Association Parameters Regression for Bivariate Failure-Time Data

A. Adam Ding
Northeastern University
Department of Mathematics
Boston, MA 02115
U.S.A.
a.ding@math.neu.edu

Jin-Jian Hsieh
National Chung Cheng University
Department of Mathematics
Chia-Yi
Taiwan, R.O.C.
jjhsieh@math.ccu.edu.tw

Weijing Wang
National Chiao-Tung University
Institute of Statistics
Hsin-Chu
Taiwan, R.O.C.
wjwang@stat.nctu.edu.tw

Abstract: Copula models are often used to model the dependence structure in bivariate failure-time data. We consider a covariate effect regression method on the copula parameter for Archimedean copulas. The proposed method can handle three different data structures, namely typical bivariate data, semi-competing risks data and dependent truncation data. We derive large-sample properties of the proposed estimators, and study their finite-sample performances via simulations and application to a well-known data set.

Key–Words: Archimedean copulas, Censoring, Clayton model, Truncation, Local odds ratio, Semi-competing risks data.

1 Introduction

Let T_1 and T_2 be a pair of continuous lifetime random variables. The dependence structure between them can be described the copula according to Sklar's theorem: $F(t_1, t_2) = C\{F_1(t_1), F_2(t_2)\}$, where F is the joint cumulative distribution (or survival) function, and F_1 and F_2 are corresponding marginal functions. The copula function $C : [0, 1]^2 \to [0, 1]$ summarized the dependence information between T_1 and T_2.

Parametric copula models have been used for statistical inference in analyzing bivariate survival data subject to independent censoring [1, 2], semi-competing risks data [3, 4, 5, 6, 7, 8, 9] and dependent truncation data [10, 11].

When covariates are present, most analysis focused on the covariate effects on the marginal distributions, and assume that the copula is constant across different covariate values. Recently, a few work allows the copula to change with covariates. The semi-competing risks data marginal regression using estimating equations [9] and nonparametric maximum likelihood estimation [12] allowed covariate-varying copula. [13] considered the nonparametric estimation of the covariate-varying Clayton copula parameter using local kernel methods for three types of bivariate failure-time data in the following. Here we consider parametric regression of the covariate-varying Archimedean copulas parameter. To simplify the illustration, external censoring is temporarily ignored but will taken into account in Section 2.2.

Data Structure 1 - Typical failure-time data. There is no specific relationship between T_1 and T_2 so that possible observations fall in the region of $\{(s, t) : 0 < s < \infty, 0 < t < \infty\}$. Such data are suitable for measuring the failure times occurred to different biological units such as twins, family members or paired organs on the same person.

Data Structure 2 - Semi-competing risks data. In analysis of multiple events data, let T_1 be the time to a non-terminal event such as onset of a disease and T_2 the time to a terminal event such as death. Notice that T_2 is a competing risk for T_1 but not vise versa. One can only observe $(T_1 \wedge T_2, T_2, I(T_1 \leq T_2))$. Complete observations of (T_1, T_2) can only fall in $\{(s, t) : 0 < s \leq t < \infty\}$. Incomplete observations are those with $T_1 > T_2$ in which the value of T_2 is still observable.

Data Structure 3 - Dependent truncation data. Assume (T_1, T_2) are observable only if $T_1 < T_2$. We say T_1 is subject to right truncation by T_2 or T_2 subject to left truncation by T_1. In an example of transfusion-related AIDS, (T_1, T_2) may refer to the incubation time from infection to AIDS and the lapse time measured from infection to the end of study, respectively. Only those who developed AIDS before the end of study can be included in the sample.

In Section 2, we introduce the proposed methodology. Section 3 describes the simplified version under Clayton copula. The resulting estimator is shown to be asymptotically normally distributed. In addition a method to check the Clayton association Assumption is proposed. Section 4 studies the proposed estimator through synthetic data and a real data set.

ISBN: 978-960-474-380-3

2 The Proposed Approach

We propose to conduct the regression analysis through time-varying local dependence measures. We first introduce these measures. Then we derive an estimating equation using such measures.

2.1 Time-varying Association Measures

For two correlated uniform variables U_1 and U_2 on the unit interval, their joint distribution can be defined as the copula function

$$C_\alpha(u_1, u_2) = \Pr(U_1 \le u_1, U_2 \le u_2),$$

where α measures the degree of association and is related to Kendall's tau as $\tau = 4\int_0^1\int_0^1 C_\alpha(u_1,u_2)C_\alpha(du_1,du_2) - 1$. Local dependence structure between U_1 and U_2 can be described by the cross ratio function [14]:

$$\theta(u_1,u_2) = \frac{\{C_\alpha(u_1,u_2)\}\{D_1D_2C_\alpha(u_1,u_2)\}}{\{D_1C_\alpha(u_1,u_2)\}\{D_2C_\alpha(u_1,u_2)\}}, \quad (1)$$

where D_j means taking derivative respect to u_j for $j = 1, 2$. Equivalently the above function can be written as the local odds ratio function:

$$\theta(u_1,u_2) = \frac{\Pr(\Delta^o_{ij} = 1|\tilde{\boldsymbol{U}}_{ij} = (u_1,u_2))}{\Pr(\Delta^o_{ij} = 0|\tilde{\boldsymbol{U}}_{ij} = (u_1,u_2))}, \quad (2)$$

where $\Delta^o_{ij} = I\{(U_{1i} - U_{1j})(U_{2i} - U_{2j}) > 0\}$ is the concordance indicator for (U_{1i}, U_{2i}) and (U_{1j}, U_{2j}), which are independent replications of (U_1, U_2), and $\tilde{\boldsymbol{U}}_{ij} = (U_{1i} \vee U_{1j}, U_{2i} \vee U_{2j})$.

Conditional on the covariate value $Z = z$, the copula structure on lifetime variables (T_1, T_2) is

$$F_z(t_1,t_2) = C_{\alpha(z)}\{F_{1,z}(t_1), F_{2,z}(t_2)\}, \quad (3)$$

where F_z is a joint function of (T_1, T_2) and $F_{k,z}$ is a marginal function for T_k for $k = 1, 2$. We follow the conventions for the three data structures discussed in Section 1. For data structures 1 and 2, $F_{k,z}(t_k) = \Pr_z(T_k > t_k)$ $(k = 1, 2)$ and $F_z(t_1, t_2) = \Pr_z(T_1 > t_1, T_2 > t_2)$; but for the third type, $F_{1,z}(t_1) = \Pr_z(T_1 \le t_1)$, $F_{2,z}(t) = \Pr_z(T_2 > t_2)$ and $F_z(t_1, t_2) = \Pr_z(T_1 \le t_1, T_2 > t_2)$ for $0 \le t_1 < t_2 < \infty$. The covariate effects on the marginal status is reflected in $F_{1,z}$ and $F_{1,z}$. In this paper, we focus on the covariate effect on association which is characterized by $\alpha(z)$. We propose to study $\alpha(z)$ through θ_z. The θ_z is defined as θ in (2) conditional on $Z = z$, and is related to $\alpha(z)$ through equation (1).

Since we can not directly observe $\tilde{\boldsymbol{U}}_{ij}$, we first find the local odds ratio θ_z on the original time scale. Let (T_{1i}, T_{2i}, Z_i) $(i = 1, \ldots, n)$ be a random sample of (T_1, T_2, Z). Assume that Z takes discrete values. Conditional on $Z_i = Z_i = z$ we can define $\Delta_{ij} = I\{(T_{1i} - T_{1j})(T_{2i} - T_{2j}) > 0\}$ which only involves the original time variables. Then the value of Δ^o_{ij}, which is defined under equation (2), can be known from Δ_{ij}. Specifically, $\Delta_{ij} = \Delta^o_{ij}$ for the first two data structures, and $\Delta_{ij} = 1 - \Delta^o_{ij}$ for the third data structure. Based on Δ_{ij}, we define $\tilde{\theta}_z(t_1, t_2)$ as

$$\frac{\Pr(\Delta_{ij} = 1|\tilde{\boldsymbol{T}}_{ij} = (t_1,t_2), Z_i = Z_j = z)}{\Pr(\Delta_{ij} = 0|\tilde{\boldsymbol{T}}_{ij} = (t_1,t_2), Z_i = Z_j = z)} \quad (4)$$

where $\tilde{\boldsymbol{T}}_{ij} = (T_{1i} \wedge T_{1j}, T_{2i} \wedge T_{2j})$ and hence $\tilde{\theta}_z(t_1, t_2) = \theta_z(u_1, u_2)$ for the first two data structures; $\tilde{\boldsymbol{T}}_{ij} = (T_{1i} \vee T_{1j}, T_{2i} \wedge T_{2j})$ and hence $\tilde{\theta}_z(t_1, t_2) = 1/\theta_z(u_1, u_2)$ for the third type. Here, $u_1 = F_{1,z}(t_1)$ and $u_2 = F_{2,z}(t_2)$. Accordingly the time-varying concordance probability

$$\tilde{\pi}_z(t_1,t_2) = \tilde{\theta}_z(t_1,t_2)/[1 + \tilde{\theta}_z(t_1,t_2)] \quad (5)$$

denotes $\Pr(\Delta_{ij} = 1|\tilde{\boldsymbol{T}}_{ij} = (t_1, t_2), Z_i = Z_j = z)$.

When the underlying copula function is specified, we have explicit formula for $\theta_z(u_1, u_2)$, then can derive $\tilde{\pi}_z(t_1, t_2)$ using (4) and (5). We can then estimate the covariate effects on $\alpha(z)$ by estimating $\tilde{\pi}_z(t_1, t_2)$. [15, 16] studied the estimation of time-varying association measures θ for bivariate survival data. Our focus is different in that our objective is not in θ_z but using this to infer the covariate effect on $\alpha(z)$.

