Inference on Association Measure for Bivariate Survival Data with Hybrid Censoring

By SUHONG ZHANG, YING ZHANG, KATHRYN CHALONER

Department of Biostatistics, The University of Iowa,

C22 GH, 200 Hawkins Drive, Iowa City, IA 52242, U.S.A.

suhong-zhang@uiowa.edu ying-j-zhang@uiowa.edu kathryn-chaloner@uiowa.edu

AND JACK T. STAPLETON

Department of Internal Medicine, The University of Iowa and Iowa City VA Medical Center, SW54-15 GH, 200 Hawkins Drive, Iowa City, IA, 52242 U.S.A.

jack-stapleton@uiowa.edu

Summary

A two-stage semiparametric estimator is proposed to estimate the association measure for bivariate survival data that are subject to hybrid censoring: one event time is right censored and the other is observed as current status data, or subject to interval censoring case 1. The bivariate data are assumed to follow a copula model, in which the association parameter is of primary interest. The consistency and asymptotic normality of the proposed estimator are established based on empirical process theories. Simulation studies indicate that the estimator performs quite well with a moderate sample size. The method is applied to a motivating HIV example, which studies the effect of GB virus type C (GBV-C) co-infection on the survival of HIV infected individuals.

Some key words: Association measure; Bivariate survival model; Copula; Current status data; Empirical process; GBV-C; HIV; Kendall's τ; Right censored data

1. INTRODUCTION

Event times are often subject to various types of censoring. The most common censoring case is right censoring, which happens when the event has not occurred at the end of the study or the subject withdraws from the study. Interval censoring occurs when the event, such as the clearance of an infection, is only known to occur within an interval. A special case of interval censored data is current status data, or interval censoring case 1 data (Groeneboom & Wellner, 1992), which happens when it is only feasible to know whether the event has occurred or not by a random monitoring time C. Specifically, let T denote the event time, then one observes (C, Δ), where $\Delta = I(T \leq C)$ and $I(\cdot)$ is the indicator function.

In this manuscript, we consider a pair of positive random event times (T_1^0, T_2^0), where T_1^0 is right censored by a random time C_1, and the only observation on T_2^0 is its status: whether or not T_2^0 exceeds a random monitoring time C_2. For each individual, one observes

$$\mathbf{X} = \{(T_1, T_2, \Delta_1, \Delta_2) : T_1 = \min(T_1^0, C_1), T_2 = C_2, \Delta_1 = I(T_1^0 \leq C_1), \Delta_2 = I(T_2^0 \leq C_2)\}. \tag{1}$$

This hybrid censoring data structure is observed in a motivating example, which studies the association between the HIV survival, time from HIV seroconversion to death, and the infection of a harmless virus called GB virus type C (GBV-C).

Some prior studies suggest that GBV-C delays the progression of HIV disease (Xiang

et al., 2001; Tillmann et al., 2001; Williams et al., 2004), while a few other studies fail to find a beneficial effect (Birk et al., 2002; Björkman et al., 2004; Van der Bij et al., 2005). These studies compared the survival curves for HIV-infected subjects with or without GBV-C infection. However, the two sample comparison method used in these studies does not adjust for the duration of GBV-C infection, which may vary from subject to subject due to its self-clearance nature, and may be the potential source of contradictive results. Williams et al. (2004) conducted the most comprehensive GBV-C study to date, and found that GBV-C is associated with prolonged survival when a selected cohort from the Multicenter AIDS Cohort Study (MACS) is examined at 5-6 years after HIV seroconversion, but no association has been found when examined at 12-18 months. Longitudinal GBV-C testing on more than two time points in HIV infected individuals are not readily available from any other studies. Besides the evaluation at baseline (early measurement to select co-infected individuals), GBV-C status is only monitored once during the follow up for each individual in the MACS sub-cohort. Therefore, the GBV-C status is well fitted to the current status data structure. It is well understood that HIV survival is subject to right censoring. The lack of a proper analytical tool for this type of data motivates us to re-analyze the MACS sub-cohort data from Williams et al. (2004), by developing a new bivariate analysis method through modeling the association of HIV survival and the duration of GBV-C infection.

