Published in final edited form as:

Lifetime Data Anal. 2008 June ; 14(2): 134–153. doi:10.1007/s10985-007-9058-9.

Efficient estimation for the proportional hazards model with bivariate current status data

Lianming Wang,

Biostatistics Branch, National Institute of Environmental Health Sciences, MD A3-03, P.O. Box 12233, Research Triangle Park, NC 27709, USA

Jianguo Sun, and

Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211, USA

Xingwei Tong

Department of Statistics and Financial Mathematics, Beijing Normal University, Beijing 100875, P.R. China

Lianming Wang: wangl3@niehs.nih.gov

Abstract

We consider efficient estimation of regression and association parameters jointly for bivariate current status data with the marginal proportional hazards model. Current status data occur in many fields including demographical studies and tumorigenicity experiments and several approaches have been proposed for regression analysis of univariate current status data. We discuss bivariate current status data and propose an efficient score estimation approach for the problem. In the approach, the copula model is used for joint survival function with the survival times assumed to follow the proportional hazards model marginally. Simulation studies are performed to evaluate the proposed estimates and suggest that the approach works well in practical situations. A real life data application is provided for illustration.

Keywords

Bivariate current status data; Efficient estimation; Counting processes; Sieve method; Copula model

1 Introduction

Current status data arise in many fields including demographical studies, epidemiology and tumorigenicity experiments (Keiding 1991; Sun 2006). In these situations, the failure time of interest is not observed but known only to be smaller or larger than an observation or a monitoring time. Bivariate current status data occur if there exist two related failure times of interest and only current status data are available for both times. In this paper, we consider regression analysis of bivariate current status data using the efficient estimation approach

(Bickel et al. 1993; Huang 1996). In practice, current status data are also often referred to as case I interval-censored data (Groeneboom and Wellner 1992).

Regression analysis of univariate current status data has been studied by a number of authors. For example, Huang (1996) proposed to use the efficient estimation approach under the proportional hazards model and Huang and Rossini (1997) studied the use of the sieve method under the proportional odds model. Using the additive hazards model, Lin et al. (1998) gave an estimating equation approach and Martinussen and Scheike (2002) derived the efficient estimates of regression parameters.

For bivariate current status data, in addition to the problems considered in the case of univariate data, some other problems of interest include estimation of the association parameter and the testing of independence. For example, Wang and Ding (2000) and Ding and Wang (2004) investigated these two problems, respectively. Other references about bivariate current status data include Dunson and Dinse (2002), which discussed the use of Bayesian approaches, and Jewell et al. (2005), which considered estimation of univariate cumulative distribution functions.

Let T_1 and T_2 denote the two failure times of interest and suppose that each study subject is observed or monitored only once at time C , which is usually referred to as univariate censoring and common for bivariate current status (Jewell et al. 2005; Wang and Ding 2000). Define $\delta_1 = I(T_1 \leq C)$ and $\delta_2 = I(T_2 \leq C)$, the censoring indicators and suppose that there exists a p -dimensional time independent covariate vector X . Thus, for each subject, the data structure is $(C, \delta_1, \delta_2, X)$. In the following, we assume that the joint survival function of (T_1, T_2) is specified by a copula model as

$$S(s, t) = C_\alpha(S_1(s), S_2(t)).$$

Here S_1 and S_2 denote the marginal survival functions of T_1 and T_2 , respectively, C_α is a genuine survival function on the unit square, and $\alpha \in R$ is a global association parameter.

The copula model is one of the commonly used models for bivariate data (Clayton 1978; Genest and MacKay 1986; Genest and Rivest 1993; Shih and Louis 1995; Sun et al. 2006). One attractive feature of this model is its modeling flexibility since it includes as special cases many useful bivariate failure time models such as the Archimedean copula family

$$C_\alpha(u, v) = \varphi_\alpha\{\varphi_\alpha^{-1}(u) + \varphi_\alpha^{-1}(v)\}, \quad 0 \leq u, v \leq 1,$$

where $0 < \varphi_\alpha < 1$, $\varphi_\alpha(0) = 1$, $\varphi_\alpha' < 0$, and $\varphi_\alpha'' > 0$. Here $\varphi_\alpha'(u) = \frac{d}{du}\varphi_\alpha(u)$ and $\varphi_\alpha''(u) = \frac{d}{du}\varphi_\alpha'(u)$. Taking $\varphi_\alpha(u) = (1 + u)^{1/(1-\alpha)}$, the Laplace transformation of a gamma distribution, we obtain the Clayton family (Clayton 1978),

$$C_{\alpha}(u, v) = (u^{1-\alpha} + v^{1-\alpha} - 1)^{1/(1-\alpha)},$$

which models positive association with parameter $\alpha > 1$. The same form of Clayton model can be generalized to model negative association with $0 < \alpha < 1$ with some modifications (Genest and MacKay 1986). Another attractive feature of copula models is that the marginal distributions do not depend on the choice of the association parameter and thus one can model the marginal distributions and the association parameter separately.

The reminder of this paper is organized as follows. In Sect. 2, we define the notation and describe the assumptions used below. Section 3 derives the efficient score and information for regression and association parameters and Sect. 4 presents the estimation procedure. Simulation studies are conducted in Sect. 5 and followed by an application in Sect. 6. Section 7 concludes with some remarks.

2 Notation and model specification

Let T_1 and T_2 be defined as above and suppose that they follow the proportional hazards model marginally (Cox 1972) with

$$S_k(t) = \exp(-\Lambda_{0k}(t) \exp(\beta' X)), \quad k=1, 2.$$

In the above, β is a p -dimensional vector of regression parameters representing covariate effects on T_k and $\Lambda_{0k}(t)$ denotes the baseline cumulative hazard function of T_k , $k = 1, 2$. Note that here we assume that the covariate effect is the same on T_1 and T_2 and the inference approach can be developed similarly otherwise.

In the following, we will focus on situations where C is a continuous variable defined in a finite interval $(0, M_0]$, where M_0 is a constant such that $P(T_1 \leq M_0) > 0$ and $P(T_2 \leq M_0) > 0$. For real data, M_0 can be taken to be the largest observation time. Sometimes it is natural to define $C = \min(C', M_0)$, where C' is a positive random variable. It will be assumed that C is independent of T_1 and T_2 given X and has the hazard function given by $\lambda_c(t|X) = \lambda_{c0}(t) \exp(\gamma' X)$, where γ denotes the possible covariate effect on C . An advantage of this Cox model assumption for censoring time C is that one does not need to estimate the baseline hazard function $\lambda_{c0}(t)$ in the estimation procedure proposed below.