2.2 Effect of External Censoring

Now we incorporate external censoring in the three data structures which occurs due to drop-out or the end-of-study effect. For the three data structures, we use $(X_{1i}, X_{2i}, \delta_{1i}, \delta_{2i}, Z_i)$ $(i = 1, \ldots, n)$ to denote observed data but the definitions depend on the data type. For typical failure-time data, assume that (T_{1i}, T_{2i}) is subject to independent censoring by (C_{1i}, C_{2i}) such that $X_{ki} = T_{ki} \wedge C_{ki}$ and $\delta_{ki} = I(T_{ki} \le C_{ki})$ for $k = 1, 2$. For semi-competing risks data, assume that (T_{1i}, T_{2i}) are subject to a common external censoring variable C_i so that $X_{1i} = T_{1i} \wedge T_{2i} \wedge C_i$, $X_{2i} = T_{2i} \wedge C_i$ and $\delta_{ki} = I(X_{ki} = T_{ki})$ for $k = 1, 2$. For the third type, we consider left truncated and right censored data such that T_{2i} may be censored by an external censoring variable C_i. Observed variables are $X_{1i} = T_{1i}$, $X_{2i} = T_{2i} \wedge C_i$, $\delta_{1i} = 1$ and $\delta_{2i} = I(T_{2i} \le C_i)$ subject to $X_{2i} > X_{1i} = T_{1i}$.

It is important to note that a random sample of (T_{1i}, T_{2i}, Z_i) $(i = 1, \ldots, n)$ is not available in the latter two data structures even without external censoring. When censoring is present, the value of Δ_{ij} may

ISBN: 978-960-474-380-3

be unknown. Now we derive the condition that Δ_{ij} can be fully observed and provides unbiased information about $\tilde{\pi}_z(t_1, t_2)$. Let $\tilde{\boldsymbol{T}}_{ij} = (\tilde{T}_{1,ij}, \tilde{T}_{2,ij})$. The for data structure 1, we define $D_{ij}(z) = I(\tilde{T}_{1,ij} < \tilde{C}_{1,ij}, \tilde{T}_{2,ij} < \tilde{C}_{2,ij}, Z_i = Z_j = z)$ where $\tilde{C}_{k,ij} = C_{1i} \wedge C_{kj}$ $(k = 1, 2)$. For data structure 2, $D_{ij}(z) = I(\tilde{T}_{1,ij} < \tilde{T}_{2,ij} < \tilde{C}_{ij}, Z_i = Z_j = z)$, where $\tilde{C}_{ij} = C_i \wedge C_j$. For data structure 3, $D_{ij}(z) = I(\tilde{T}_{1,ij} < \tilde{T}_{2,ij} < \tilde{C}_{ij}, Z_i = Z_j = z)$. For each data structure, it is easy to see that the value of Δ_{ij} is known when $D_{ij}(z) = 1$. In addition, define

$$\tilde{p}_z(t_1, t_2) = \Pr(\Delta_{ij} = 1 | \tilde{\boldsymbol{T}}_{ij} = (t_1, t_2), D_{ij}(z) = 1).$$

Then $\tilde{p}_z(t_1, t_2) = \tilde{\pi}_z(t_1, t_2)$ for all (t_1, t_2) values in the model range of (3). We will estimate the covariate effect through $\tilde{p}_z(t_1, t_2)$.

2.3 Association Parameter Regression.

We consider a parametric model $\alpha(z) = \alpha(z, \beta)$ with parameter β representing the covariate effect on the association parameter. We consider the inference under Archimedean copulas [17]:

$$C_\alpha(u_1, u_2) = \phi_\alpha^{-1}\{\phi_\alpha(u_1) + \phi_\alpha(u_2)\}, \qquad (6)$$

where the generating function $\phi_\alpha(\cdot) : [0, 1] \to [0, \infty]$ satisfies that $\phi_\alpha(1) = 0$, $\phi'_\alpha(t) < 0$ and $\phi''_\alpha(t) > 0$. For analytic $\phi_\alpha(\cdot)$ (which is satisfied by all common Archimedean copulas), the covariate effect model is identifiable [18, 19].

For Archimedean copulas, the local association measure $\tilde{\theta}(t_1, t_2)$ only dependents on the joint function F_z and not the marginal functions $F_{1,z}$ and $F_{2,z}$. Let $\bar{\theta}_\alpha(v) = -v\phi''_\alpha(v)/\phi'_\alpha(v)$ for the first two data structures, and $\bar{\theta}_\alpha(v) = -\phi'_\alpha(v)/[v\phi''_\alpha(v)]$ for the third data structure. Then $\tilde{\theta}_z(t_1, t_2) = \bar{\theta}_{\alpha(z,\beta)}[F_z(t_1, t_2)]$. Hence by (5) we can denote the regression model as

$$\tilde{\pi}_z(t_1, t_2) = \eta(z, \beta, \gamma_z). \qquad (7)$$

Here the nuisance parameter $\gamma_z = F_z(t_1, t_2)$ is the only time-varying component. Using the subsample with $Z = z$, we can get standard nonparametric estimation $\hat{\gamma}_z = \hat{F}_z(t_1, t_2)$. We then find the least-squares estimator $\hat{\beta}$ which minimize

$$U(\beta) = \sum_z \sum_{i<j} \{W_z(\tilde{X}_{1,ij}, \tilde{X}_{2,ij}) D_{ij}(z) [\Delta_{ij} - \eta(z, \beta, \hat{\gamma}_z(\tilde{X}_{1,ij}, \tilde{X}_{2,ij}))]^2\}. \qquad (8)$$

