2847–2864
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.10.6.12

SURVIVAL COPULA PARAMETERS ESTIMATION FOR ARCHIMEDEAN FAMILY UNDER SINGLY CENSORING

Nesrine Idiou[1] and Fatah Benatia

ABSTRACT. Given $(Z_i, \delta_i) = \left\{\min(T_i, C_i), I_{(T_i<C_i)_{i=1,2}}\right\}$, as dependent or independent right-censored variables, general formulas are proven for a semi-parametric estimation of the proposed method. As a logical continuation of results established by N.IDIOU et al 2021 [16], a new estimator of $\tilde{C}$ is proposed by considering that the underlying copula is Archimedean, under singly censoring data. As an application, two Archimedean copulas models have been chosen to illustrate our theoretical results. A simulation study follows, which sheds light on the behavior of the process estimation method shown that the proposed estimator performs well in terms of relative bias and RMSE. The methodology of the proposed estimator is also illustrated by using lifetime data from the Diabetic Retinopathy Study, where its efficiency and robustness are observed.

1. INTRODUCTION

In medical domain, researchers were mostly confronted with competing risk issues, that is, event times may be dependent and they are censoring each other [1–3, 10–12]. Likewise, in survival analyses, it is popular to observe two or

[1] *Nesrine IDIOU*

2020 *Mathematics Subject Classification.* Primary 62G05; Secondary 62G20.

Key words and phrases. Archimedean copulas; Individually censoring data; Moment estimator; Survival copula; Semi-parametric estimation.

Submitted: 30.05.2021; *Accepted:* 14.06.2021; *Published:* 21.06.2021.

more lifetimes for the same customer, patient, or equipment. For example, the lifetimes of a pair of organs can be observed in a pair of kidneys, an ear, or an eye in patients; or the lifetimes of engines in a two-engine vehicle. In most cases, these variables are related and this pattern of bivariate data is well-suited to the copula model, particularly the Archimedean copula models.

As an outcome, we suggest that the two failure times T_1 and T_2 can be modelled by an Archimedean copula model and it is subject to dependence or independence right-censoring with the censoring vector (C_1, C_2), we also propose that the vector (C_1, C_2) follows an arbitrary bivariate continuous distribution. Hence, we can only observe $Z_i = \min(T_i, C_i)$, $\delta_i = I_{\{T_i \leq C_i\}_{i=1,2}}$ where $I_{(.)}$ represents the indicator function. Sometimes, the problem in right-censoring is how modeling the dependence concept among a bivariate censoring vectors (T_1, T_2) and (C_1, C_2), when both variables are censored at the same time (see N.Idiou et al [16]). The issue now is how to construct the dependency structure between this vector when only one variable is right-censored. Let's look at the bivariate pattern (T_1, T_2), with the joint distribution function (df) $F(t_1, t_2) = P(T_1 \leq t_1, T_2 \leq t_2)$, which can be presented by the following form $F(t_1, t_2) = C(F_1(t_1), F_2(t_2))$, where F_1, F_2 are continuous margins and C is the related copula function ordinarily known for all (u, v) in $[0, 1]^2$ by:

$$C(u, v) = F((F_1^{-1}(u), F_2^{-1}(v)),$$

when $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}$ is the generalized inverse of a non-decreasing function F. A joint survival function S of (T_1, T_2) is said to have an Archimedean association dependence structure if for all $t_1, t_2 \geq 0$, it can be interpreted as follows $S(t_1, t_2) = \varphi^{-1}(\varphi(S_1(t_1)) + \varphi(S_2(t_2)))$, where φ is a continuous and convex function defined on $[0, 1] \to [0, \infty]$ with $\varphi(1) = 0$, and S_1, S_2 are the marginal survival functions of T_1 and T_2 respectively (see [4], [5]) and [6]. In the context of multivariate survival analysis, many models have been proposed to model multivariate survival data among them, Archimedean copulas models (see [2], [19] and [20]). Specifically, for the couple $(u, v) \in [0, 1]^2$ an Archimedean copula is noted as $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$, where φ^{-1} is the inverse function of φ and φ usually called the Archimedean generator of C. Hence, the function $\tilde{C}$ define from $[0, 1]^2 \to [0, 1]$, which couples S_1 and S_2 and known by the survival copula of (T_1, T_2) via

$$\tilde{C}(u, v) = u + v - 1 + C(1 - u, 1 - v), \tag{1.1}$$

(Nelsen, (2006) [15]). Supposing that (T_1, T_2) follows an Archimedean copula, where $S_1$and S_2 are the marginal survival functions, Genest and Rivest (1993) [6] have proved that $U = \frac{\varphi(S_1(T_1))}{\varphi(S_1(T_1))+\varphi(S_2(T_2))}$ and $V = \tilde{C}(S_1(T_1), S_2(T_2)) = \varphi^{-1}(\varphi(S_1(T_1)) + \varphi(S_2(T_2)))$, are independently distributed random variables when U follows a uniform distribution on $[0,1]$ and V follows a so-called Kendall distribution with the density function: $k_C(t) = \frac{\varphi(t)\varphi''(t)}{(\varphi'(t))^2}$ defined on $(0,1]$, as a function of t depends on the unidentified parameter θ.

The main aim of this paper is to present a new semi-parametric estimation procedure and its application to health-related survival data, given (T_1, T_2) as individually censored. General formulas for all possible parameters estimate of a survival copula $\tilde{C}$ are also presented under the assumption that the copula is Archimedean.

Important results are reviewed in section 2, where general formulas are proposed for the marginal survival functions of T_1 and T_2. As an application of our results, a simple way of the estimation of the unknown parameters is declared in section 3, where an estimator of V based on the classical moments method is proposed, followed by two examples of Clayton and Gumbel copula models. Under the Archimedean dependence structure assumption for censored data, a simulation study evaluates the performance of our estimator presented in Section 4, relatively on bias and RMSE; where the robustness and efficiency of the estimator proven. In section 5, we illustrate the methodology presented in section 3 on real data from the Diabetic Retinopathy Study, which is available in the "survival" package [22], [23] of the R software. Our paper ends with some discussions in Section 6.