Define counting processes

$$\begin{aligned} N_{11}(t) &= \delta_1 \delta_2 I(C \leq t), \quad N_{10}(t) = \delta_1 (1 - \delta_2) I(C \leq t), \\ N_{01}(t) &= (1 - \delta_1) \delta_2 I(C \leq t), \quad N_{00}(t) = (1 - \delta_1) (1 - \delta_2) I(C \leq t). \end{aligned}$$

It can be easily shown using the arguments similar to those in Lin et al. (1998) that the intensity process of N_{jm} is

$$Y(t)\lambda_c(t|X)S_{jm}(t,\theta), \quad j,m=0,1.$$

In the above, $Y(t) = I(C \leq t)$, $\theta = (\beta, \alpha)'$,

$$\begin{aligned} S_{11}(\theta, t) &= P(T_1 \leq t, T_2 \leq t) = 1 - S_1(t) - S_2(t) + C_\alpha(S_1(t), S_2(t)), \\ S_{01}(\theta, t) &= P(T_1 > t, T_2 \leq t) = S_1(t) - C_\alpha(S_1(t), S_2(t)), \\ S_{10}(\theta, t) &= P(T_1 \leq t, T_2 > t) = S_2(t) - C_\alpha(S_1(t), S_2(t)), \\ S_{00}(\theta, t) &= P(T_1 > t, T_2 > t) = C_\alpha(S_1(t), S_2(t)). \end{aligned}$$

Under the above assumptions, up to an additive constant, the log-likelihood contribution is given by

$$\begin{aligned} l(\theta, \Lambda_{01}, \Lambda_{02}) &= \delta_1 \delta_2 \log(S_{11}(\theta, C)) + (1 - \delta_1) \delta_2 \log(S_{01}(\theta, C)) + \delta_1 (1 - \delta_2) \log(S_{10}(\theta, C)) + (1 - \delta_1)(1 - \delta_2) \log(S_{00}(\theta, C)) \\ &= \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \log S_{jm}(\theta, t) dN_{jm}(t). \end{aligned} \quad (1)$$

3 Efficient score and information bound

This section will consider the efficient score and information for the parameter θ (Bickel et al. 1993). For this, consider general parametric submodels of Λ_{01} and Λ_{02} characterized by a parameter η . Let $\dot{l}_\theta$ and $\dot{l}_\eta$ denote the derivatives of the log-likelihood $l(\theta, \Lambda_{01}, \Lambda_{02})$ given in (1) with respect to θ and η , respectively. Also let $\dot{l}_\eta^*$ denote the projection of $\dot{l}_\theta$ onto the linear span formed by all possible $\dot{l}_\eta$. Then the efficient score for θ is given by $\dot{l}_{\theta^*} = \dot{l}_\theta - \dot{l}_\eta^*$ that satisfies

$$E(\dot{l}_{\theta^*} * \dot{l}_\eta) = 0 \quad (2)$$

(Huang 1996; Martinussen and Scheike 2002).

To derive $\dot{l}_{\theta^*}$ and obtain the information, in the following, we will first consider the situation where T_1 and T_2 have the same marginal survival function. We then discuss the situation where the marginal survival functions may be different.

3.1 Efficient score with common marginal survival functions

In this subsection, we assume that $\Lambda_{01}(t) = \Lambda_{02}(t)$ and let $\Lambda_0(t)$ denote the common baseline cumulative hazard function. In this case, we can write $C_\alpha(S_1(t), S_2(t))$ as $C_\alpha(S_1(t))$ for simplicity. Define

$$\begin{aligned} D_\alpha &= \frac{\partial C_\alpha(S_1)}{\partial \alpha}, \quad D_u = \frac{\partial C_\alpha(u)}{\partial u} \Big|_{u=S_1}, \\ a_{jm} &= -\exp(\beta' X) S_1 \Lambda_0 [(-1)^j m + (-1)^m j + 2(-1)^{m+j} D_u], \end{aligned}$$

and

$$Z_{jm} = (X', (-1)^{m+j} D_\alpha / a_{jm})', \quad j, m = 0, 1.$$

First it is easy to show that for $j, m = 0, 1$,

$$S_{jm}(\theta, t) = j m + (-1)^j m S_1 + (-1)^m j S_1 + (-1)^{m+j} C_\alpha(S_1).$$

This yields $S_{jm}(\theta, t) / \theta = Z_{jm} a_{jm}$. Using the notation defined above, we can write the score function of θ as

$$\dot{l}_\theta = \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{Z_{jm} a_{jm}}{S_{jm}} dN_{jm}(t) = \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{Z_{jm} a_{jm}}{S_{jm}} dM_{jm}(t).$$

In the above,

$$dM_{jm}(t) = dN_{jm}(t) - Y(t) \lambda_c(t) S_{jm} dt,$$

which is a martingale, $j, m = 0, 1$, and the second equality results from the fact that

$$\sum_{j=0}^1 \sum_{m=0}^1 S_{jm}(\theta, t) = 1, \text{ which implies that } \sum_{j=0}^1 \sum_{m=0}^1 Z_{jm} a_{jm} = 0.$$

Suppose $\log \Lambda_0(t) / \eta = b(t)$. Then we have

$$\dot{l}_\eta = \dot{l}_\eta(b) = \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{b a_{jm}}{S_{jm}} dN_{jm} = \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{b a_{jm}}{S_{jm}} dM_{jm}.$$

To find $\dot{l}_{\theta^*}$, we need to find b^* such that

$$\begin{aligned} 0 &= E(\dot{l}_\theta - \dot{l}_\eta(b^*)) \dot{l}_\eta(b) \\ &= \sum_{j=0}^1 \sum_{m=0}^1 E \left\{ \int_0^{M_0} \frac{(Z_{jm} - b^*) a_{jm}}{S_{jm}} dM_{jm}(t) \int_0^{M_0} \frac{b a_{jm}}{S_{jm}} dM_{jm}(t) \right\} \\ &= \sum_{j=0}^1 \sum_{m=0}^1 E \left\{ \int_0^{M_0} \frac{(Z_{jm} - b^*) a_{jm}^2}{S_{jm}} b Y \lambda_c dt \right\} \end{aligned}$$

for any b . To this end, one has,

$$\sum_{j=0}^1 \sum_{m=0}^1 E \left[\frac{(Z_{jm} - b^*) a_{jm}^2 Y \lambda_c}{S_{jm}} \right] = 0$$

for any $t \in (0, M_0]$. Solving this equation, we can obtain that

$$b^* = \frac{\sum_{j,m} E[Z_{jm} a_{jm}^2 Y \lambda_c / S_{jm}]}{\sum_{j,m} E[a_{jm}^2 Y \lambda_c / S_{jm}]}.$$

It follows that the efficient score for θ has the form

$$\dot{l}_{\theta^*} = \dot{l}_{\theta} - \dot{l}_{\eta}(b^*) = \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jm} - b^*) a_{jm}}{S_{jm}} dN_{jm}(t) \quad (3)$$

and the information for θ is given by

$$I(\theta) = E \dot{l}_{\theta^*}^{\otimes 2} = \sum_{j=0}^1 \sum_{m=0}^1 E \int_0^{M_0} ((Z_{jm} - b^*) a_{jm})^{\otimes 2} S_{jm}^{-1} Y \lambda_c dt, \quad (4)$$

where $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$, and $a^{\otimes 2} = a a'$.

Remember that $\lambda_c = \lambda_{c0} \exp(\gamma' X)$ as assumed in Sect. 2. Thus λ_{c0} can be canceled out in the above general formula for b^* since λ_{c0} is contained in both the numerator and denominator and free of any random variables. In other words, we have

$$b^* = \frac{\sum_{j,m} E[Z_{jm} a_{jm}^2 Y \exp(\gamma' X) / S_{jm}]}{\sum_{j,m} E[a_{jm}^2 Y \exp(\gamma' X) / S_{jm}]}.$$

3.2 Efficient score with different marginal survival functions

Now we consider the situation where Λ_{01} and Λ_{02} may be different. Let D_u and D_v denote the derivatives of $C_a(u, v)$ with respect to u and v and at $u = S_1$ and $v = S_2$, respectively.