Bivariate and multivariate survival data have been studied extensively in statistical literatures. Liang et al. (1995) and Oakes (2000) reviewed some recent developments for analysis of multivariate failure time data. Copula based survival models are considered, for example, by Hougaard (1989), Oakes (1989), Shih & Louis (1995) and Wang & Ding (2000), to

study the association between two event times. Shih & Louis (1995) examined the association of the bivariate data that are both subject to right censoring, through a two-stage semiparametric estimation procedure. At the first stage, the marginal survival functions are estimated consistently by nonparametric maximum likelihood estimator (NPMLE). At the second stage, a dependency structure is imposed by using a copula model, and the NPMLEs of two marginal survival functions are plugged into the likelihood to form a pseudolikelihood, then the association parameter is estimated through a pseudolikelihood approach. Wang & Ding (2000) proposed a parallel two-stage semiparametric method for the bivariate current status data. Both papers show that the proposed estimators of the dependence measure converge in distribution to normal distributions with the $n^{1/2}$ rate, without showing the consistencies in the first place. In this manuscript, we model the association of bivariate event times using copula models and estimate the association parameter through the two-stage procedure as well, but we focus specifically on the data structure where one of the paired event time data is right-censored and the other is observed as current status data, as stated in (1).

The classical method for estimating the NPMLE of marginal survival function for right censored data is widely cited as the Kaplan-Meier estimator (1958). For current status data, Turnbull (1976) derived a self-consistency equation and used the EM algorithm to compute the NPMLE of distribution function. Groeneboom & Wellner (1992) introduced the Convex Minorant Algorithm (CMA) to solve for the NPMLE of distribution function as well. Huang & Wellner (1997) reviewed recent progress in interval censored survival data, which includes current status data as a special case.

Our main goal in this manuscript is to develop an inference procedure to study the association of bivariate survival data with hybrid censoring structure aforementioned. A direct application of this development is to investigate the association between HIV survival and the duration of GBV-C infection to see if they are positively correlated or not. The knowledge of this association may lead to a potential HIV treatment.

2. SOME PRELIMINARIES

2.1 *Copula models*

A copula is often referred to as the multivariate distribution function whose marginal distributions are uniform over $[0, 1]$. Consider a bivariate uniform random variable (U_1, U_2), a copula C is defined as

$$C(u_1, u_2) = Pr(U_1 \leq u_1, U_2 \leq u_2).$$

Any continuous random variable can be transformed from the uniform random variable on $[0, 1]$, therefore, copula can be used to construct a multivariate distribution with any marginal distributions.

For Bivariate distribution function H with univariate marginal distribution functions F and G, the associated copula function C is

$$C_\alpha : [0, 1]^2 \to [0, 1] \text{ that satisfies}$$

$$H_\alpha(x, y) = C_\alpha(F(x), G(y)),$$

where α is called the association parameter, see Nelsen (2006). A bivariate survival function can be defined in a similar way.

Copula model provides a convenient way to express the joint distribution of two or more random variables. A copula facilitates the joint distribution into two contributions: the marginal distributions of the individual variables, and the interdependency between margins. It's sometime useful to do so if we mainly focus on either the marginal distributions only or the interdependency only.

The association parameter α defines the strength of dependency between two margins. The Kendall's tau, denoted by τ, is related to α as below:

$$\tau = 4 \int_0^1 \int_0^1 C_\alpha(u, v) du dv - 1.$$

A collection of copulas called Archimedean copulas have been studied extensively in literature. Suppose that $\psi_\alpha : [0, \infty] \to [0, 1]$ is a strictly decreasing function such that $\psi(0)_\alpha = 1$, then an Archimedean copula can be generated as

$$C_\alpha(u, v) = \psi_\alpha(\psi_\alpha^{-1}(u) + \psi_\alpha^{-1}(v)), u, v \in [0, 1].$$

Examples of Archimedean copulas include the following three popular sub-families:

1. Gumbel (Gumbel-Hougaard) copula:

$$C_\alpha(u, v) = exp\big\{ -[(-\log u)^\alpha + (-\log v)^\alpha]^{1/\alpha} \big\},$$

$$\alpha \geq 1, 0 \leq u, v \leq 1$$

2. Clayton copula:

$$C_\alpha(u, v) = [\max(u^{-\alpha} + v^{-\alpha} - 1, 0)]^{-1/\alpha},$$

$$\alpha > -1 \quad \text{and} \quad \alpha \neq 0, 0 \leq u, v \leq 1$$

3. Frank copula:

$$C_\alpha(u,v) = -\frac{1}{\alpha}\log\left\{1+\frac{(e^{-\alpha u}-1)(e^{-\alpha v}-1)}{e^{-\alpha}-1}\right\},$$
$$\alpha \neq 0, 0 \leq u, v \leq 1$$

Archimedean copulas are widely used in applications due to their simplicity, flexibility of dependence structures, and the ability to extend to a higher dimensional problem via the associativity property. A collection of twenty-two one-parameter families of Archimedean copulas can be found in Table 4.1 of Nelsen (2006).

2.2 *Univariate survival function estimation*

Let S_j and F_j, $j = 1, 2$, denote the survival function and distribution function of T_j^0, respectively.

2.2.1 Right censored data Kaplan & Meier (1958) introduced the product limit estimator of survival function, which is a basic tool to estimate the probability of survival at time t for right censored data. S_1 can therefore be estimated by

$$\hat{S}_1(t) = \prod_{i:t_i<t} \frac{n_i - d_i}{n_i},$$

where n_i is the number at risk just prior to observation time t_i, and d_i is the number of deaths at time t_i.

2.2.2 Current status data For i.i.d. current status data $(C_{2i}, \Delta_{2i}), i = 1, 2, \cdots, n$, the NPMLE $\hat{F}_2$ of the distribution function F_2 maximizes the following likelihood:

$$l(F_2) = \sum_{i=1}^{n} \{\Delta_{2i} \log F_2(C_{2i}) + (1 - \Delta_{2i}) \log(1 - F_2(C_{2i}))\},$$

under the assumption that T_2^0 and C_2 are independent. Turnbull (1974) derived a self-consistency equation for $\hat{F}_2$:

$$\hat{F}_2(c) = \mathrm{E}_{\hat{F}_2}\{\tilde{F}_{2n}(c) | C_{21}, \cdots, C_{2n}, \Delta_{21}, \cdots, \Delta_{2n}\},$$

where $\tilde{F}_{2n}$ is the (unobservable) empirical distribution function of the random variables $T_{21}^0, \cdots, T_{2n}^0$. This equation immediately yields the iteration steps of the Expectation Maximization (EM) algorithm (Dempster et al., 1977), which can be used to solve for $\hat{F}_2$.

Groeneboom & Wellner (1992) introduced an algorithm called Convex Minorant Algorithm (CMA) to maximize the above likelihood from a different characterization of NPMLE. Let $C_{2(i)}$ be the ith order statistics of C_2 and let $\Delta_{2(i)}$ be the corresponding indicator, then $\hat{F}_2$ can be obtained explicitly through the "max-min" formula:

$$\hat{F}_2(c_{2(i)}) = \max_{m \le i} \min_{k \ge i} \frac{\sum_{m \le j \le k} \Delta_{2(j)}}{k - m + 1}.$$

As suggested by Groeneboom & Wellner (1992), the CMA is considerably faster than the commonly used EM method, especially when the sample size is large. The NPMLE of survival function $\hat{S}_2$ can be obtained by $1 - \hat{F}_2$.

2.3 *Bivariate survival models*

First, we write the joint survival function into a copula structure

$$S_\alpha(t_1, t_2) = C_\alpha(S_1(t_1), S_2(t_2)) \quad \alpha \in \mathbf{R}^1,$$

where α is the association parameter. Let $F_\alpha(t_1, t_2)$ denote the joint distribution function corresponding to $S_\alpha(t_1, t_2)$.