2. IMPORTANT RESULTS

Assume that (T_1, T_2) are two positive random variables whose distributions can be modelled by an Archimedean copula either dependently or independently right-censored by a censoring vector (C_1, C_2) that follows an arbitrary bivariate continuous distribution. Take the available observation in the case of absence data $(Z_{1i}, Z_{2i}, \delta_{1i}, \delta_{2i})_{1\le i\le n}$: the independent copies of a non-negative random variable of the vector $(Z_1, Z_2, \delta_1, \delta_2)$. As a result, the variable $Z_i = \min(T_i, C_i)$ is only observed when $T_i \le C_i$ for $i = 1, 2$, then $\delta_i = I_{\{T_i \le C_i\}_{i=1,2}}$ equal to one which represents the indicator function of censored data.

Considering only T_1 is right-censored, in other words, $C_2 = \infty$ almost surely, which is a particular situation from another case of doubly right-censored (see M.Boukeloua 2020 [14]). In this case, the empirical distribution function is $\tilde{F}_n(t_1, t_2) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{T_{1i}\leq t_1, T_{2i}\leq t_2\}}$. As per Stute (1993) [18], the relating consistent estimator of F was given by

$$\tilde{F}_n(t_1, t_2) = \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_{1i}}{\hat{S}_1\left(Z_{1i}^{-}\right)}1_{\{Z_{1i}\leq t_1, Z_{2i}\leq t_2\}}, \tag{2.1}$$

recognizing it was provided that $\hat{S}_1(t) = \prod_{k/Z'_{1k}<t}(1-\frac{\sum_{i=1}^{n}1_{\{Z_{1i}=Z'_{1k},\delta_{1i}=0\}}}{\sum_{i=1}^{n}1_{\{Z_{1i}\geq Z'_{1k}\}}})$, where $\hat{S}_1$ as the Kaplan-Meier estimate of S_1 and $((Z'_{1k})_{1\leq k\leq m}, m\leq n)$ is the distinct values of $(Z_{1i})_{1\leq i\leq n}$. Suppose that the copula C is twice continuously differentiable and the variable T_1's support is lower than the variable T_2's support. Following Gribkova and Lopez (2015) [9] and noted that $F_{1n}(t_1) = \lim_{t_2\to\infty} F_n(t_1, t_2)$, $F_{2n}(t_2) = \lim_{t_1\to\infty} F_n(t_1, t_2)$, the empirical copula function C_n have estimated by:

$$C_n(u, v) = \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_{1i}}{\hat{S}_1\left(Z_{1i}^{-}\right)}1_{\{F_{1n}(Z_{1i})\leq u, F_{2n}(Z_{2i})\leq v\}}, (u, v)\in[0,1]^2,$$

The weak convergence of C_n has proved under some assumptions (see [16]). Hence, the empirical survival copula of such form:

$$\tilde{C}_n(u, v) \;=\; u + v - 1 + \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_{1i}}{\hat{S}_1\left(Z_{1i}^{-}\right)}1_{\{\bar{F}_{1n}(Z_{1i})\geq u, \bar{F}_{2n}(Z_{2i})\geq v\}},$$

where $(u, v)\in[0,1]^2$. Following [16], the asympthotic normality of the empirical survival copula $\tilde{C}_n$, has proved under some assumptions in Theorem 2. Because the dependence between T_i and C_i, $i = 1, 2$ can be modeled by an Archimedean copula, Wang and Oakes(2008) proved that the distribution function of V formulated by

$$F(v, c_1, c_2) = \frac{1}{\tilde{C}(c_1, c_2)}\left\{v - \frac{\varphi(v) - \varphi\left(\tilde{C}(c_1, c_2)\right)}{\varphi'(v)}\right\},\ 0\leq v\leq \tilde{C}(c_1, c_2), \tag{2.2}$$

where T_1 and T_2 are both right-censored [21]. By analogy, when only one variable is censored the distribution function of V become as follows:

- $F_1(v, c_1, t_2) = \frac{\varphi'\left(\tilde{C}(c_1,t_2)\right)}{\varphi'(v)}, 0 \leq v \leq \tilde{C}(c_1, t_2)$, when only T_1 is right-censored;

- $F_2(v, t_1, c_2) = \frac{\varphi'\left(\tilde{C}(t_1,c_2)\right)}{\varphi'(v)}, 0 \leq v \leq \tilde{C}\left(t_1, c_2\right),$ when only T_2 is right-censored.

Proof. see [21]. ☐

By the way used 2.2, the k^{th} moments of V in the case of doubly right censoring have established by:

$$\mathbb{E}(V^k \big| T_1 > c_1, T_2 > c_2) = \frac{\left(\tilde{C}\left(c_1, c_2\right)\right)^k}{k+1} - k\left(\tilde{C}\left(c_1, c_2\right)\right)^{k-1} \varphi\left(\tilde{C}\left(c_1, c_2\right)\right) \int_0^1 \frac{v^{k-1}}{\varphi'\left(v\tilde{C}\left(c_1, c_2\right)\right)} dv + k\left(\tilde{C}\left(c_1, c_2\right)\right)^{k-1} \int_0^1 \frac{v^{k-1}\varphi\left(v\tilde{C}\left(c_1, c_2\right)\right)}{\varphi'\left(v\tilde{C}\left(c_1, c_2\right)\right)} dv, \; k \geq 1 \tag{2.3}$$

Proof. see [16]. ☐

Based on the results discussed previously, we can show

Corollary 2.1. *Let (T_1, T_2) be a random pair whose distribution can be modelled by an Archimedean copula. Assuming that (T_1, T_2) is subject to dependent or independent right censoring by a censoring vector (C_1, C_2) that follows an arbitrary bivariate continuous distribution, then we have:*