For $j, m = 0, 1$, define

$$\begin{aligned} a_{jm}^{(1)} &= (-1)^j \exp(\beta' X) \Lambda_{01} S_1 [m + (-1)^m D_u], \\ a_{jm}^{(2)} &= (-1)^m \exp(\beta' X) \Lambda_{02} S_2 [j + (-1)^j D_v], \end{aligned}$$

and $B_{jm} = (a_{jm}^{(1)} + a_{jm}^{(2)})^{-1} (-1)^{m+j} D_{\alpha}$. Also for $j, m = 0, 1$, let $a_{jm} = (a_{jm}^{(1)}, a_{jm}^{(2)})'$ and define

$$Z_{jm} = \begin{pmatrix} X & X \\ B_{jm} & B_{jm} \end{pmatrix},$$

a $(p + 1) \times 2$ matrix. Suppose that $\frac{\partial}{\partial \eta} \log \Lambda_{01}(t) = b_1(t)$ and $\frac{\partial}{\partial \eta} \log \Lambda_{02}(t) = b_2(t)$ and let $b = (b_1, b_2)$. Note that here for simplicity, we use the same notation a_{jm} , Z_{jm} and b as in the previous section. Then we have

$$\frac{\partial S_{jm}(\theta, t)}{\partial \theta} = Z_{jm} \cdot a_{jm}$$

and

$$\frac{\partial S_{jm}(\theta, t)}{\partial \eta} = b \cdot a_{jm}.$$

To derive the efficient score function and information for θ , following the same procedure as that in the previous section, we can show that they have the same forms as given in (3) and (4), respectively, with a_{jm} and Z_{jm} defined in this subsection and $b^* = H^* G^{-1}$, where,

$$G = \sum_{j=0}^1 \sum_{m=0}^1 E \left[a_{jm}^{\otimes 2} Y \lambda_c / S_{jm} \right]$$

and

$$H = \sum_{j=0}^1 \sum_{m=0}^1 E \left[Z_{jm} a_{jm}^{\otimes 2} Y \lambda_c / S_{jm} \right],$$

which are 2×2 and $(p + 1) \times 2$ matrices, respectively.

For the same reason as mentioned in the previous subsection, we can simply cancel out the baseline hazard function λ_{c0} in both G and H defined above when computing b^* .

4 Estimation procedure

Consider a survival study involving n independent subjects and in which the observed data have the form $\{(C_i, \delta_{1i}, \delta_{2i}, X_i); i = 1, \dots, n\}$. Let a_{jmi} , S_{jmi} , Z_{jmi} , and Y_i be defined as in the previous section but with respect to subject i . To estimate, it is natural to use the empirical version of the efficient score function derived in the previous section, which has the form

$$\dot{l}_{\theta^*,n}(\theta, \Lambda_{01}, \Lambda_{02}, \gamma) = \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jmi} - b_n^*) a_{jmi}}{S_{jmi}} dN_{jmi}(t), \quad (5)$$

and to define the estimate as the solution to $\dot{l}_{\theta^*,n}(\theta, \Lambda_{01}, \Lambda_{02}, \gamma) = 0$ assuming that Λ_{01} , Λ_{02} and γ are known. In (5),

$$b_n^* = \frac{\sum_{i=1}^n \sum_{j,m} [Z_{jmi} a_{jmi}^2 Y_i \exp(\gamma' X_i) / S_{jmi}]}{\sum_{i=1}^n \sum_{j,m} [a_{jmi}^2 Y_i \exp(\gamma' X_i) / S_{jmi}]}$$

for the case where $\Lambda_{01}(t) = \Lambda_{02}(t)$ and

$$b_n^* = H_n(\Lambda_{01}, \Lambda_{02}, \gamma, t) * G_n^{-1}(\Lambda_{01}, \Lambda_{02}, \gamma, t)$$

otherwise, where

$$G_n = \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \left[a_{jmi}^{\otimes 2} Y_i \exp(\gamma' X_i) / S_{jmi} \right]$$

and

$$H_n = \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \left[Z_{jmi} a_{jmi}^{\otimes 2} Y_i \exp(\gamma' X_i) / S_{jmi} \right].$$

Suppose that some consistent estimates denoted by $\hat{\Lambda}_{01}$ and $\hat{\Lambda}_{02}$ of Λ_{01} and Λ_{02} are available with $n^{1/3}$ convergence rate and an estimate $\hat{\gamma}$ of γ with convergence rate $n^{1/2}$. Let $U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma})$ denote the score function in (5) with Λ_{01} , Λ_{02} and γ replaced by $\hat{\Lambda}_{01}$, $\hat{\Lambda}_{02}$ and $\hat{\gamma}$ respectively. It can be shown that the estimate $\hat{\theta}$ defined as the solution to the equation $U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma}) = 0$ is consistent and has an asymptotic normal distribution. Furthermore, it is efficient and its variance can be consistently estimated by $\hat{I}_n^{-1}(\hat{\theta})$ with replacing Λ_{01} , Λ_{02} , and γ by their estimates, where

$$\hat{I}_n(\hat{\theta}) = \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \left(\frac{(\hat{Z}_{jmi} - \hat{b}_n^*) \hat{a}_{jmi}}{\hat{S}_{jmi}} \right)^{\otimes 2} dN_{jmi}(t). \quad (6)$$

The proof is sketched in the Appendix.

To implement the estimation procedure above, we need $\hat{\gamma}$, $\hat{\Lambda}_{01}$ and $\hat{\Lambda}_{02}$. Since we have complete or right-censored data available for C in the current status data, partial likelihood

approach gives an estimate $\hat{\gamma}$ of γ with convergence rate $n^{1/2}$ (Anderson and Gill 1982). To obtain proper $\hat{\Lambda}_{01}$ and $\hat{\Lambda}_{02}$, one way is to use the univariate non-parametric maximum likelihood estimate discussed in Huang (1996). The resulted estimates have a convergence rate $n^{1/3}$, but no close form. Here we propose to use the sieve method to obtain the needed consistent estimates, which approximates $\Lambda_0(t)$ or Λ_{01} and Λ_{02} by step functions with $k = n^\varepsilon$ jump time points, where $0 < \varepsilon < 1$. The approach has been studied by many authors including Rossini and Tsiatis (1996) and Huang and Rossini (1997), who considered regression analysis of univariate current status data using the proportional odds model.

For simplicity, in the following, we only describe the estimation procedure for θ when Λ_{01} and Λ_{02} are same. The procedure is similar when Λ_{01} and Λ_{02} are different.

Step 1: Use the marginal likelihood approach to obtain a consistent estimator of the baseline cumulative hazard function Λ_0 in terms of k parameters (Kim and Xue 2002). Specifically, we estimate β and Λ_0 by maximizing the full likelihood given by assuming that T_1 and T_2 are independent. The same approach has been used by Goggins and Finkelstein (2000) and Wei et al. (1989) among others. Denote the estimates obtained by $(\hat{\Lambda}, \hat{\beta})$.

Step 2: Obtain $\hat{\gamma}$ using the partial likelihood approach based on complete or right-censored data on C and calculate $\hat{b}_n^*$ with replacing Λ_0 by $\hat{\Lambda}$ given in step 1 and γ by $\hat{\gamma}$.

Step 3: Solve the efficient estimating equation $U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma}) = 0$ for θ and calculate the variance estimator $\hat{I}(\theta)^{-1}$ through (6). In this step, one can use $\hat{\beta}$ given in step 1 as the initial value for β .

5 Simulation studies

Several simulation studies were conducted to assess the performance of the proposed methods in terms of efficiency and robustness. In particular, we considered the general performance, comparison with univariate regression, and the performance when the association parameter depends on covariates and there exists model misspecification. In these studies, we focused on the two sample situation with $X = 0$ or 1 with probability 0.5 . Also it was assumed that both T_1 and T_2 had the marginal hazard $\exp(\beta X)$ and the censoring time C followed the exponential distribution with the hazard $\exp(\gamma X)$. For estimation of $\Lambda_{0j}(t)$, we used the sieve method described in Sect. 4 with $k = [n^{1/3}] + 1$ and chose jump time points to contain approximately equal number of observations within each interval. In all the stimulations, the proposed methods were applied based on the Clayton model introduced in Sect. 1.