We consider a special case when T_1^0 is right censored by a random time C_1 and T_2^0 is subject to interval censoring case 1 by a monitoring time C_2. The observed data, $\mathbf{X}$, is described in (1). Throughout the manuscript, we assume the independent and noninformative censoring.

3. Copula based pseudolikelihood estimation of association parameter

Let $(T_{1i}, T_{2i}, \Delta_{1i}, \Delta_{2i}), i = 1, \cdots, n$, be an i.i.d. sample, each with density $h(t_1, t_2, \delta_1, \delta_2)$ given by

$$\begin{aligned}
&\lim_{\substack{h_1 \to 0+ \\ h_2 \to 0+}} \frac{P[t_1 \le T_1 < t_1 + h_1, t_2 \le T_2 < t_2 + h_2, \Delta_1 = \delta_1, \Delta_2 = \delta_2]}{h_1 h_2} \\
= \;& \big[\frac{\partial}{\partial t_1} F_\alpha(t_1, t_2)\big]^{\delta_1\delta_2} \big[- \frac{\partial}{\partial t_1} S_\alpha(t_1, t_2)\big]^{\delta_1(1-\delta_2)} \\
& \times \big[S_1(t_1) - S_\alpha(t_1, t_2)\big]^{(1-\delta_1)\delta_2} \big[S_\alpha(t_1, t_2)\big]^{(1-\delta_1)(1-\delta_2)}.
\end{aligned}$$

Let $C_{1\alpha}(u, v) = \frac{\partial}{\partial u} C_\alpha(u, v)$. Note that $F_\alpha(t_1, t_2) = 1 - S_1(t_1) - S_2(t_2) + S_\alpha(t_1, t_2)$. Given two marginal survival functions S_1, S_2, the likelihood of the association parameter α based on all observations is

$$\begin{aligned}
&\prod_{i=1}^{n} \big[1 - C_{1\alpha}\big(S_1(t_{1i}), S_2(t_{2i})\big)\big]^{\delta_{1i}\delta_{2i}} \big[C_{1\alpha}\big(S_1(t_{1i}), S_2(t_{2i})\big)\big]^{\delta_{1i}(1-\delta_{2i})} \\
\times &\big[S_1(t_{1i}) - C_\alpha\big(S_1(t_{1i}), S_2(t_{2i})\big)\big]^{(1-\delta_{1i})\delta_{2i}} \big[C_\alpha\big(S_1(t_{1i}), S_2(t_{2i})\big)\big]^{(1-\delta_{1i})(1-\delta_{2i})},
\end{aligned} \tag{2}$$

by omitting the parts that irrelevant in estimating α. Our main interest is to estimate the association parameter α. We propose to apply a two-stage pseudolikelihood approach. At

the first stage, the marginal survival function S_1, which corresponds to the right censored data, is estimated by Kaplan-Meier estimator $\hat{S}_1$, and S_2, which corresponds to current status data, is estimated by $\hat{S}_2$ using Convex Minorant Algorithm. At the second stage, the estimates $\hat{S}_1$ and $\hat{S}_2$ are plugged into the likelihood (2), and the resulted log pseudolikelihood is then maximized to get the estimator of α, $\hat{\alpha}_n$, which is the solution to the pseudo score equation:

$$U_\alpha(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) = \sum_{i=1}^{n} \frac{\partial}{\partial \alpha} l(\alpha, \hat{S}_1(t_{1i}), \hat{S}_2(t_{2i}), \delta_{1i}, \delta_{2i}) = 0, \tag{3}$$

where

$$\begin{aligned} & l(\alpha, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) \qquad (4) \\ = \; & \delta_1\delta_2 \log\left(1 - C_{1\alpha}(\hat{S}_1(t_1), \hat{S}_2(t_2))\right) + \delta_1(1-\delta_2)\log C_{1\alpha}(\hat{S}_1(t_1), \hat{S}_2(t_2)) \\ & +(1-\delta_1)\delta_2 \log\left(\hat{S}_1(t_1) - C_\alpha(\hat{S}_1(t_1), \hat{S}_2(t_2))\right) \\ & +(1-\delta_1)(1-\delta_2)\log C_\alpha(\hat{S}_1(t_1), \hat{S}_2(t_2)). \end{aligned}$$

The pseudolikelihood estimation approach allows the functional form of the marginal survival functions to be flexible, and is determined by data. It's also computationally easy since only the association parameter is left as unknown in the pseudolikelihood.