(1) *For $k \geq 1$, the k^{th} moments of V when only T_1 is right-censored is*

$$\mathbb{E}(V^k \big| T_1 > c_1, T_2 = t_2) = \left(\tilde{C}\left(c_1, t_2\right)\right)^k - k\left(\tilde{C}\left(c_1, t_2\right)\right)^k \varphi'\left(\tilde{C}\left(c_1, t_2\right)\right) \int_0^1 \frac{v^{k-1}}{\varphi'\left(v\tilde{C}\left(c_1, t_2\right)\right)} dv.$$

(2) *For $k \geq 1$, the k^{th} moments of V when only T_2 is right-censored is*

$$\mathbb{E}(V^k \big| T_1 = t_1, T_2 > c_2) = \left(\tilde{C}\left(t_1, c_2\right)\right)^k - k\left(\tilde{C}\left(t_1, c_2\right)\right)^k \varphi'\left(\tilde{C}\left(t_1, c_2\right)\right) \int_0^1 \frac{v^{k-1}}{\varphi'\left(v\tilde{C}\left(t_1, c_2\right)\right)} dv.$$

Proof. For $k > 1$ the k^{th} moments is defined by:

$$\mathbb{E}(V^k \big| T_1 > c_1, T_2 = t_2) = \int_0^{\tilde{C}(c_1,t_2)} v^k dF_1\left(v, c_1, t_2\right),$$

based on theorem given by Wang, we use the conditional distribution of V when only T_1 is censored $(V|\, T_1 > c_1, T_2 = t_2)$, we have

$$\begin{aligned}\mathbb{E}(V^k \big| T_1 &> c_1, T_2 = t_2) = \int_0^{\tilde{C}(c_1,t_2)} v^k dF_1\left(v, c_1, t_2\right) \\ &= \int_0^{\tilde{C}(c_1,t_2)} v^k \left\{ \frac{-\varphi''\left(v\right)\varphi'\left(\tilde{C}\left(c_1, t_2\right)\right)}{\left(\varphi'\left(v\right)\right)^2} \right\} dv \\ &= I\end{aligned}$$

To simplify I we pass directly to integration by parts, and we have:

$$I = \left(\left[v^k \frac{\varphi'\left(\tilde{C}\left(c_1, t_2\right)\right)}{\varphi'\left(v\right)} \right]_0^{\tilde{C}(c_1,t_2)} - k \int_0^{\tilde{C}(c_1,t_2)} v^{k-1} \frac{\varphi'\left(\tilde{C}\left(c_1, t_2\right)\right)}{\varphi'\left(v\right)} dv \right),$$

it follows by changing variables:

$$I = (\tilde{C}\left(c_1, t_2\right))^k - k(\tilde{C}\left(c_1, t_2\right))^k \varphi'\left(\tilde{C}\left(c_1, t_2\right)\right) \int_0^1 \frac{v^{k-1}}{\varphi'(v\tilde{C}\left(c_1, t_2\right))} dv,$$

which is the k^{th} moments of the variable V, where only T_1 is censored. Because of the symmetry of the copula, the same proof used previously can apply for equation 2 of Corollary1 to get the k^{th} moments of the variable V, where only T_2 is censored. □

3. Parameters estimation

As an application of our results proved in the previous section, we propose a simple way of estimating the unknown parameters in Archimedean copula models. We set up the procedure based on the classical moments method. Assume that $Z_{1:n} < ... < Z_{n:n}$ the order statistics, pertaining to the sample $\{Z_i, \delta_i; 1 \leq i \leq n\}$ with their associated concomitants $\delta_{[i:n]}, ..., \delta_{[n:n]}$. Then, $\delta_{[j:n]} = \delta_i$ if $Z_{j:n} = Z_i$ for $1 \leq j \leq n$. Since we are focusing on the datasets that contain extreme values that include distributions such as Burr, Fréchet, generalized Pareto...etc. However, the selected Pareto model is well known as a heavy-tailed censored data model and it is obvious that the heavy-tailed distribution class plays a significant role in the theory of extreme value. Then it would be natural to assume that both survival functions $S_1 = 1 - F_1$ and $S_2 = 1 - F_2$ are regularly varying at infinity with tail indices $\gamma_1 > 0$ and $\gamma_2 > 0$ respectively. In other word, if we assume that both F_1 and F_2 are heavy-tailed (mentioned that F_1 and F_2 are completely known), so there exist two constants $\gamma_1 > 0$ and $\gamma_2 > 0$ such that: $\lim_{t\to\infty} \frac{S_1(tx)}{S_1(t)} = x^{\frac{-1}{\gamma_1}}$ and $\lim_{t\to\infty} \frac{S_2(tx)}{S_2(t)} = x^{\frac{-1}{\gamma_2}}$, $x > 0$. By a logical sequence and since F_1 and F_2 are heavy-tailed, then the censoring distribution is assumed to be heavy tailed too (ie: the CDF of the observed Z's noted by H and given by $\bar{H} = S_1 S_2$ is heavy-tailed too), hence: $\lim_{t\to\infty} \frac{\bar{H}(tx)}{\bar{H}(t)} = x^{\frac{-1}{\gamma}}$, $x > 0$. Therefore, the extreme value index of the distribution function (d.f.) of (Z, δ) denoted by γ and given by $\gamma = \frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}$ (see Resnick, 2006 [17] and Gomes and Neves [8] two examples of censored data with heavy tails). Let (T_1, T_2) random variables whose distribution can be modelled by an Archimedean copula and is subject to dependent or independent right censoring, $V = \tilde{C}\left(S_1(t_1), S_2(t_2)\right)$ is a conditionally distributed variable follows a so-called Kendall distribution K_C with the density function: $k_C\left(t\right) = \frac{\varphi(t)\varphi''(t)}{(\varphi'(t))^2}$, defined on $(0, 1]$. We noted $M_k(V|c_1, t_2)$ the k^{th}-moments of V, when only T_1 censored and is given by:

$$M_k(V|c_1, t_2) = E(V^k|T_1 > c_1, T_2 = t_2), \quad \text{for } k \geq 1.$$

Relying on the results in Corollary2.1 we have:

$$\begin{aligned} M_k(V|c_1, t_2) &= E(V^k|T_1 > c_1, T_2 = t_2) = \left(\tilde{C}\left(c_1, t_2\right)\right)^k \\ &-k\left(\tilde{C}\left(c_1, t_2\right)\right)^k \varphi'\left(\tilde{C}\left(c_1, t_2\right)\right)\int_0^1 \frac{v^{k-1}}{\varphi'\left(v\tilde{C}\left(c_1, t_2\right)\right)}dv. \end{aligned} \tag{3.1}$$