Note that for different copula models, the spaces of the association parameter α are quite different and also their interpretation differs from case to case. For this, in the simulation, we focused on another global association parameter, Kendall's τ , defined as

$$\tau = Pr\{(T_{1i} - T_{1j})(T_{2i} - T_{2j}) > 0\} - Pr\{(T_{1i} - T_{1j})(T_{2i} - T_{2j}) < 0\}$$

for i.i.d. replicates (T_{1i}, T_{2i}) and (T_{1j}, T_{2j}) of (T_1, T_2) (Hougaard 2000). Kendall's τ is always between -1 and 1 with 0 indicating no association, so it is easy to interpret. In addition, Kendall's τ is more robust and invariant to monotone transformation. For these reasons, we present the estimation of τ instead of α in the following.

For the results given below, we use M1, M2, and M3 to denote the proposed methods in the situations of the same regression parameter and the same baseline, the same regression parameter but different baselines, and both different, respectively, for T_1 and T_2 , and M4 the univariate regression method. For each setup considered, we recorded the bias calculated as the average of point estimates minus the true value (Bias), the sample standard deviation (SSE), the average of estimated standard errors (SEE), and the 95% empirical coverage probability (CP) based on 1,000 replicates.

In the first simulation, we considered the general performance of the proposed method by comparing it to the full maximum likelihood approach. For this, data were generated with sample size $n = 100$ or 200 , $\beta = 0$ or 1 and $\tau = 1/3$ or $1/2$. Table 1 shows the comparison results about the estimation of regression parameter β and Kendall's τ obtained by applying M1 and by maximizing the full likelihood assuming the common baseline being a step function with the same number of jump joints, respectively, in the case of covariate-independent censoring with $\gamma = 0$. These results suggest that the proposed method works as well as the full maximum likelihood method for practical sample sizes. In particular, they gave similar biases, and SSE and CP were close to SEE and 0.95 , respectively, for both methods. As expected, the performance of the proposed method as well as that of the maximum likelihood method is improved as sample size n increases, yielding smaller biases and variance estimates.

Note that the proposed methods give efficient estimates for both regression parameters and association parameter. In some situations, one may be only interested in regression parameters and in this case, an alternative approach is to apply the univariate regression analysis, which could be in general simpler and more robust. However, the univariate analysis can be less efficient in addition not to allow one to estimate the association parameters. To investigate this, we conducted the second simulation study and Table 2 shows the comparison results of the proposed methods and the univariate regression under the same setups as in the first simulation study. From the results in Table 2, M3 is quite comparable to M4, while M1 and M2 show much better results than M4 with smaller biases and smaller variances. Results in Table 2 suggest that the proposed method (M1 or M2) should be used instead of the univariate regression method (M4) when the correlated failure times share the same parameters (either regression parameters or baselines) in their marginal distributions.

The Copula model used to characterize the structure of the joint survival functions essentially implies that the global association parameter is independent of covariates. Sometimes this may not be true and to check the robustness of the proposed method with respect to this, we applied the proposed methods to the case of covariate-dependent association in the third simulation study. Under the similar situations considered above, let $\alpha(X) = 2$ when $X = 0$ and $\alpha(X) = 3$ when $X = 1$ corresponding to $\tau = 1/3$ in the control group

and 1/2 in the treatment group, respectively. Data were generated from this adjusted Clayton model with α replaced by $\alpha(X)$. Although the method seems to give little larger biases compared to the results in Table 2, the results in Table 3 (with sample size $n = 100$) suggest similar conclusions to those drawn from Table 2 in comparing the proposed methods and the univariate regression method, and in particular, the proposed methods seem to be robust. The bias becomes smaller when the sample size increases.

In the fourth simulation study, we studied the robustness of the proposed approach when there is model misspecification. To do this, we generated data from different copula models other than Clayton model but based the proposed methods on the Clayton model. It can be seen from Table 4, giving several commonly used copula models, that different copula models can have different ranges of association. In this study, we generated data from AMH, GPM, and Gumbel models with $\tau = 1/6$, and from A12, A14, and Gumbel with $\tau = 1/2$ and $2/3$, respectively. These Copula models were chosen because they have explicit and easy forms of τ as functions of the association parameter α (Genest and Rivest 1993; Huard et al. 2006). Table 5 gives the estimation results obtained by applying M1 when the true $\beta = 1$ and $\gamma = 0$. The obtained variance estimates in Table 5 seem to be biased in some cases. This bias decreased as we increased sample size in the same setup. Although the results in Table 5 are not as good as those given in Table 1, they seem to be reasonable.

The simulation results given above were for the case of covariate-independent censoring with $\gamma = 0$. We also considered the case of covariate-dependent censoring with $\gamma = 1$ and obtained similar results.

6 An application

In this section we apply our proposed method to a tumor dataset arising from a 2-year tumorigenicity study conducted by National Toxicology Program. In the study, groups of 50 male and female F344/N rats and B6C3F₁ mice were exposed to chloroprene at concentrations 0, 12.8, 32, or 80 ppm by inhalation, 6 h per day, 5 days per week, for 2 years. Each animal was examined for various tumors in different organs when died. Some animals died naturally during the study, and those who survived at the end of study were sacrificed for examinations. Here, we focus on two types of lung tumors, Alveolar/Bronchiolar Carcinoma and Alveolar/Bronchiolar Adenoma, for male B6C3F₁ mice in the control group (0 ppm) and high dose group (80 ppm). Alveolar/Bronchiolar Carcinoma is a more severe tumor and may arise from Alveolar/Bronchiolar Adenoma, but it does not regress. In the following, we use A/B C and A/B A to denote these two types of tumors for convenience. Since the tumor status was only examined at death time, we only have bivariate current status data and the common censoring (observation) time is death time.

Let T_1 and T_2 denote the ages of onset of A/B C and A/B A tumors, respectively, and C the death time ranging from 16 months to 25 months, the end time of the study. Let X denote the group indicator with 1 for the animals in the 80 ppm dose group and 0 otherwise. In both groups, there were many mice that survived at the end of the study and were sacrificed for tumor examinations. Thus, we have right censored data available for the death time C . For the analysis, we first considered the possible group effect on the death time and for this, the

partial likelihood approach gave $\hat{\gamma} = 0.7588$ with the estimated standard error of 0.2676. This result suggests that the animals in the 80 mmp dose group had a higher death rate than those in the control group.

Next we need to check if T_1 and T_2 have the same baseline cumulative hazard functions and to this end, the nonparametric maximum likelihood estimates (NPMLE's) of the survival functions of T_1 and T_2 for the animals in the two groups were obtained separately. The NPMLE's show great difference between the control and dose group for both tumors but does not show much difference between the two tumors in either group. The plot of NPMLE's suggests that the marginal proportional hazard assumption is valid for both T_1 and T_2 . In the following, we apply the proposed methods M1, M2 and M3 as mentioned in Sect. 5. The baseline hazard functions were taken to be step functions with jumps at 18, 20, 22, 24, and 25 months, and the joint survival function was assumed to be specified by a Clayton type model allowing negative association. Table 6 lists the results obtained by the proposed methods as well as the univariate regression method, including point estimates, estimated standard errors, and p -values. The p -values were calculated by using the Wald's test under $H_0 : \beta = 0$, corresponding to no dose effect in the high dose group, and $H_0 : \tau = 0$, corresponding to no association between the onset times to the two lung tumors.

The results in Table 6 show a significant positive dose effect and suggest that the animals in the dose group had a significantly higher chance of getting these two lung tumors than those in the control group. The proposed methods gave a negative association between the two lung tumors. All the test results suggest significant negative association between the two tumors based on the reported p -values. This negative association makes sense because the numbers of these two tumors go to different directions as some A/B A tumors become more serious and transit to A/B C tumors.