4. Asymptotic properties of the pseudo estimator $\hat{\alpha}_n$

We now define some notations to be used in the sequel. Let $l(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$ be defined as in (4) on $[0, t_{01}] \times [0, t_{02}]$, $t_{01} = \sup\{t : P(T_1 > t, C_1 > t) > 0\}$ and $t_{02} = \sup\{t : P(C_2 > t) > 0\}$. Suppose α is in an open set A in the real line. We let D be a constant, which may represent different values at different places.

Before we formally state the asymptotic results, we need to define the following notations:

$$
\begin{aligned}
V_{\alpha}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial}{\partial \alpha} l(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) \\
V_{\alpha^2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^2}{\partial \alpha^2} l(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) \\
V_{\alpha,1}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^2}{\partial \alpha \partial u} l(\alpha, u, S_2(t_2), \delta_1, \delta_2)|_{u=S_1(t_1)} \\
V_{\alpha,2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^2}{\partial \alpha \partial v} l(\alpha, S_1(t_1), v, \delta_1, \delta_2)|_{v=S_2(t_2)} \\
V_{\alpha^2,1}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^3}{\partial \alpha^2 \partial u} l(\alpha, u, S_2(t_2), \delta_1, \delta_2)|_{u=S_1(t_1)} \\
V_{\alpha^2,2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^3}{\partial \alpha^2 \partial v} l(\alpha, S_1(t_1), v, \delta_1, \delta_2)|_{v=S_2(t_2)} \\
V_{\alpha,1^2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^3}{\partial \alpha \partial u^2} l(\alpha, u, S_2(t_2), \delta_1, \delta_2)|_{u=S_1(t_1)} \\
V_{\alpha,1,2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^3}{\partial \alpha \partial u \partial v} l(\alpha, u, v, \delta_1, \delta_2)|_{u=S_1(t_1), v=S_2(t_2)} \\
V_{\alpha,2^2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) &= \frac{\partial^3}{\partial \alpha \partial v^2} l(\alpha, S_1(t_1), v, \delta_1, \delta_2)|_{v=S_2(t_2)}
\end{aligned}
$$

Suppose the following regularity conditions hold:

(A1) $l(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$ is three-times differentiable with respect to α on $[0, t_{01}] \times [0, t_{02}]$, for each $\alpha \in A$, and all derivatives are continuous and uniformly bounded by some constant D.

(A2) $V_{\alpha,1}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$, $V_{\alpha,2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$, $V_{\alpha^2,1}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$, $V_{\alpha^2,2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$, $V_{\alpha,1^2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$, $V_{\alpha,1,2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$, and $V_{\alpha,2^2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$ are continuous and uniformly bounded by some constant D on $[0, t_{01}] \times [0, t_{02}]$, for all $\alpha \in A$.

(A3) For each $\alpha \in A$, $0 < \mathrm{E}_{\alpha}[V_{\alpha}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)]^2 < \infty$.

(A4) Let F_2, G_2 be distribution functions of T_2^0 and C_2, respectively. $G_2 \ll F_2, F_2 \ll G_2$,

and G_2 has density g_2 with respect to the Lebesgue measure.

(A5) $(\psi_2/g_2) \circ S_2^{-1}$ is bounded and Lipchitz on $[0,1]$, where ψ_2 is the derivative of the influence curve $IC_2(t_2)$, which is defined in the appendix.

(A6) S_2, g_2 and ψ_2 satisfy

$$\int_0^{t_{02}} \frac{S_2(t_2)(1-S_2(t_2))}{g_2(t_2)} \psi_2(t_2) dt_2 < \infty.$$

The above regularity conditions hold for the bivariate copula models mentioned earlier given that the marginal distribution functions are smooth.