Here W is a positive weight function. $(\tilde{X}_{1,ij}, \tilde{X}_{2,ij})$ is defined similarly as $\tilde{\boldsymbol{T}}_{ij}$, and depends on the data type. For the first two data structures, $\tilde{X}_{k,ij} = X_{ki} \wedge X_{kj}$ $(k = 1, 2)$, and for truncation data, $\tilde{X}_{1,ij} = X_{1i} \vee X_{2j}$ and $\tilde{X}_{2,ij} = X_{2i} \wedge X_{2j}$. We use the weight function $W_{z,a,b}(x_1, x_2)$ of the form,

$$\frac{n_z}{\sum_{i=1}^n I\{X_{1i} \geq \min(a, x_1), X_{2i} \geq \min(b, x_2), Z_i = z\}}, \qquad (9)$$

where n_z is the sample size of $Z = z$; a and b are constants. With $a = b = 0$, the function reduces to $W_z^0 = 1$ which is the un-weighted case. With $a = b = \infty$, the weight function becomes $W_z^\infty = n_z / \sum_{i=1}^n I\{X_{1i} \geq x_1, X_{2i} \geq x_2, Z_i = z\}$. For truncation data, there is no information in the wedge $T_1 > T_2$, therefore, we consider alternative weight function $W_z^*(x_1, x_2) = n_z / \sum_{i=1}^n I\{X_{1i} \leq x_1, X_{2i} \geq x_2, Z_i = z\}$.

3 Simplification for Clayton Copula.

One type of Archimedean copulas is Clayton copula

$$C_\alpha(u_1, u_2) = (u_1^{1-\alpha} + u_2^{1-\alpha} - 1)_+^{1/(1-\alpha)} \ (\alpha > 0).$$

Under the Clayton family, the local odds ratio $\theta_z(u_1, u_2) = \alpha(z)$ is no longer time-varying. Hence (7) can be simplified as $\tilde{\pi}_z(t_1, t_2) = \eta(z, \beta)$, and the estimation (8) no longer require estimation of the time-varying nuisance parameter γ_z.

We now derive the asymptotic properties of the estimator $\hat{\beta}$ minimizing (8) under the Clayton copula assumption. Let β^* be the true value of β. We will show that $\hat{\beta}$ is a consistent estimator of β, and $\sqrt{n}(\hat{\beta} - \beta^*)$ converges to a multivariate normal distribution. Let the list of possible association covariate values be $\mathcal{Z} = \{z_1, ..., z_K\}$. We assume the following technical conditions:

(T1) n_z/n converge to a constant $0 < c_z < 1$ for each possible values $z \in \mathcal{Z}$.

(T2) The weight function $W_z(u, v)$ has a uniformly bounded deterministic limit $\tilde{W}_z(u, v)$.

(T3) The function $\eta(z, \beta)$ is twice differentiable against parameters β with uniformly bounded partial derivatives against β_k, $k = 1, 2, ..., p$. $\eta(z, \beta) = \eta(z, \beta^*)$ for all $z \in \mathcal{Z}$ only for $\beta = \beta^*$. When $|\beta| \to \infty$, $\eta(z, \beta) \to 0$ or 1.

(T4) Let $I(\beta) = (I_{ij})$ with $I_{ij} = E[-n^{-2}\partial^2 U(\beta)/\partial\beta_i\partial\beta_j]$, $i = 1, 2, ..., p, j = 1, 2, ..., p$ denote the Fisher information matrix. Let $\tilde{U}(\beta)$ denote the statistic $U(\beta)$ in (8) with the weight function W_z replaced by its deterministic limit $\tilde{W}_z$. Let $\tilde{I}(\beta) = \left(\tilde{I}_{ij}\right)$ with $\tilde{U}(\beta)$ replacing $U(\beta)$ in the definition of of $I(\beta)$. We assume that the Fisher information $\tilde{I}(\beta^*)$ is nonsingular at true parameter value $\beta = \beta^*$.