Note that in the above analysis, we assumed that the association parameter is the same for different covariate values or independent of the covariate. As a referee pointed out, one way for checking this, which can also be used as a simpler alternative to the proposed method, is to fit the Copula model to each covariate group separately to estimate the association parameter since the covariate only takes two values. The application of this alternative produced $\hat{\tau} = 0.0233$ with the standard error 0.5043 for the control group and $\hat{\tau} = -0.3317$ with the standard error 0.1192 for the high dose group. These results suggest that the onset times of the two types of tumors seem to be independence for the animals in the control group but significantly negatively correlated for the animals in the high dose group. One reason for the independence or insignificant association in the control group may be the rare cases of left censoring (corresponding to the cases that tumors had occurred) in the group.

7 Concluding remarks

In this paper, we considered the efficient estimation of regression and association parameters for bivariate current status data for situations where the joint survival function of the two related survival times follows a copula model. We derived the efficient score and information for these parameters and the simulation results suggest that the proposed estimates are efficient and also seem to be robust against misspecified Copula models for

practical situations. In particular, in the case where the two failure times are highly correlated, the presented approach could give more efficient estimates of regression parameters than the use of univariate regression approaches. This is especially the case when the two failure times have the same marginal survival functions.

To implement the proposed procedure, a key step is to find uniformly consistent estimates of baseline cumulative hazard functions with $n^{1/3}$ convergence rate. For this, we proposed to use the sieve method to approximate the baseline cumulative hazard function. As other authors have commented before (Rossini and Tsiatis 1996), in the sieve method, the choice of the number of sieves could affect the results. If the number is too small, the step function can not fully represent the variability of the data, while if the number is too large, there may not be enough information to estimate all parameters. We also observed this in the simulation study and it does not seem to exist approaches that provide the optimal number of sieves.

In this paper, we have assumed Cox model for the censoring time C for the ease of computation since in this case the estimation procedure does not need estimating the baseline hazard function of C . This assumption is not essential in the proposed methods and can be relaxed to any flexible model allowing C to be continuous or discrete as commonly seen in real-life examples. Since we have complete or right censored data available for the censoring time C in current status data, model checking on C is relatively easy. The proposed methods still apply when using a general model on C but may need much more computational effort.

In the proposed approach, it was assumed that the association is independent of covariates and as seen in the example, this may not be true in practice. Estimation of association is challenging in general when the association is covariate-dependent, let alone considering efficient estimation. In the special case that the covariates are all discrete, one could fit the Copula model to each covariate group to estimate the association as long as there are enough samples in each covariate group. The last part in Sect. 6 provides an illustrative example of such a method. Effort is clearly needed to study the general case. Other future research interests related to this paper include checking whether an assumed Copula model is valid and whether the association is equal for all the subjects.

Acknowledgments

The authors wish to thank the Editor-in-Chief, an associate editor, and two referees for their critical and constructive comments that greatly improved the paper. Dr. Wang's research is supported by the Intramural Research Program of the NIH, NIEHS, Dr. Sun's work is partly supported by a NIH grant, and Dr. Tong's work is partly supported by NSFC 11671106.

References

- Anderson PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat.* 1982; 10:1100–1120.
- Bickel, P.; Klaassen, C.; Ritov, Y.; Wellner, J. Efficient and adaptive estimation for semiparametric models. The Johns Hopkins University Press; Baltimore: 1993.
- Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika.* 1978; 65:141–151.

- Cox DR. Regression models and life-table (with discussion). *J Roy Stat Soc B*. 1972; 33:187–220.
- Ding AA, Wang W. Testing independence for bivariate current status data. *J Am Stat Assoc*. 2004; 99:145–155.
- Dunson D, Dinse G. Bayesian models for multivariate current status data with informative censoring. *Biometrics*. 2002; 58:79–88. [PubMed: 11890330]
- Foutz RV. On the unique consistent solution to the likelihood equations. *J Am Stat Assoc*. 1977; 72:147–148.
- Genest C, MacKay J. The joy of copulas: bivariate distributions with uniform marginals. *Am Stat*. 1986; 4:280–283.
- Genest C, Rivest L-P. Statistical inference procedures for bivariate Archimedean copulas. *J Am Stat Assoc*. 1993; 88:1034–1043.
- Goggins WB, Finkelstein DM. A proportional hazards model for multivariate interval-censored failure time data. *Biometrics*. 2000; 56:940–943. [PubMed: 10985240]
- Groeneboom, P.; Wellner, JA. Information bounds and nonparametric maximum likelihood estimation. Birkhäuser; Boston: 1992.
- Hougaard, P. Analysis of multivariate survival data. Springer; 2000.
- Huang J. Efficient estimation for the proportional hazards model with interval censoring. *Ann Stat*. 1996; 24:540–568.
- Huang J, Rossini JA. Sieve estimation for the proportional odds model with interval-censoring. *J Am Stat Assoc*. 1997; 92:960–967.
- Huard D, Evin G, Favre A-C. Bayesian Copula selection. *Comput Stat Data Anal*. 2006; 51:809–822.
- Jewell NP, van de Laan M, Lei X. Bivariate current status data with univariate monitoring times. *Biometrika*. 2005; 92:847–862.
- Keiding N. Age-specific incidence and prevalence: a statistical perspective. *J Roy Stat Soc A*. 1991; 154:371–412.
- Kim MY, Xue X. The analysis of multivariate interval-censored survival data. *Stat Med*. 2002; 21:3715–3726. [PubMed: 12436466]
- Lin DY, Oakes D, Ying Z. Additive hazards regression with current status data. *Biometrika*. 1998; 85:289–298.
- Martinussen T, Scheike TH. Efficient estimation in additive hazards regression with current status data. *Biometrika*. 2002; 89:649–658.
- Rossini A, Tsiatis AA. A semiparametric proportional odds regression model for the analysis of current status data. *J Am Stat Assoc*. 1996; 91:713–721.
- Shih JH, Louis TA. Inference on the association parameter in copula models for bivariate survival data. *Biometrics*. 1995; 51:1384–1399. [PubMed: 8589230]
- Sun, J. The statistical analysis of interval-censored data. Springer; 2006.
- Sun L, Wang L, Sun J. Estimation of the association for bivariate interval-censored failure time data. *Scand J Stat*. 2006; 33:637–649.
- Wang W, Ding AA. On assessing the association for bivariate current status data. *Biometrika*. 2000; 87:879–893.
- Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J Am Stat Assoc*. 1989; 84:1065–1073.

Appendix

In this appendix, we will show the consistency and asymptotic normality of the proposed estimate $\hat{\theta}$. Also it will be shown that it is asymptotically efficient in the sense that the variance of $\hat{\theta}$ reaches the information lower bound $I^{-1}(\theta_0)$, where θ_0 denotes the true value of θ . For the proof, we only consider the case of different marginal survival functions and it is simpler and similar for the case of common marginal survival function.

Let γ_0 be the true vector of γ and $\hat{Z}_{jmi}$, $\hat{a}_{jmi}$, $\hat{S}_{jmi}$, $\hat{H}_n$, and $\hat{G}_n$ the plug-in estimates of Z_{jmi} , a_{jmi} , S_{jmi} , H_n and G_n with Λ_{01} , Λ_{02} and γ replaced by $\hat{\Lambda}_{01}$, $\hat{\Lambda}_{02}$ and $\hat{\gamma}$, respectively. Define

$$U_n(\theta, \Lambda_{01}, \Lambda_{02}, \gamma_0) = \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jmi} - H_n G_n^{-1}) a_{jmi}}{S_{jmi}} dN_{jmi}(t).$$

By definition, $\hat{\theta}$ is the solution to the equation $U_n(\hat{\theta}, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma}) = 0$, where $\hat{\Lambda}_{01}$ and $\hat{\Lambda}_{02}$ are the estimates of Λ_{01} and Λ_{02} both with $n^{1/3}$ convergence rate and $\hat{\gamma}$ is an estimate of γ with $n^{1/2}$ convergence rate. We assume that the $I(\theta_0)$ defined in (4) is positive definite.