Some technical lemmas are needed for proving the asymptotic results and they are stated as follows:

LEMMA 1. *Let* $\mathcal{F}_j = \{f : f$ *is a survival function on* $[0, t_{0j}]\}$, $j = 1, 2$, *and the class* $\mathcal{G}_{\mathcal{F}} = \{V_{\alpha,1}(\alpha, f_1(t_1), f_2(t_2), \delta_1, \delta_2);\ f_j \in \mathcal{F}_j,\ j = 1, 2\}$. *Let* P *denote the probability measure of* $(T_1, T_2, \Delta_1, \Delta_2)$, *then under condition (A1)-(A2),* $\mathcal{G}_{\mathcal{F}}$ *is a P-Glivenko-Cantelli class, for all* $\alpha \in A$.

LEMMA 2. *Let* $\mathcal{F}_j = \{f : f$ *is a survival function on* $[0, t_{0j}]\}$, $j = 1, 2$ *and the class* $\mathcal{H}_{\mathcal{F}} = \{V_\alpha(\alpha, f_1(t_1), f_2(t_2), \delta_1, \delta_2) - V_\alpha(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2) :\ f_j \in \mathcal{F}_j,\ j = 1, 2\}$. *Let* P *denote the probability measure of* $(T_1, T_2, \Delta_1, \Delta_2)$, *then under condition (A1)-(A2),* $\mathcal{H}_{\mathcal{F}}$ *is a P-Donsker Class, for all* $\alpha \in A$.

Under the regularity conditions stated previously, the estimator $\hat{\alpha}_n$, the solution to equation (3), is consistent and has the asymptotic normal distribution as stated in the following two theorems:

THEOREM 1. *Assume that the joint distribution of* (T_1^0, T_2^0) *follows an Archimedean copula model with the true association parameter* $\alpha = \alpha_0$. *Let* $\hat{S}_1(\cdot)$ *be K-M estimator of* $S_1(\cdot)$ *and* $\hat{S}_2(\cdot)$ *be the NPMLE estimator of* $S_2(\cdot)$ *by CMA. Under the regularity conditions (A1)-(A2),* $\hat{\alpha}_n \xrightarrow{p} \alpha_0$ *as* $n \to \infty$.

THEOREM 2. *Under the regularity conditions (A1)-(A6),* $\sqrt{n}(\hat{\alpha}_n - \alpha_0) \xrightarrow{d} N(0, \sigma^2)$, *where*

$$\sigma^2 = \frac{Var(Q(\alpha_0, S_1, S_2, t_1, t_2, \delta_1, \delta_2))}{W^2(\alpha_0, S_1, S_2, \delta_1, \delta_2)}$$

with

$$\begin{aligned} W(\alpha_0, S_1, S_2, \delta_1, \delta_2) &= -\int \left[V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)\right]^2 dP(t_1, t_2, \delta_1, \delta_2) \\ Q(\alpha_0, S_1, S_2, t_1, t_2, \delta_1, \delta_2) &= V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2) \\ &\quad + I_1(T_1, \Delta_1, \alpha_0) - \tilde{l}(t_2, \delta_2, S_2, G_2, \psi_2), \end{aligned}$$

in which

$$\begin{aligned} I_1(T_1, \Delta_1, \alpha_0) &= \int_0^{t_{01}} \int_0^{t_{02}} M_{\alpha,u}(\alpha_0, S_1(t_1), S_2(t_2)) f(t_1, t_2) I_1^0(T_1, \Delta_1)(t_1) dt_1 dt_2 \\ \tilde{l}(t_2, \delta_2, S_2, G_2, \psi_2) &= -[\delta_2 - (1 - S_2(t_2))] \frac{\psi_2(t_2)}{g_2(t_2)} I[g_2(t_2 > 0)], \end{aligned}$$

where

$$M_{\alpha,u}(\alpha_0, S_1(t_1), S_2(t_2)) = -E_{\delta_1\delta_2|t_1t_2} V_{\alpha,u}(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$$

and

$$I_1^0(T_1, \Delta_1)(t_1) = -S_1(t_1)\Big\{ \int_0^{t_1} \frac{1}{P(T_1 \geq u)} dN_1(u) - \int_0^{t_1} \frac{I[T_1 \geq u]}{P(T_1 \geq u)} d\Lambda_1(u) \Big\}.$$

$N_1(u)$ *is defined as* $I[T_1 \leq u, \Delta_1 = 1]$ *and* Λ_1 *is the cumulative hazard function of* T_1^0.