Assuming that V belongs to a parametric family $V_\theta = \tilde{C}_\theta(u, v)_{\theta\in\mathbb{R}^d}$, then it follows that $\varphi = \varphi_\theta$ and $K_C = K_\theta$, for the unknown parameter $\theta \in \mathbb{R}^d$.

If we suppose that $M_k(V|c_1, t_2) = M_k(\theta|c_1, t_2)$, the equation (3.1) can be written as:

$$\begin{aligned} M_k(\theta|c_1, t_2) = \left(\tilde{C}_\theta\left(c_1, t_2\right)\right)^k - k\left(\tilde{C}_\theta\left(c_1, t_2\right)\right)^k \\ \cdot\varphi'_\theta\left(\tilde{C}_\theta\left(c_1, t_2\right)\right)\int_0^1 \frac{v_\theta^{k-1}}{\varphi'_\theta\left(v_\theta\tilde{C}_\theta\left(c_1, t_2\right)\right)}dv_\theta. \end{aligned}$$

As a result and because of the copula symmetry, the equation (2) in Corollary1 via :

$$M_k(\theta|t_1, c_2) = \left(\tilde{C}_\theta\left(t_1, c_2\right)\right)^k - k\left(\tilde{C}_\theta\left(t_1, c_2\right)\right)^k \cdot \varphi'\left(\tilde{C}_\theta\left(t_1, c_2\right)\right)\int_0^1 \frac{v_\theta^{k-1}}{\varphi'\left(v_\theta\tilde{C}_\theta\left(t_1, c_2\right)\right)}dv_\theta,$$

which shows the k^{th}moments of V, when only T_2 is censored. Given, the empirical version of the moment estimator presented by $\hat{M}_k(\hat{V}|H_j)$:

$$\hat{M}_k(\hat{V}|H_j) = \frac{1}{N}\sum_{i=1}^{n}(\tilde{C}_n(\hat{S}_i\left(t_i\right))|H_j)^k, \text{ for } k \geq 1, \ j = 1; 2,$$

where $\hat{V}$ is the survival empirical copula $\tilde{C}_n$ and H_j represent each case of censoring (H_1 means only T_1 is censored). Then, as the natural estimators of moments copula, it is necessary to solve the equation system given by

$$M_k(\theta|H_j) = \hat{M}_k(\hat{V}|H_j), \text{ for } \theta = (\theta_1, ..., \theta_d) \text{ and } j = 1; 2.$$

To obtain the unique solution $\hat{\theta}^{CCM} = (\hat{\theta}_1, ..., \hat{\theta}_d)$ called the censored copula moment (CCM) estimator of θ.

3.1. **APPLICATION: ILLUSTRATIVE EXAMPLES.** From now, only T_1 considered as a censored variable. Therefore, two models evaluated, the first is for the Clayton model of one-parameter and the second is for the Gumbel model of two parameters.

- **Clayton model**

For the Clayton model of one-parameter the survival copula is known by

$$\tilde{C}_\alpha(u, v) = u + v - 1 + ((1-u)^{-\alpha} + (1-v)^{-\alpha} - 1)^{\frac{-1}{\alpha}},$$

with generator $\varphi_\alpha(t) = t^{-\alpha} - 1$, $\alpha > 0$. Applying Corollary1, we can simplify the estimating equations as follow:

$$\mathbb{E}(V^k\big| T_1 > c_1, T_2 = t_2) = (m)^k - km^k\varphi'\left(m\right)\int_0^1 \frac{v^{k-1}}{\varphi'\left(vm\right)}dv, \tag{3.2}$$

for $k > 0$ and when $m = \tilde{C}\left(c_1, t_2\right)$, represent the ordinary copula. If we simplify more the formula (6) we can obtain:

$$\mathbb{E}(V^k\big| T_1 > c_1, T_2 = t_2) = (m)^k - km^{k-\alpha-1}\int_0^1 \frac{v^{k-1}}{(vm)^{-\alpha-1}}dv.$$

By an elementary calculation we get the kth moments $M_k(\alpha) = m^k - \frac{km^{k-1}}{k+\alpha+1}$, where $m = \tilde{C}\left(c_1, t_2\right)$. Hence, for $k = 1$ the first moments is normally given by $M_1(\alpha) = m - \frac{1}{\alpha+2}$, which allows us easily find the unique estimator of α given by $\hat{\alpha} = 2 - \frac{1}{m-\hat{M}_1}$.