ISBN: 978-960-474-380-3

Theorem 1 *Under model (7) for Clayton copulas, the estimator $\hat{\beta}$ minimizing (8) consistent. And $\sqrt{n}(\hat{\beta} - \beta^*)$ converges in distribution to a multivariate normal distribution with variance Σ which can be consistently estimated by $\hat{\Sigma} = \hat{I}^{-1}\hat{J}(\hat{I}^{-1})'$, where $\hat{I} = \left(\hat{I}_{ij}\right)$ and $\hat{J} = \left(\hat{J}_{ij}\right)$ are matrices of dimension $p \times p$ with*

$$\hat{I}_{ij} = -\frac{1}{n^2}\frac{\partial^2 U(\beta)}{\partial\beta_i\partial\beta_j}|_{\beta=\hat{\beta}}, \tag{10}$$

$$\begin{aligned}\hat{J}_{ij} = \ & \tfrac{2}{n^3}\sum_z[\sum_{k<l<m}(\hat{Q}^{(i)}_{kl,z}\hat{Q}^{(j)}_{km,z} + \hat{Q}^{(i)}_{kl,z}\hat{Q}^{(j)}_{lm,z} \\ & +\hat{Q}^{(i)}_{lm,z}\hat{Q}^{(j)}_{km,z}) + \sum_{k<l}(\hat{Q}^{(i)}_{kl,z}\hat{Q}^{(j)}_{kl,z})],\end{aligned} \tag{11}$$

$$\begin{aligned}\hat{Q}^{(k)}_{ij,z} = \ & 2W_z(\tilde{X}_{1,ij}, \tilde{X}_{2,ij})D_{ij}(z) \\ & [\Delta_{ij} - \eta(z,\hat{\beta})](-\tfrac{\partial\eta(z,\beta)}{\partial\beta_k})|_{\beta=\hat{\beta}}.\end{aligned}$$

Proof: We first show that the statistics $n^{-2}U(\beta)$ in (8) has a positive limiting deterministic function in β that is minimized at $\beta = \beta^*$. Note that $\tilde{U}(\beta)$ is

$$\tilde{U}(\beta) = \sum_z\sum_{i<j}\tilde{W}_z(\tilde{X}_{ij}, \tilde{Y}_{ij})D_{ij}(z)[\Delta_{ij} - \eta(z,\beta)]^2.$$

Then

$$n^{-2}|U(\beta) - \tilde{U}(\beta)| \le \frac{1}{2}\sup_{z,u,v}|\tilde{W}_z(u,v) - W_z(u,v)|$$

which by condition (T2) converges to zero in probability. Therefore we only need to show that the limit of $\tilde{U}(\beta)$ in probability is positive and is minimized at $\beta = \beta^*$. Under the model assumptions, for those $Z_i = Z_j = z$ and $D_{ij}(z) = 1$, Δ_{ij} is a Bernoulli random variable with $\Pr(\Delta_{ij} = 1) = \eta(z,\beta^*)$. Hence $E\{[\Delta_{ij} - \eta(z,\beta)]^2 | Z_i = Z_j = z, D_{ij}(z) = 1\} = \eta(z,\beta^*)[1-\eta(z,\beta^*)] + [\eta(z,\beta^*) - \eta(z,\beta)]^2$. So this decompose $n_z^{-2}E\{\sum_{i<j}\tilde{W}_z(\tilde{X}_{ij},\tilde{Y}_{ij})D_{ij}(z)[\Delta_{ij} - \eta(z,\beta)]^2\}$ into two terms, by the law of large numbers, converges to positive constant d_z and $e_z(\beta)$ respectively. Then $E[n^{-2}\tilde{U}(\beta)] \to \sum_z c_z^2 d_z + \sum_z c_z^2 e_z(\beta)$, which is minimized if and only if $\sum_z c_z^2 e_z(\beta) = 0$ when $\beta = \beta^*$ by condition (T3). By the law of large numbers, we can see that $n^{-2}\tilde{U}(\beta)$ converges to its expectation. Hence $n^{-2}U(\beta)$ also converges to this limit which is uniquely minimized by $\beta = \beta^*$.

The consistency of $\hat{\beta}$ then comes from the uniform convergence of $n^{-2}U(\beta)$ within a neighborhood of $\beta = \beta^*$ and the condition (T3).

Let $u(\beta) = \nabla_\beta U(\beta)$ denote the gradient of $U(\beta)$. Since the local minimizer of $U(\beta)$ also solves $u(\beta) = 0$, without loss of generality, we can take $\hat{\beta}$ as a consistent root of $u(\beta) = 0$. Let $\tilde{u}(\beta)$ denote $u(\beta)$ with W_z replaced by $\tilde{W}_z$. Under model (7), the limit of $E[u(\beta^*)]$ is $E[\tilde{u}(\beta^*)] = 0$. And the limit of $I(\beta^*)$ is the nonsingular matrix $\tilde{I}(\beta^*)$. So without loss of generality, by Taylor expansion:

$$\begin{aligned}\sqrt{n}(\hat{\beta} - \beta^*) &= [\tilde{I}(\beta^*)]^{-1}n^{-3/2}u(\beta^*) + o_p(1) \\ &= \left(n^{-3/2}\textstyle\sum_z\sum_{i<j}Q^{(k)}_{ij,z}\right) + o_p(1),\end{aligned}$$

where $Q^{(k)}_{ij,z} = 2\tilde{W}_z(\tilde{X}_{ij},\tilde{Y}_{ij})D_{ij}(z)[\Delta_{ij} - \eta(z,\beta^*)](-\nabla_\beta\eta(z,\beta)|_{\beta=\beta^*})$. By the central limit theorem for U-statistic and Slutsky's theorem: $\sqrt{n}(\hat{\beta} - \beta^*)$ converges in distribution to a multivariate normal distribution with variance Σ which is consistently estimated by $\hat{\Sigma} = \hat{I}^{-1}\hat{J}(\hat{I}^{-1})'$, where $\hat{I}$ and $\hat{I}$ are defined in (10) and (11). □