The proof of the two lemmas and the two theorems are given in the Appendix.

5. Simulation studies

The preceding sections provide a two-stage pseudo-likelihood estimation procedure for the association parameter between the two survival times. Although the estimator is shown to be asymptotically consistent and normally distributed, it is crucial to ascertain its finite sample performance before applying it to real problems. Simulation studies are preformed to evaluate the proposed estimator.

We consider the Gumbel copula function

$$C_\alpha(u,v) = exp\big\{ -[(-\log u)^\alpha + (-\log v)^\alpha]^{1/\alpha}\big\},$$

where $\alpha \geq 1$, and two margins are both assumed to be exponentially distributed with unit rate 1.

A sample of bivariate copula random variables is generated based on conditional distribution function. Suppose that the joint distribution of the bivariate data (T_1^0, T_2^0) is $C_\alpha(F_1(t_1), F_2(t_2))$. We generate (T_1^0, T_2^0) through the following steps:

- Generate two independent uniform $(0,1)$ random variables u, w.
- Set $w = P(V \leq v|U = u) = \partial C_\alpha(u,v)/\partial u$, solve for v.
- Set $T_1^0 = F_1^{-1}(u), T_2^0 = F_2^{-1}(v)$.

Meanwhile, a sample of bivariate censoring times (C_1 and C_2) are each independently drawn from a uniform distribution on $[0, 2.3]$. In this setting, about 50% of T_1^0 is right censored by C_1, and about 50% of T_2^0 is subject to interval censoring case 1 by C_2 as well.

Kendall's τ is chosen as a global association measure. For Gumbel copula, $\tau = 1 - 1/\alpha$. Three different values of α are set such that the corresponding Kendall's τ is 0.25, 0.5, and 0.75. For each value of α, we conduct Monte-Carlo simulations with $1,000$ replications for sample size $n = 50,\ 100,\ 200$ and 400, respectively.

We compute the two-stage pseudolikelihood estimator $\hat{\alpha}_n$ as proposed previously. We also compute $\tilde{\alpha}_n$, the maximum likelihood estimator when the two margins are completely specified. The latter estimator serves as a benchmark to evaluate the performance of the pseudolikelihood estimator.

For each of the $1,000$ simulations, Wald confidence interval is constructed based on the asymptotic normality, in which the asymptotic variance of $\hat{\alpha}_n$ is computed using 200 bootstrap resamples. The empirical estimate of the coverage probability is obtained based on the Wald confidence interval over $1,000$ replications.

Table 1 summarizes the simulation results for the two-stage pseudolikelihood estimator. It provides results for estimation bias, standard deviation from $1,000$ replicates (sd), mean of bootstrap standard deviation (bsd^*), and 95% empirical coverage probability.

As the sample size increases, for a wide range of α, the bias of both $\hat{\alpha}_n$ and $\hat{\tau}_n$ decreases considerably; the standard deviation and the bootstrapped standard error are dropping down as well and they are getting closer and closer. In addition, the empirical coverage probability converges to the nominal level when the sample size increases.

Given the same sample size, the stronger the dependency, the bigger the bias and the

Table 1

Simulation results of the two-stage pseudolikelihood estimator based on 1000 Monte-Carlo samples with sample size chosen as 50,100,200,400 for $\alpha = 4/3,\ 2,\ 4$.