- **Gumbel model**

The second model is about the Gumbel copula family in the bivariates case and for two parameters, which known by:

$$C_{\alpha,\beta}\left(u,v\right) = \left(\left(\left(u^{-\alpha}-1\right)^{\beta} + \left(v^{-\alpha}-1\right)^{\beta}\right)^{\frac{1}{\beta}} + 1\right)^{-\frac{1}{\alpha}},$$

with the generator: $\varphi_{\alpha,\beta}(t) = (t^{-\alpha}-1)^{\beta}$, where $\alpha > 0$ and $\beta \geq 1$ (see [6]). Obviously, the survival copula of the Gumbel family given by:

$$\tilde{C}_{\alpha,\beta}\left(u,v\right) = u + v - 1 + \left(\left(\left((1-u)^{-\alpha}-1\right)^{\beta} + \left((1-v)^{-\alpha}-1\right)^{\beta}\right)^{1/\beta} + 1\right)^{-1/\alpha}.$$

By analogy, the k^{th} moments, is given by:

$$M_k(\theta|\, c_1, t_2) = m^k - \frac{k(\beta-1)\left(m^{-\alpha}-1\right)^{\beta}}{\alpha m^{2\alpha+1}} \frac{\Gamma\left(1-\beta\right)\Gamma\left(\frac{1}{\alpha}\left(k+\alpha\beta+1\right)\right)}{\Gamma\left(\frac{1}{\alpha}\left(k+2\alpha+1\right)\right)},$$

where $\theta = (\alpha, \beta)$. As a result, the two first moments M_1 , M_2 are given by:

(3.3)

$$\begin{cases} M_1(\theta|\, c_1, t_2) = M_1(\alpha,\beta) = m - \frac{(\beta-1)\left(m^{-\alpha}-1\right)^{\beta}}{\alpha m^{2\alpha+1}} \frac{\Gamma(1-\beta)\Gamma\left(\frac{1}{\alpha}(\alpha\beta+2)\right)}{\Gamma\left(\frac{2}{\alpha}(\alpha+1)\right)} = M_1(\alpha,\beta) \\ M_2(\theta|\, c_1, t_2) = M_2(\alpha,\beta) = m^2 - \frac{2(\beta-1)\left(m^{-\alpha}-1\right)^{\beta}}{\alpha m^{2\alpha+1}} \frac{\Gamma(1-\beta)\Gamma\left(\frac{1}{\alpha}(\alpha\beta+3)\right)}{\Gamma\left(\frac{1}{\alpha}(2\alpha+3)\right)} = M_2(\alpha,\beta) \end{cases}.$$

Consequently, the estimator $\hat{\theta}$ of θ is the unique solution of the system:

$$\begin{cases} M_1(\theta) = \hat{M}_1 \\ M_2(\theta) = \hat{M}_2 \end{cases}.$$

4. Simulation studies

Taking into consideration that only T_1 to be right-censored. A simulation study was carried out to evaluate the performance of the proposed estimators, based on the Monte Carlo procedure under the Gumbel Archimedean dependence assumption. The results are shown in Tables 1–4 accordingly. We first generate bivariate data from the Gumbel copula model of T_1 and T_2 with Pareto

TABLE 1. Moments estimator performance based on Clayton survival copula of one parameter under singly right censored variable.

N	$n=30$		$n=50$		$n=100$		$n=500$		$n=1000$	
% of censures	R.Bias	RMSE	R.Bias	RMSE	R.Bias	RMSE	R.Bias	RMSE	R.Bias	RMSE
$\tau = 0.05$, $\alpha = 0.1$										
5%	-0.0547	0.0640	-0.0513	0.0615	-0.0549	0.0631	-0.0535	0.0626	-0.0542	0.0625
10%	-0.0532	0.0621	-0.0520	0.0609	-0.055	0.0632	-0.0529	0.0620	-0.0548	0.0642
15%	-0.0554	0.0642	-0.0538	0.0627	-0.0539	0.0632	-0.0532	0.0623	-0.0541	0.0630
20%	-0.0536	0.0624	-0.0543	0.0632	-0.0525	0.0614	-0.0530	0.0619	-0.0553	0.0638
$\tau = 0.5$, $\alpha = 0.2$										
5%	-0.0257	0.0298	-0.0260	0.0301	-0.0253	0.0291	-0.0255	0.0299	-0.027	0.0311
10%	-0.0262	0.0304	-0.0257	0.0299	-0.0260	0.0299	-0.0264	0.0305	-0.0267	0.0308
15%	-0.0262	0.0302	-0.0262	0.0301	-0.0254	0.0297	-0.0258	0.0299	-0.0254	0.0295
20%	-0.0257	0.0299	-0.0257	0.0299	-0.0257	0.0300	-0.0257	0.0299	-0.0261	0.0302
$\tau = 0.7$, $\alpha = 0.4$										
5%	-0.0130	0.0149	-0.0126	0.0146	-0.0125	0.0145	-0.0131	0.0150	-0.0127	0.0148
10%	-0.0126	0.0145	-0.0129	0.0149	-0.0126	0.0146	-0.0123	0.0144	-0.0129	0.0149
15%	-0.0132	0.0151	-0.0126	0.0146	-0.0124	0.0144	-0.0125	0.0145	-0.0126	0.0146
20%	-0.0127	0.0147	-0.0126	0.0147	-0.0130	0.0149	-0.0127	0.0148	-0.0126	0.0146

margins of parameters γ_1 and γ_2 respectively. We also generate the censoring variable C_1 whose marginal distribution is a Pareto with γ_c parameter. We suppose that $\gamma_1 = \gamma_2 = 0.3$ and that the corresponding percentage of observed data is given by $p_1 = \frac{\gamma_c}{\gamma_1+\gamma_c}$, we choose parameter values corresponding to p_1 values $0.95, 0.90, 0.85, 0.80$, and we solve the equation $p_1 = \frac{\gamma_c}{\gamma_1+\gamma_c}$ to get the pertaining γ_c-values. Based on the parameters estimate procedure in Section 3, 1000 replicas to be generated for each common size n varied for $n = 30, 50, 100, 500, 1000$, to pick our final performance as empirical evidence of the results gained across all replicates. Table 1 describes the results obtained for the Clayton model of one-parameter (3.2) with unit Pareto margins of shape parameter 0.3, whose estimator looked with $\hat{\alpha} = 2 - \frac{1}{m-\hat{M}_1}$. Where we can see the R.Bias and the RMSE are very close to zero. Once the rate of dependence τ is increased, we see an improvement in the results of the estimated parameters $\hat{\alpha}$ due to a large decrease in R.Bias and RMSE, which are inversely proportional readings.