3.1 Checking the Clayton Assumption

We propose a generalized version of Shih [20]'s test to verify the Clayton assumption for all three data structures. Let $U_1(\beta)$ and $U_2(\beta)$ follow the same form as $U(\beta)$ with W_z being specified as two different weight functions $W_{z,1}$ and $W_{z,2}$ respectively. We will use $W_{z,1} = W_z^0$, and use W_z^∞ or W_z^* as $W_{z,2}$ according to the data structures. Let $\hat{\beta}_{W_{z,k}}$ be the solution to $U_k(\beta) = 0$ $(k = 1, 2)$. The proposed test statistic can be expressed as

$$T = n(\hat{\beta}_{W_{z,1}} - \hat{\beta}_{W_{z,2}})'\hat{\Gamma}^{-1}(\hat{\beta}_{W_{z,1}} - \hat{\beta}_{W_{z,2}}),$$

where $\hat{\Gamma} = \left(\hat{\Gamma}_{ij}\right)$,

$$\begin{aligned}\hat{\Gamma}_{ij} = \ & n^{-3}\sum_z[2\sum_{k<l<m}(\hat{Q}^{*(i)}_{kl,z}\hat{Q}^{*(j)}_{km,z} + \hat{Q}^{*(i)}_{kl,z}\hat{Q}^{*(j)}_{lm,z} \\ & +\hat{Q}^{*(i)}_{lm,z}\hat{Q}^{*(j)}_{km,z}) + \sum_{k<l}(\hat{Q}^{*(i)}_{kl,z}\hat{Q}^{*(j)}_{kl,z})],\end{aligned} \tag{12}$$

and $\hat{Q}^{*(i)}_{kl,z}$ is defined in the proof below.

Theorem 2 *Under conditions for Theorem 1, if the Clayton model is correctly specified, T converges in distribution to χ^2_{p+1}. That is, for a γ-level test, we reject the null hypothesis if $T > \chi^2_{p+1,\gamma}$, where $Pr(\chi^2_{p+1} > \chi^2_{p+1,\gamma}) = \gamma$.*

Proof: Under model (7), the distributions of $\hat{\beta}_{W_{z,1}}$ and $\hat{\beta}_{W_{z,2}}$ are centered around the same β^*. By the results of Theorem 1, we have

$$\sqrt{n}(\hat{\beta}_{W_{z,1}} - \hat{\beta}_{W_{z,2}}) = n^{-3/2}\sum_z\sum_{i<j}\tilde{Q}^*_{ij,z} + o_p(1),$$

where $\tilde{Q}^*_{ij,z} = \left(Q^{*(k)}_{ij,z}\right) = I_1^{-1}\tilde{Q}_{1,ij,z} - I_2^{-1}\tilde{Q}_{2,ij,z}$, $\tilde{Q}_{m,ij,z} = \left(Q^{(k)}_{m,ij,z}\right)$, I_m is $\tilde{I}(\beta^*)$ with W_z replaced

by $W_{z,m}$, and $Q^{(k)}_{m,ij,z}$ is $Q^{(k)}_{ij,z}$ with $\tilde{W}_z$ replaced by $\tilde{W}_{z,m}$ $(m = 1, 2)$. By the central limit theorem for U-statistic and Slutsky's theorem: $\sqrt{n}(\hat{\beta}_{W_{z,1}} - \hat{\beta}_{W_{z,2}})$ converges in distribution to a mean-zero multivariate normal distribution with variance Γ which can be consistently estimated by $\hat{\Gamma} = \left(\hat{\Gamma}_{ij}\right)$, with $\hat{\Gamma}_{ij}$ defined in (12). Therefore, $T = n(\hat{\beta}_{W_{z,1}} - \hat{\beta}_{W_{z,2}})'\hat{\Gamma}^{-1}(\hat{\beta}_{W_{z,1}} - \hat{\beta}_{W_{z,2}})$ converges in distribution to χ^2_{p+1}.