To establish the asymptotic properties of $\hat{\theta}$, first we will prove that as $n \rightarrow \infty$,

$$n^{-1/2} U_n(\theta_0, \Lambda_{01}, \Lambda_{02}, \gamma_0) \xrightarrow{D} N(0, I(\theta_0)), \quad (\text{A.1})$$

and

$$n^{-1/2} U_n(\theta_0, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma}) = n^{-1/2} U_n(\theta_0, \Lambda_{01}, \Lambda_{02}, \gamma_0) + o_p(1). \quad (\text{A.2})$$

For (A.1), using the facts that $\sum_{j=0}^1 \sum_{m=0}^1 Z_{jmi} a_{jmi} = 0$ and $\sum_{j=0}^1 \sum_{m=0}^1 a_{jmi} = 0$, we can rewrite $U_n(\theta_0, \Lambda_{01}, \Lambda_{02}, \gamma_0)$ as

$$\begin{aligned} U_n(\theta_0, \Lambda_{01}, \Lambda_{02}, \gamma_0) &= \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jmi} - H_n G_n^{-1}) a_{jmi}}{S_{jmi}} dM_{jmi}(t) + \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} (Z_{jmi} - H_n G_n^{-1}) a_{jmi} Y_i \lambda_{ci} dt \\ &= \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jmi} - H_n G_n^{-1}) a_{jmi}}{S_{jmi}} dM_{jmi}(t) \\ &= \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jmi} - H G^{-1}) a_{jmi}}{S_{jmi}} dM_{jmi}(t) + o_p(n^{1/2}). \end{aligned}$$

The last equality holds since H_n and G_n are the empirical versions of H and G , respectively. That above result suggests that $U_n(\theta_0, \Lambda_{01}, \Lambda_{02}, \gamma_0)$ can be approximated by a sum of i.i.d. martingales. Thus (A.1) follows from the martingale central limit theorem.

For (A.2), note that

$$\begin{aligned} n^{-1/2} U(\theta_0, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma}) &= n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(\hat{Z}_{jmi} - \hat{H}_n \hat{G}_n^{-1}) \hat{a}_{jmi}}{\hat{S}_{jmi}} dM_{jmi}(t) + n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(\hat{Z}_{jmi} - \hat{H}_n \hat{G}_n^{-1}) \hat{a}_{jmi}}{\hat{S}_{jmi}} \times S_{jmi} Y_i \lambda_{ci} dt \\ &= V_{1n} + V_{2n}. \end{aligned}$$

To show (A.2), we need to show that V_{1n} is asymptotically equivalent to the quantity at the right side of (A.2) and V_{2n} converges in probability to zero. Define

$$Q_{jmi} = Q_{jmi}(\Lambda_{01}, \Lambda_{02}, \gamma_0) = (Z_{jmi} - H_n G_n^{-1}) a_{jmi} / S_{jmi}$$

and denote $Q_{jmi}^{(1)}$, $Q_{jmi}^{(2)}$ and $Q_{jmi}^{(3)}$ the derivatives of $Q_{jmi}(u, v, \gamma)$ with respect to u , v and γ , respectively, evaluated at $u = \Lambda_{01}$, $v = \Lambda_{02}$ and $\gamma = \gamma_0$. Denoted by $\hat{Q}_{jmi}$, $\hat{Q}_{jmi}^{(1)}$, $\hat{Q}_{jmi}^{(2)}$ and $\hat{Q}_{jmi}^{(3)}$ the corresponding estimates with Λ_{01} , Λ_{02} and γ_0 replaced by $\hat{\Lambda}_{01}$, $\hat{\Lambda}_{02}$ and $\hat{\gamma}$, respectively.

Using Taylor series expansion, one obtain

$$\begin{aligned} V_{1n} = & n^{-1/2} U_n(\theta_0, \Lambda_{01}, \Lambda_{02}, \gamma_0) + n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} Q_{jmi}^{(3)}(\hat{\gamma} - \gamma_0) dM_{jmi} \\ & + n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} [Q_{jmi}^{(1)}(\hat{\Lambda}_{01} - \Lambda_{01}) + Q_{jmi}^{(2)}(\hat{\Lambda}_{02} - \Lambda_{02})] dM_{jmi} + o_p(1) \end{aligned}$$

based on the fact that $|\hat{\Lambda}_{01} - \Lambda_{01}|^2 + |\hat{\Lambda}_{02} - \Lambda_{02}|^2 + \|\hat{\gamma} - \gamma_0\|^2 = O(n^{-2/3})$. This fact plus the boundedness of $(Q_{jmi}^{(k)}, k = 1, 2, 3)$ and martingale properties gives us that the second and third terms converge to zero in probability.

Now consider V_{2n} . First notice that

$$\sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \hat{Q}_{jmi} \hat{S}_{jmi} Y_i \hat{\lambda}_{ci} dt = \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} (\hat{Z}_{jmi} - \hat{H}_n \hat{G}_n^{-1}) \hat{a}_{jmi} Y_i \hat{\lambda}_{ci} dt = 0$$

based on the fact that $\sum_{j=0}^1 \sum_{m=0}^1 \hat{Z}_{jmi} \hat{a}_{jmi} = 0$ and $\sum_{j=0}^1 \sum_{m=0}^1 \hat{a}_{jmi} = 0$ for any i . Then we obtain

$$\begin{aligned} V_{2n} = & -n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \hat{Q}_{jmi} Y_i (\hat{S}_{jmi} \hat{\lambda}_{ci} - S_{jmi} \lambda_{ci}) dt \\ = & -n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \hat{Q}_{jmi} Y_i (\hat{S}_{jmi} - S_{jmi}) \hat{\lambda}_{ci} dt - n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \hat{Q}_{jmi} Y_i S_{jmi} (\hat{\lambda}_{ci} - \lambda_{ci}) dt \\ = & -n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} Q_{jmi} (\hat{S}_{jmi} - S_{jmi}) Y_i \lambda_{ci} dt - n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} Q_{jmi} S_{jmi} Y_i (\hat{\lambda}_{ci} - \lambda_{ci}) dt + O_p(n^{-1/6}). \end{aligned}$$

The last step uses the facts that $\hat{Q}_{jmi} - Q_{jmi} = O_p(n^{-1/3})$, $\hat{S}_{jmi} - S_{jmi} = O_p(n^{-1/3})$ and $\hat{\lambda}_{ci} - \lambda_{ci} = O_p(n^{-1/2})$. Denote the first and second terms in the above by W_{1n} and W_{2n} . By using Taylor series expansion of $\hat{S}_{jmi}$ at true value $(\Lambda_{01}, \Lambda_{02})$, we have

$$W_{1n} = -n^{-1/2} \sum_{k=1}^2 \int_0^{M_0} \Lambda_{0k}^{-1} (\Lambda_{0k} - \hat{\Lambda}_{0k}) \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 Q_{jmi} a_{jmi}^{(k)} Y_i \lambda_{ci} dt + O_p(n^{-1/6}).$$

The first term in W_{1n} reduces to zero based on the fact $\sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 Q_{jmi} a_{jmi}^{(k)} Y_i \lambda_{ci} = 0$ due to the empirical efficient score or the construction of H_n and G_n . It is easy to see that

$$W_{2n} = -n^{-1/2} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} (Z_{jmi} - H_n G_n^{-1}) a_{jmi} Y_i (\hat{\lambda}_{ci} - \lambda_{ci}) (t) dt = 0$$

based on the facts that $\sum_{j=0}^1 \sum_{m=0}^1 Z_{jmi} a_{jmi} = 0$ and $\sum_{j=0}^1 \sum_{m=0}^1 a_{jmi} = 0$ for any i . Thus, we show that V_{2n} converges to zero in probability and then complete the proof of (A.2).