Now, by considering the second model of the Gumbel survival copula of two parameters, where the two first moments are formulated as (3.3). Given Kendall's tau $\tau_{\alpha,\beta} = 4E(V_{\alpha,\beta}) - 1$ as an association index (a function of the

TABLE 2. The true parameters of the survival Gumbel copula.

τ	α	β
0.05	0.1	1.00
0.5	0.2	1.82
0.7	0.4	2.78

dependency parameter in Archimedean copula models), we select the survival copula parameter values (α, β) that correspond to specified values of τ by using the select values $0.05, 0.5$ and 0.7 of Kendall's tau dependence assumption values and the transformed of the underlying survival Gumbel copula

$$V_{\alpha,\beta} = u + v - 1 + \left(\left(\left((1-u)^{-\alpha} - 1 \right)^{\beta} + \left((1-v)^{-\alpha} - 1 \right)^{\beta} \right)^{1/\beta} + 1 \right)^{-1/\alpha},$$

as shown in Table 2.

Table 3: Moments estimator performance based on Gumbel survival copula of two parameters under singly right censored variable and for weak dependence.

$\tau = 0.05,\ \alpha = 0.1 \Rightarrow \beta = 1.00$

Sample Size	c_1	$\hat{\alpha}$ R.Bias	$\hat{\alpha}$ RMSE	$\hat{\beta}$ R.Bias	$\hat{\beta}$ RMSE
% of censoring 20					
30	0.11681	-0.05381	0.06285	0.13100	0.01145
50	0.07223	-0.05428	0.06293	0.09702	0.01185
100	0.03696	-0.05455	0.06302	0.12730	0.01179
500	0.00796	-0.05241	0.06136	0.11346	0.01171
1000	0.00411	-0.05310	0.06181	0.11126	0.01178
% of censoring 15					
30	0.15716	-0.05364	0.06258	0.14563	0.01153
50	0.09968	-0.05506	0.06414	0.17947	0.01192
100	0.05382	-0.05405	0.06289	0.14270	0.01125
500	0.01168	-0.05338	0.06219	0.12725	0.01156
1000	0.00591	-0.05223	0.06111	0.13337	0.01216

% of censoring 10					
30	0.22207	-0.05170	0.06076	0.16483	0.01172
50	0.14458	-0.05233	0.06110	0.16214	0.01164
100	0.08509	-0.05438	0.06300	0.14554	0.01151
500	0.01801	-0.05562	0.06412	0.16037	0.01148
1000	0.00892	-0.05353	0.06223	0.14950	0.0117
% of censoring 5					
30	0.31617	-0.05347	0.06244	0.17334	0.01160
50	0.25135	-0.05312	0.06190	0.17205	0.01181
100	0.15138	-0.05363	0.06259	0.17993	0.01146
500	0.03666	-0.05295	0.06245	0.1544	0.01162
1000	0.01841	-0.05247	0.06124	0.15832	0.01170

Table 4: Moments estimator performance based on Gumbel survival copula of two parameters under singly right censored variable and for moderate dependence.

$\tau = 0.5,\ \ \alpha = 0.2 \Rightarrow \beta = 1.82$

% of censoring 20					
Sample Size	c_1	$\hat{\alpha}$ R.Bias	$\hat{\alpha}$ RMSE	$\hat{\beta}$ R.Bias	$\hat{\beta}$ RMSE
30	0.12019	-0.02632	0.03015	0.06926	0.00642
50	0.07406	-0.02598	0.03007	0.08484	0.00645
100	0.03829	-0.02558	0.02983	0.08158	0.00641
500	0.00804	-0.02591	0.03011	0.10276	0.00637
1000	0.00394	-0.02579	0.02979	0.08549	0.00643
% of censoring 15					
30	0.15611	-0.02605	0.03027	0.0989	0.00657
50	0.10008	-0.02498	0.02906	0.09554	0.00633
100	0.05478	-0.02615	0.03037	0.12623	0.00641
500	0.01088	-0.02536	0.02957	0.09632	0.00631
1000	0.00546	-0.02605	0.03012	0.08785	0.00632

% of censoring 10					
30	0.21223	-0.2556	0.02972	0.17075	0.00637
50	0.14818	-0.02544	0.02948	0.09582	0.00644
100	0.08047	-0.02526	0.02945	0.13741	0.00644
500	0.01811	-0.02714	0.03121	0.09872	0.00636
1000	0.00915	-0.02675	0.03072	0.09565	0.00637
% of censoring 5					
30	0.30813	-0.02537	0.02939	0.16263	0.00644
50	0.23892	-0.02488	0.02905	0.19983	0.00646
100	0.15392	-0.02639	0.03044	0.09317	0.00640
500	0.03543	-0.02568	0.02995	0.10464	0.00630
1000	0.01844	-0.02517	0.02931	0.10486	0.00637

Table 5: Moments estimator performance based on Gumbel survival copula of two parameters under singly right censored variable and for strong dependence.

$\tau = 0.7,\ \alpha = 0.4 \Rightarrow \beta = 2.78$

% of censoring 20					
Sample Size	c_1	$\hat{\alpha}$ R.Bias	$\hat{\alpha}$ RMSE	$\hat{\beta}$ R.Bias	$\hat{\beta}$ RMSE
30	0.11466	-0.01275	0.01476	0.14167	0.00377
50	0.07528	-0.0126	0.01465	0.10573	0.00393
100	0.03747	-0.01287	0.01483	0.08515	0.00389
500	0.00787	-0.01251	0.01446	0.11032	0.00391
1000	0.00421	-0.01285	0.01478	0.08713	0.00385
% of censoring 15					
30	0.15831	-0.01260	0.01459	0.14577	0.00394
50	0.10332	-0.01282	0.01485	0.10507	0.00387
100	0.05540	-0.01303	0.01502	0.13646	0.00385
500	0.01108	-0.01221	0.01427	0.10486	0.00391
1000	0.00590	-0.01276	0.01473	0.10388	0.00389