4 Numerical Studies

4.1 Simulations

We performed simulations to assess finite-sample performances of the proposed methods. Three types of covariates are considered: $Z^{(k)}$ affects the marginal distribution of T_k for $k = 1, 2$ and Z affects the association structure. For marginal models, we let $T_k = Z^{(k)}\mu + \varepsilon_k$ for $k = 1, 2$, where $Z^{(k)}$ was generated from $Bernoulli(0.5)$ for $k = 1, 2$, ε_1 follow $\exp(0.8)$ and ε_2 follow $\exp(1)$, and set $\mu = 0.5$. For the model on association, we consider the Clayton copula with $\alpha(Z) = \exp(\beta_0 Z_0 + \beta_1 Z_1)$, where $Z = (Z_0, Z_1) = (1, 0)'$ or $(1, 1)'$. Each covariate group was generated with equal probability. We report the results when Z and $Z^{(k)}$ $(k = 1, 2)$ were generated independently. The case of common covariates ($Z = Z^{(1)} = Z^{(2)}$) produces similar results and hence is not reported. The latter two data structures were created by imposing a censoring or truncation relationship between T_1 and T_2 respectively. Right censoring is incorporated in the three data structures. For bivariate censored data, we set C_1 to be independent of C_2 and $C_k|Z^{(k)} = z \sim z\mu + U(0, 6)$. For semi-competing risks data and truncation data, we set $C|Z^{(2)} = z \sim z\mu + U(0, 6)$. More simulations are conducted with marginal accelerated failure times model and more covariate groups. The results are similar and thus are omitted from the report here.

For bivariate censored data, the censoring proportion of T_k $(k = 1, 2)$ is around 0.16. For semi-competing risks data, the censoring rate for T_1 which is subject to dependent censoring by T_2 and independent censoring by C varies from 0.11 ($\tau = 0.76$, $Z^{(1)} = 1$, $Z^{(2)} = 0$) to 0.88 ($\tau = 0.76$, $Z^{(1)} = 0$, $Z^{(2)} = 1$). For truncation data, the missing proportion $\Pr(T_1 > T_2)$ and the censoring rate $\Pr(T_2 > C|T_1 \leq T_2)$ vary with τ, $Z^{(1)}$ and $Z^{(2)}$. $\Pr(T_1 > T_2) \approx 0.06$, $\Pr(T_2 > C|T_1 \leq T_2) \approx 0.16$ when $\tau = 0.76$, $Z^{(1)} = 1$, $Z^{(2)} = 0$ and $\Pr(T_1 > T_2) \approx 0.83$, $\Pr(T_2 > C|T_1 \leq T_2) \approx 0.41$ when $\tau = 0.76$, $Z^{(1)} = 0$, $Z^{(2)} = 1$.

Tables 1 summarizes the results of the proposed estimator $\hat{\beta}$. For each data setting, we applied two weight functions W^0_z versus W^∞_z or W^*_z. Based on 1000 replications, we computed its root mean square error ($\sqrt{SE}$) and the coverage probability of the nominal 0.95 confidence interval (Cov). In general the proposed estimator for the regression parameter has nice performance in all the settings, and the 95% confidence intervals have reasonable coverage probabilities. The weight functions W^∞_z or W^*_z results in better estimation performance than the weight W^0_z.

Table 1: Simulation results.

			$n = 150$		$n = 300$	
Data	W	β_0 β_1	$\sqrt{SE}$	Cov	$\sqrt{SE}$	Cov
1	W^0_z	0.5	0.21	0.952	0.14	0.948
		0.5	0.30	0.955	0.20	0.951
		1	0.21	0.949	0.14	0.956
		1	0.32	0.947	0.21	0.953
	W^∞_z	0.5	0.19	0.947	0.13	0.953
		0.5	0.28	0.950	0.18	0.951
		1	0.20	0.941	0.13	0.950
		1	0.31	0.945	0.20	0.952
2	W^0_z	0.5	0.26	0.955	0.16	0.947
		0.5	0.37	0.949	0.24	0.949
		1	0.27	0.952	0.17	0.951
		1	0.40	0.947	0.26	0.944
	W^∞_z	0.5	0.24	0.948	0.16	0.949
		0.5	0.34	0.950	0.22	0.951
		1	0.25	0.941	0.16	0.943
		1	0.39	0.942	0.24	0.946
3	W^0_z	0.5	0.18	0.954	0.11	0.951
		0.5	0.29	0.944	0.18	0.944
		1	0.23	0.955	0.14	0.953
		1	0.43	0.951	0.28	0.955
	W^*_z	0.5	0.15	0.951	0.09	0.946
		0.5	0.24	0.943	0.14	0.943
		1	0.18	0.952	0.10	0.953
		1	0.36	0.945	0.21	0.961

Data: the type of data structure; W: the weight function.

4.2 Otitis Media Data Analysis

We applied the proposed methods to the otitis media clinical trial data [21] which belong to typical bivariate censored data. The data set collected the information of 78 children with age from 6 months to 8 years who developed chronic otitis media effusion between February 1987 and January 1990. Here (T_1, T_2) are the times (in months) to failure of ventilating tubes

ISBN: 978-960-474-380-3

surgically inserted into right and left ears respectively and $C_1 = C_2$ is the time to the end of study. The covariate Z is a binary variable with $Z = 1$ indicating that a subject was treated by oral antibiotic and $Z = 0$ if the subject was assigned to a placebo control group. The assignment for the treatment type was random. Applying the log-rank test to examine whether the treatment type affects the marginal distribution of the time to failure of ventilating tubes, we found that the result was significant (with p-value=0.0093) for the right ear, but not significant (with p-value=0.659) for the left ear. Without specifying the marginal regression models, we investigated whether and how the association between the failure times to two ears differs in the two treatment groups.