In addition to (A.1) and (A.2), we also need to establish that $n^{-1} U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma})|_{\theta=\theta_0}$ converges in probability to $-I(\theta_0)$. This is equivalent to proving that

$$n^{-1} \frac{\partial}{\partial \theta} U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma})|_{\theta=\theta_0} = n^{-1} \frac{\partial}{\partial \theta} U_n(\theta, \Lambda_{01}, \Lambda_{02}, \gamma_0)|_{\theta=\theta_0} + o_p(1) \quad (\text{A.3})$$

and

$$n^{-1} \frac{\partial}{\partial \theta} U_n(\theta, \Lambda_{01}, \Lambda_{02}, \gamma_0)|_{\theta=\theta_0} \xrightarrow{p} -I(\theta_0). \quad (\text{A.4})$$

The equality (A.3) can be proved similarly as (A.2). For (A.4), note the fact that

$$\sum_{j=0}^1 \sum_{m=0}^1 \frac{\partial}{\partial \theta} (Z_{jmi} - H_n G_n^{-1}) a_{jmi} = 0$$

following from $\sum_{j=0}^1 \sum_{m=0}^1 Z_{jmi} a_{jmi} = 0$ and $\sum_{j=0}^1 \sum_{m=0}^1 a_{jmi} = 0$ for each i and from the fact that both H_n and G_n are free of j , m and i . Hence we have

$$\begin{aligned} & n^{-1} \frac{\partial}{\partial \theta} U_n(\theta, \Lambda_{01}, \Lambda_{02}, \gamma_0)|_{\theta=\theta_0} \\ &= -\frac{1}{n} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jmi} - H_n G_n^{-1}) a_{jmi}}{S_{jmi}^2} (Z_{jmi} a_{jmi})' dN_{jmi}(t) \\ &+ \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \left[\frac{\partial}{\partial \theta} (Z_{jmi} - H_n G_n^{-1}) a_{jmi} \right] S_{jmi}^{-1} dN_{jmi}(t) \\ &- \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} \frac{(Z_{jmi} - H_n G_n^{-1}) a_{jmi}}{S_{jmi}} (Z_{jmi} a_{jmi})' Y_i \lambda_{ci} dt + o_p(1) \\ &= -\frac{1}{n} \sum_{i=1}^n \sum_{j=0}^1 \sum_{m=0}^1 \int_0^{M_0} ((Z_{jmi} - H_n G_n^{-1}) a_{jmi})^{\otimes 2} S_{jmi}^{-1} Y_i \lambda_{ci} dt + o_p(1), \end{aligned}$$

which proves (A.4).

Now for the asymptotic properties of $\hat{\theta}$, note that for any θ in a neighborhood of θ_0 , one can show that $n^{-1}U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma})$ is asymptotically equivalent to $n^{-1}U_n(\theta, \Lambda_{01}, \Lambda_{02}, \gamma_0)$ using the arguments similar to those for (A.2). Define

$$V(\theta) = \sum_{j=0}^1 \sum_{m=0}^1 E \int_0^{M_0} \frac{(Z_{jm} - HG^{-1})a_{jm}}{S_{jm}}(\theta, \Lambda_{01}, \Lambda_{02}, \gamma_0) S_{jm}(\theta, \Lambda_{01}, \Lambda_{02}) Y \lambda_c dt.$$

Then it can be easily seen that $U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma}) = V(\theta_0) = 0$ and $n^{-1}U_n(\theta, \hat{\Lambda}_{01}, \hat{\Lambda}_{02}, \hat{\gamma})$ converges almost surely to $V(\theta)$. These along with (A.3), (A.4) and the inverse function theorem (Foutz 1977) show that $\hat{\theta}$ is a consistent estimate of θ_0 . The asymptotic normality and efficiency of $\hat{\theta}$ follow directly from the Taylor series expansion and (A.1)–(A.4).

Table 1

Comparison of the proposed method (M1) and maximum likelihood method

True	Maximum likelihood method				Proposed method			
	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
$n = 100$								
0	$\hat{\beta}$ -0.0004	0.2632	0.2609	0.939	-0.0036	0.2709	0.2629	0.939
1/3	$\hat{\tau}$ -0.0152	0.1199	0.1129	0.930	-0.0180	0.1198	0.1148	0.937
1	$\hat{\beta}$ 0.0693	0.3241	0.2988	0.948	0.0663	0.3256	0.3068	0.959
1/3	$\hat{\tau}$ -0.0111	0.1185	0.1189	0.939	-0.0147	0.1183	0.1211	0.949
0	$\hat{\beta}$ 0.0034	0.2731	0.2761	0.955	-0.0033	0.2805	0.2775	0.953
1/2	$\hat{\tau}$ -0.0100	0.1058	0.1055	0.947	-0.0126	0.1062	0.1066	0.946
1	$\hat{\beta}$ 0.0754	0.3390	0.3180	0.948	0.0709	0.3490	0.3254	0.950
1/2	$\hat{\tau}$ -0.0032	0.1104	0.1123	0.939	-0.0061	0.1110	0.1135	0.939
$n = 200$								
0	$\hat{\beta}$ -0.0062	0.1710	0.1817	0.967	-0.0058	0.1735	0.1814	0.966
1/3	$\hat{\tau}$ -0.0051	0.0809	0.0790	0.953	-0.0070	0.0809	0.0795	0.952
1	$\hat{\beta}$ 0.0312	0.2075	0.2036	0.957	0.0278	0.2103	0.2063	0.950
1/3	$\hat{\tau}$ -0.0029	0.0853	0.0830	0.937	-0.0052	0.0854	0.0836	0.939
0	$\hat{\beta}$ -0.0092	0.1831	0.1919	0.958	-0.0098	0.1876	0.1914	0.955
1/2	$\hat{\tau}$ 0.0008	0.0694	0.0736	0.957	-0.0009	0.0694	0.0739	0.961
1	$\hat{\beta}$ 0.0377	0.2213	0.2162	0.958	0.0352	0.2217	0.2198	0.948
1/2	$\hat{\tau}$ 0.0030	0.0826	0.0785	0.929	0.0012	0.0829	0.0787	0.932