% of censoring 10					
30	0.21203	-0.01244	0.01443	0.12436	0.00384
50	0.14910	-0.01261	0.01460	0.16765	0.00384
100	0.08393	-0.01285	0.01475	0.15731	0.00389
500	0.01785	-0.01252	0.01455	0.1095	0.00395
1000	0.00872	-0.01297	0.01509	0.11251	0.00392
% of censoring 5					
30	0.31116	-0.0126	0.01457	0.1568	0.00391
50	0.24405	-0.01236	0.01448	0.17432	0.00389
100	0.15254	-0.01259	0.01466	0.10309	0.00383
500	0.03668	-0.01306	0.01504	0.11537	0.00387
1000	0.01838	-0.01271	0.01469	0.10606	0.00386

Tables 3-5 shows the results obtained of CCM estimator $\left(\hat{\alpha}, \hat{\beta}\right)$ of (α, β) based on survival copula under the censored variable T_1, generated from the Gumbel copula model of two parameters given by 3.3 with unit Pareto margins of shape parameter 0.3. By looking at three different values of dependency weak (0.05) moderate (0.5) and strong (0.7), the R.Bias and the RMSE of the two parameters estimate $\hat{\alpha}$ and $\hat{\beta}$ were calculated and are usually given lower values especially when the dependency increases.

5. Application to a real data set

In this part of the paper, we examine the performance of our estimation procedure for the real data set of diabetic retinopathy, which is available in the R software via the survival package [22, 23]. Diabetic retinopathy is a disease that affects people with diabetes and can outcome in vision loss and blindness. In the study a significant number of diabetic patients (times of follow-up for 197 diabetic patients under 60 years old , who represent a 50% sample of high risk patients for loss of vision) was followed for an extended period.

The primary aim of the study was to assess the efficacy of photocoagulation as a treatment for proliferative retinopathy. For each patient, one eye was treated with laser photocoagulation, and the other eye was taken as a control. To model this data, the evaluation that piques our interest is concerned by (2.3), where the two variables are both censored.

To fit the failure times (T_1, T_2), we use a bivariate Gumbel family of two parameters with extreme value margins (Pareto $(\gamma = 0.3)$) for both T_1 and T_2. Taking T_1 as the time to visual loss for the treatment eye and T_2 the time to visual loss for the control eye. The percentage of censure times for T_1 is 73% (143 observations) and 49% (96 observations) for T_2.

a. Scatter plots b. Cens and obs data for T_1 c. Cens and obs data for T_2

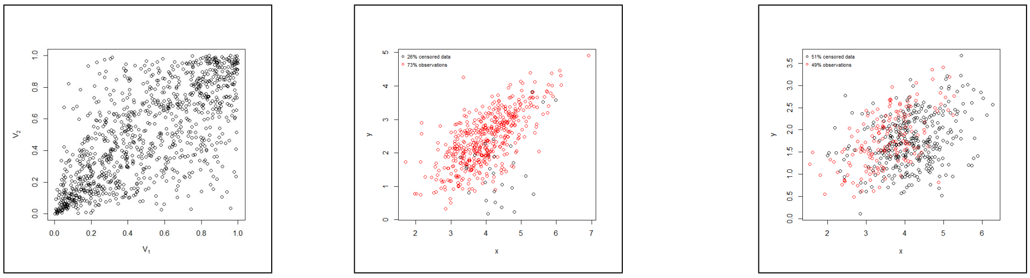

FIGURE 1. Censored and observed points for each T_1 and T_2 separately of bivariate survival Gumbel copula.

To model this data, we ran the algorithm presented in section3, by considering Kendall's tau as the association index (a function of the dependency parameter).

Table 6: Relative bias and RMSE of Moments estimator based on a Gumbel survival copula model from the Diabetic Retinopathy study data.

$\tau = 0.7$, $\alpha = 0.1 \Rightarrow \beta = 2.78$

sample Size	% of cens	$\hat{\alpha}$ R.Bias	$\hat{\alpha}$ RMSE	$\hat{\beta}$ R.Bias	$\hat{\beta}$ RMSE	Association τ before cens	Association τ after cens
$n = 50$	5%	-0.0126	0.0146	0.3941	0.0042	0.6967	0.6412
	10%	-0.0126	0.0146	0.4147	0.0041	0.7056	0.5925
	15%	-0.0128	0.0148	0.3547	0.0041	0.7002	0.5373
	20%	-0.0125	0.0145	0.3440	0.0042	0.7001	0.4930
$n = 100$	5%	-0.0125	0.0146	0.4243	0.0043	0.7004	0.6415
	10%	-0.0122	0.0143	0.3818	0.0041	0.6994	0.5849
	15%	-0.0133	0.0152	0.3613	0.0042	0.7016	0.5376
	20%	-0.0130	0.0149	0.3585	0.0042	0.6982	0.4859
$n = 500$	5%	-0.0125	0.0146	0.4102	0.0042	0.7007	0.6426
	10%	-0.0131	0.0150	0.3784	0.0043	0.7000	0.5872
	15%	-0.0126	0.0146	0.3648	0.0041	0.7000	0.5369
	20%	-0.0130	0.0150	0.3415	0.0042	0.6998	0.4885

	5%	-0.0127	0.0147	0.4077	0.0042	0.7003	0.6409
$n = 1000$	10%	-0.0128	0.0148	0.3841	0.0042	0.7001	0.5881
	15%	-0.0125	0.0145	0.3573	0.0042	0.6999	0.5361
	20%	-0.0125	0.0145	0.3388	0.0042	0.7003	0.4891

To assess the performance of the considered estimator, we have used the RMSE and the relative bias (R.Bais) define by:

$$\text{R.Bais} = \frac{1}{N}\frac{\left[\sum_{i=1}^{N}\hat{\theta}_i - \theta\right]}{\theta}, \quad \text{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta\right)^2},$$

where $\hat{\theta}_i$ is the CCM estimator (from the considered model) of θ.