Applying the testing procedure discussed in Section 3.1, $T = 5.069$ $(d.f. = 2)$ which corresponds to p-value=0.079. This implies that the Clayton model is still acceptable for this data set. Accordingly we assume that $\theta_a(s, t; Z) = \exp(\beta_0 + \beta_1 Z)$, where $\exp(\beta_0)$ is the odds ratio for the baseline (placebo) group and $\exp(\beta_1)$ represents the difference of the odds ratio between the antibiotic group and the baseline group. For weights W_z^{∞}, the estimators and the corresponding standard errors given in the parentheses for β are $\hat{\beta}_0$=0.5966 (0.2068), $\hat{\beta}_1$=-0.4871 (0.2767), and $\hat{\beta}_0 + \hat{\beta}_1$=0.1095 (0.1838). Correspondingly the odds for placebo and antibiotic groups are $\exp(\hat{\beta}_0)$=1.8159 and $\exp(\hat{\beta}_0 + \hat{\beta}_1)$=1.1157 which are both greater than 1 indicating positive association between T_1 and T_2 but only the former is significant. The odds ratio between the two groups is $\exp(\hat{\beta}_1) = 0.6144$ which is significant at 0.1 level providing some evidence that the level of association differs in the two groups, but which is not significant at 0.05 level.

References:

[1] J.H. Shih and T.A. Louis, Inference on the association parameter in copula models for bivariate survival data. *Biometrics* 51, 1995, 1384-99.

[2] W. Wang and M.T. Wells, Model selection and semiparametric inference for bivariate failure time data. *J. Am. Statist. Assoc.* 95, 2000, 62-72.

[3] R. Day, J. Bryant and M. Lefkopolou, Adaptation of bivariate frailty models for prediction, with applications to biological markers as prognostic indicators. *Biometrika* 84, 1997, 45–56.

[4] J.P. Fine, H. Jiang and R. Chappell, On semicompeting risks data. *Biometrika* 88, 2001, 907-919.

[5] W. Wang, Estimation of the association parameter for copula models under dependent censoring. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 65, 2003, 257-273.

[6] H. Jiang, J. P. Fine and R. Chappell, Semiparametric Analysis of Survival Data with Left Truncation and Dependent Right Censoring. *Biometrics* 61, 2005, 567-575.

[7] L. Peng and J.P. Fine, Regression modeling of semicompeting risks data. *Biometrics* 63, 2007, 96-108.

[8] L. Lakhal, L.-P. Rivest and B. Abdous, Estimating Survival and Association in a Semicompeting Risks Model. *Biometrics* 64, 2008, 180-188.

[9] J. Hsieh, W. Wang, and A. A. Ding, Regression Analysis based on Semicompeting Risks Data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 70, 2008, 3-20.

[10] L. Chaieb, L.-P. Rivest, and B. Abdous, Estimating Survival under a Dependent Truncation. *Biometrika* 93, 2006, 655-669.

[11] T. Emura, W. Wang and H.N. Hung, Semiparametric inference for copula models for truncated data. *Statistica Sinica* 21, 2011, 349-367.

[12] Y-H. Chen, Maximum likelihood analysis of semicompeting risks data with semiparametric regression models. *Lifetime Data Anal* 18, 2012, 3657.

[13] A. A. Ding, J. J. Hsieh and W. Wang, Local linear estimation of concordance probability with application to covariate effects models on association for bivariate failure-time data. *Lifetime Data Anal.*, 2013, 1-33.

[14] D. Oakes, Bivariate Survival Models Induced by Frailties. *J. Amer. Statist. Assoc.* 84, 1989, 487-493.

[15] J. Fan, L. Hsu and R.L. Prentice, Dependence estimation over a finite bivariate failure time region, *Lifetime Data Anal.* 6, 2000, 343-355.

[16] T. Hu, B. Nan, X. Lin and J. Robins, Time dependent cross-ratio estimation for bivariate failure times, *biometrika* 98, 2011, 341-354.

[17] R. B. Nelsen, *An Introduction to Copulas* Springer – New York. 2006

[18] A. A. Ding, Identifiability Conditions for Covariate Effects Model on Survival Times under Informative Censoring. *Statistics and Probability Letters* 80, 2010, 911-915.

[19] A. A. Ding, Copula Identifiability Conditions for Dependent Truncated Data Model. *Lifetime Data Anal.* 18, 2012, 397-407.

[20] J.H. Shih, A goodness-of-fit test for association in a bivariate survival model. *Biometrika* 85, 1998, 189-200.

[21] C. T. Le and B. Lindgren, Duration of Ventilating Tubes: A Test for Comparing Two Clustered Samples of Censored Data. *Biometrics* 52, 1996, 328-334.