Table 2

Comparison of the proposed methods and the univariate regression method

(β, ϑ)	$n = 100$				$n = 200$			
	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
(0, 1/3)								
M1 $\hat{\beta}$	0.0091	0.2787	0.2636	0.938	0.0029	0.1905	0.1817	0.943
M2 $\hat{\beta}$	-0.0118	0.2886	0.2664	0.935	-0.0077	0.1941	0.1827	0.937
M3 $\hat{\beta}_1$	0.0065	0.3535	0.3333	0.949	-0.0009	0.2395	0.2255	0.948
$\hat{\beta}_2$	0.0107	0.3626	0.3356	0.940	0.0074	0.2336	0.2254	0.951
M4 $\hat{\beta}_1$	0.0070	0.3587	0.3350	0.944	-0.0012	0.2419	0.2300	0.936
$\hat{\beta}_2$	0.0128	0.3612	0.3367	0.935	0.0079	0.2362	0.2300	0.949
(1, 1/3)								
M1 $\hat{\beta}$	0.0446	0.3170	0.3037	0.960	0.0206	0.2101	0.2056	0.949
M2 $\hat{\beta}$	0.0388	0.3255	0.3036	0.956	0.0198	0.2126	0.2058	0.951
M3 $\hat{\beta}_1$	0.0936	0.4090	0.3980	0.965	0.0478	0.2639	0.2626	0.957
$\hat{\beta}_2$	0.0910	0.4078	0.3984	0.966	0.0467	0.2635	0.2596	0.961
M4 $\hat{\beta}_1$	0.0875	0.4109	0.3841	0.961	0.0448	0.2673	0.2590	0.954
$\hat{\beta}_2$	0.0876	0.4115	0.3847	0.964	0.0439	0.2681	0.2572	0.949
(0, 1/2)								
M1 $\hat{\beta}$	0.0073	0.2865	0.2783	0.943	-0.0009	0.2003	0.1915	0.940
M2 $\hat{\beta}$	-0.0185	0.2957	0.2797	0.942	-0.0138	0.2033	0.1923	0.939
M3 $\hat{\beta}_1$	0.0124	0.3440	0.3301	0.947	-0.0041	0.2293	0.2227	0.947
$\hat{\beta}_2$	0.0069	0.3497	0.3322	0.948	0.0028	0.2327	0.2236	0.942
M4 $\hat{\beta}_1$	0.0109	0.3523	0.3330	0.945	-0.0025	0.2372	0.2295	0.940
$\hat{\beta}_2$	0.0125	0.3516	0.3345	0.948	0.0054	0.2414	0.2299	0.943
(1, 1/2)								
M1 $\hat{\beta}$	0.0575	0.3289	0.3213	0.957	0.0277	0.2210	0.2187	0.953
M2 $\hat{\beta}$	0.0441	0.3368	0.3198	0.949	0.0193	0.2234	0.2175	0.953

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

(β, σ)	$n = 100$						$n = 200$					
	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP	Bias	SSE	SEE	CP
M3	$\hat{\beta}_1$	0.0957	0.3966	0.3927	0.973	0.0478	0.2603	0.2604	0.951			
	$\hat{\beta}_2$	0.0974	0.3901	0.3925	0.962	0.0497	0.2551	0.2591	0.968			
M4	$\hat{\beta}_1$	0.0857	0.4011	0.3810	0.959	0.0458	0.2658	0.2590	0.953			
	$\hat{\beta}_2$	0.0905	0.3915	0.3812	0.962	0.0481	0.2615	0.2580	0.964			

Table 3

Simulation results with sample size $n = 100$ when the association parameter is covariate dependent

		$\beta_1 = \beta_2 = 0$						$\beta_1 = \beta_2 = 1$					
		Bias		SSE	SEE	CP		Bias		SSE	SEE	CP	
M1	$\hat{\beta}$	0.0545	0.2798	0.2705	0.947	0.0897	0.3300	0.3151	0.945				
M2	$\hat{\beta}$	0.0186	0.2878	0.2712	0.939	0.0705	0.3392	0.3105	0.934				
M3	$\hat{\beta}_1$	0.0499	0.3504	0.3302	0.943	0.1323	0.4145	0.3993	0.950				
	$\hat{\beta}_2$	0.0432	0.3412	0.3312	0.958	0.1319	0.4117	0.3984	0.958				
M4	$\hat{\beta}_1$	0.0082	0.3611	0.3343	0.940	0.0821	0.4164	0.3849	0.942				
	$\hat{\beta}_2$	0.0025	0.3477	0.3348	0.953	0.0798	0.4141	0.3854	0.951				

Table 4

Copulae used in the simulation studies

Copulas	$C_a(u, v)$	Ω_a	$\tau = \tau(a)$	Ω_τ
Clayton	$[u^{-(a-1)} + v^{-(a-1)} - 1]^{-1/(a-1)}$	$(1, \infty)$	$\tau = \frac{\alpha-1}{\alpha+1}$	$(0, 1]$
AMH	$\frac{uv}{1-\alpha(1-u)(1-v)}$	$(-1, 1)$	$1 - g(a)^a$	$(\eta, \frac{1}{3})_a$
FGM	$uv + auv(1-u)(1-v)$	$[-1, 1]$	$\frac{2\alpha}{9}$	$[-\frac{2}{9}, \frac{2}{9}]$
A12	$\{1 + [(u^{-1} - 1)^a + (v^{-1} - 1)^a]^{1/a}\}^{-1}$	$[1, \infty)$	$1 - \frac{2}{3\alpha}$	$[\frac{1}{3}, 1)$
A14	$\{1 + [(u^{-1/a} - 1)^a + (v^{-1/a} - 1)^a]^{1/a}\}^{-a}$	$[1, \infty)$	$1 - \frac{2}{1+2\alpha}$	$[\frac{1}{3}, 1)$
Gumbel	$\exp\{-[(-\ln u)^a + (-\ln v)^a]^{1/a}\}$	$[1, \infty)$	$1 - a^{-1}$	$[0, 1)$

^aNote $g(\alpha) = \frac{2}{3} \frac{\alpha^2 \ln(1-\alpha) - 2\alpha \ln(1-\alpha) + \alpha + \ln(1-\alpha)}{\alpha^2}$ and $\eta = \frac{5}{3} - \frac{8}{3} \ln 2 \approx -0.1817$

Table 5
Simulation results based on Clayton model when data are generated from other Copula models

True t	Copula	Estimate	Bias	SSE	SEE	CP
$t = 1/6$	AMH	$\hat{\beta}$	0.0180	0.2977	0.2830	0.943
		$\hat{\tau}$	-0.0034	0.0987	0.1265	0.984
	GPM	$\hat{\beta}$	0.0283	0.2807	0.2809	0.953
		$\hat{\tau}$	-0.0167	0.0964	0.1266	0.989
$t = 1/2$	Gumbel	$\hat{\beta}$	0.0330	0.2947	0.2778	0.947
		$\hat{\tau}$	-0.0180	0.0956	0.1272	0.988
	A12	$\hat{\beta}$	0.0561	0.3481	0.3178	0.949
		$\hat{\tau}$	-0.0268	0.1190	0.1142	0.927
$t = 2/3$	A14	$\hat{\beta}$	0.0444	0.3362	0.3105	0.941
		$\hat{\tau}$	-0.0371	0.1181	0.1140	0.937
	Gumbel	$\hat{\beta}$	0.0149	0.3360	0.2973	0.932
		$\hat{\tau}$	-0.0694	0.1189	0.1167	0.925
$t = 2$	A12	$\hat{\beta}$	0.0397	0.3628	0.3291	0.942
		$\hat{\tau}$	-0.0265	0.1048	0.0995	0.949
	A14	$\hat{\beta}$	0.0197	0.3480	0.3227	0.949
		$\hat{\tau}$	-0.0398	0.1005	0.1005	0.941
	Gumbel	$\hat{\beta}$	0.0213	0.3633	0.3188	0.934
		$\hat{\tau}$	-0.0534	0.1020	0.1017	0.943

Table 6

Results of lung tumor data analysis

Model		Estimate	Standard error	<i>p</i> -Value
M1	$\hat{\beta}$	1.9272	0.2822	0.0000
	$\hat{\tau}$	-0.3027	0.1221	0.0132
M2	$\hat{\beta}$	1.9068	0.2825	0.0000
	$\hat{\tau}$	-0.3099	0.1213	0.0106
M3	$\hat{\beta}_1$	2.1972	0.4804	0.0000
	$\hat{\beta}_2$	1.6935	0.4171	0.0000
	$\hat{\tau}$	-0.2998	0.1210	0.0132
M4	$\hat{\beta}_1$	2.1647	0.4748	0.0000
	$\hat{\beta}_2$	1.6877	0.4189	0.0001