Figure1 (a) shows the scatter plots of the survival Gumbel copula with two parameters. Figure 1 (b) and Figure 1 (c) shows the censored and observed data of the bivariate survival Gumbel copula for each variable T_1 and T_2 respectively. Table 6 shows the relative bias (R.Bias) and the RMSE of the parameters estimates under different doubly right-censoring values, the dependency value before and after censoring (Association τ before cens, after cens). For this data set, the estimator gave the smaller relative bais and RMSE values, which proves its effectiveness and robustness.

6. Conclusion and perspective

In this paper, we have presented a semi-parametric estimation method of a survival copula $\tilde{C}$ based on the classical method of moments under individually censored of (T_1, T_2). As a logical continuation of results established by Idiou et al 2020 [16], general formulas are given for marginal survival copula $\tilde{C}$ of such data by the assumption that their underlying copula is Archimedean. Two models are proposed for this study, the Clayton model of one-parameter and the Gumbel model of two-parameters proved our theoretical results obtained. Under the Archimedean dependence structure assumption for censored data, a simulation study evaluates the performance of our estimator, relative bias, and RMSE formulas for estimator are evaluated. This study shows that the new estimator works well, where the values obtained are tending towards zero for each case of small and even large samples. The methodology presented in section 3, was applied to real data from the Diabetic Retinopathy Study, which is available in the "survival" package [22], [23] of the R software. For this data set, the estimator gave the smaller relative bias and RMSE values, which proves its effectiveness and robustness.

Consequently, this method is preferable if we compare it with the maximum likelihood method and other methods ([7], [4]), because of its easy analytical mathematical form.

Our main result for this study is based on the copula approaches and the survival analysis, under the Archimedean dependence structure assumption for censored data. Based on these results, we can establish a new methods checking process of Archimedean copula models for censored variables. This is one of our recent research areas and the idea was already established in another paper that is under preparation.

REFERENCES

[1] A. WANG: *On the nonidentifiability property of Archimedean copula models under dependent censoring*, Statistics and Probability Letters, (2012) 621-625.

[2] K.J. BANDEEN-ROCHE, K.Y. LIANG: *Modelling failure-time associations in data with multiple levels of clustering*, Biometrika (83) (1996), 29-39.

[3] D.G. CLAYTON: *A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence*, Biometrika, **65**(1) (1978), 141-151.

[4] D. OAKES: *Bivariate survival models induced by frailties*, Statist. Assoc. **84** (1989), 487–493.

[5] C. GENEST, R.J. MACKAY:*The joy of copulas: Bivariate distributions with uniform marginals*, Amer. Statist. **40** (1986), 280-283.

[6] C. GENEST, L.P. RIVEST: *Statistical inference procedures for bivariate Archimedean copulas*, Statist. Assoc. **88** (1993), 1034-1043.

[7] G.K. BHATTACHARYYA, R.A. JOHNSON: *Maximum Likelihood Estimation and Hypothesis Testing in the Bivariate Exponential Model of Marshall and Olki*, Journal of the American Statistical Association, **68**(343) (1973), 704-706.

[8] M.I. GOMES, M.M. NEVES: *Estimation of the extreme value index for randomly censored data*, Biometrical Lett., **48**(1) (2011), 1-22.

[9] S. GRIBKOVA, O. LOPEZ: *Non-parametric Copula estimation under bivariate censoring*, Scandinavian Journal of Statistics, **42**(4) (2015), 925-946.

[10] H. JOE: *Asymptotic efficiency of th two-stage estimation method for copula based models*, Journal of Multivariate Analysis, **94** (2005), 401-419.

[11] P. HOUGAARD: *Survival models for heterogeneous populations derived from stable distributions*, Biometrika, **73**(2) (1986), 387-396.

[12] P. HOUGAARD: *A class of multivariate failure time distributions*, Biometrika, **73**(3) (1986), 671-678.

[13] A.W. MARSHALL, I. OLKIN: *Families of multivariate distributions*, Journal of the American Statistical Association, **83** (1988), 834-840.

[14] M. BOUKELOUA: *Study of semiparametric copula models via divergences with bivariate censored data*, Communications in Statistics - Theory and Methods, 2020.

[15] R.B. NELSEN: *An Introduction to Copulas*. second ed. Springer, New York, 2006.

[16] N.IDIOU ET AL: *A semi-parametric estimation of copula models based on moments method under right censoring*. Journal of TWMS J. App and Eng. Math. (Accepted and to be published), 2021.

[17] S. RESNICK: *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer, 2006.

[18] W. STUTE: *Consistent estimation under random censorship when covariables are present*, Journal of Multivariate Analysis **45**(1) (1993), 89-103.

[19] W. WANG, M.T. WELLS: *Model selection and semiparametric inference for bivariate failure-time data*, J. Amer. Statist. Assoc. **449** (2000), 62–72.

[20] W. WANG, M.T. WELLS: *Estimation of Kendall's tau under censoring*, Statist. Sinica, **10** (2000), 1199–1218.

[21] A. WANG, O. DAVID: *Some properties of the Kendall distribution in bivariate Archimedean copula models under censoring*, Statistics and Probability Letters, **78**(16) (2008), 2578-2583.

[22] AM. J. OPHTHALMOL : *The Diabetic Retinopathy Study Research Group*, Preliminary report on the effect of photocoagulation therapy. **81** (1976), 383-396.

[23] T.M. THERNEAU: *A Package for Survival Analysis*. Version 2.38, 2015.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MOHAMED KHIDER
SIDI OKBA, BP 145 RP 07000, BISKRA,
ALGERIA.
Email address: nesrine.idiou@univ-biskra.dz

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MOHAMED KHIDER
SIDI OKBA, BP 145 RP 07000, BISKRA,
ALGERIA.
Email address: f.benatia@univ-biskra.dz