		n=50		n=100		n=200		n=400	
		$\hat{\alpha}_n$	$\hat{\tau}_n$	$\hat{\alpha}_n$	$\hat{\tau}_n$	$\hat{\alpha}_n$	$\hat{\tau}_n$	$\hat{\alpha}_n$	$\hat{\tau}_n$
$\tau = 0.25$	Bias	0.219	0.043	0.059	0.013	0.023	0.005	-0.005	-0.002
$\alpha = 1.333$	*sd*	0.845	0.172	0.233	0.113	0.142	0.076	0.098	0.055
	bse^*	8.799	0.161	0.334	0.109	0.154	0.077	0.099	0.054
	95% Cover.P	0.968		0.966		0.963		0.954	
$\tau = 0.50$	Bias	1.120	0.051	0.194	0.021	0.102	0.014	0.032	0.003
$\alpha = 2.0$	*sd*	8.090	0.158	0.563	0.098	0.320	0.070	0.208	0.050
	bse^*	26.276	0.156	2.523	0.101	0.359	0.069	0.213	0.048
	95% Cover.P	0.985		0.976		0.966		0.957	
$\tau = 0.75$	Bias	9.460	0.176	0.695	0.037	0.189	0.017	0.058	0.004
$\alpha = 4.0$	*sd*	51.64	0.117	4.305	0.081	1.002	0.054	0.646	0.038
	bse^*	60.48	0.124	15.726	0.079	1.597	0.054	0.696	0.038
	95% Cover.P	0.991		0.980		0.974		0.959	

bse^* : Bootstrap standard error.

standard error for the estimator $\hat{\alpha}_n$, as greater variations are usually expected for larger values. Therefore, to preserve high efficiency, large sample size is desired to achieve the reasonable performance of $\hat{\alpha}_n$ when a strong dependence exists. Interestingly, we observe that the standard deviation of $\hat{\tau}_n$ decreases when the dependence becomes stronger. By delta method, $\sigma_{\hat{\tau}_n} \approx \sigma_{\hat{\alpha}_n}/\alpha^2$, where $\sigma_{\hat{\tau}_n}$ and $\sigma_{\hat{\alpha}_n}$ are standard deviation of $\hat{\tau}_n$ and $\hat{\alpha}_n$, respectively. Therefore, as τ increases, the bias of $\hat{\tau}$ increases but the variance of $\hat{\tau}_n$ may still decrease. Additionally, we can also observe that the estimated standard error of both $\hat{\tau}_n$ and $\hat{\alpha}_n$ closely fit the association suggested by delta method when the sample size is over 200.

Table 2 gives the results of $\tilde{\alpha}$, the maximum likelihood estimator α, when the two marginal survival functions are known. $\tilde{\alpha}_n$ performs better than $\hat{\alpha}_n$, as expected, while their difference is substantially reduced when the number of observations increases, for example, $n \geq 200$. The small difference between the two estimators assures us the use of two-stage pseudo estimation procedure, in which we gain advantage of having flexibility for not modeling the marginal distributions without loss of too much estimation efficiency, given a reasonable sample size ($n \geq 200$ for a wide range of α).

6. Application to the Motivating Example

This study is designed to determine the effect of GBV-C virus on the survival benefit among HIV patients. The sub-cohort of MACS from Williams et al. (2004) is re-analylized. MACS consists of gay men who were enrolled between 1984 and 1990 and whose blood samples were obtained every 6 months. The sub-cohort includes 271 subjects from MACS who were initially HIV negative when they entered the study but HIV positive during the

Table 2

Simulation results of MLE when S_1 and S_2 are fixed at known. 1000 monto carlo datasets are generated and sample sizes are chosen as 50,100,200,400 for $\alpha = 4/3,\ 2,\ 4$.

		n=50		n=100		n=200		n=400	
		$\tilde{\alpha}_n$	$\tilde{\tau}_n$	$\tilde{\alpha}_n$	$\tilde{\tau}_n$	$\tilde{\alpha}_n$	$\tilde{\tau}_n$	$\tilde{\alpha}_n$	$\tilde{\tau}_n$
$\tau = 0.25$	Bias	0.065	0.031	0.019	0.011	0.016	0.004	-0.004	-0.001
$\alpha = 1.333$	sd	0.334	0.145	0.192	0.102	0.136	0.076	0.097	0.053
	bse^*	1.253	0.136	0.218	0.101	0.136	0.073	0.094	0.053
	95% Cover.P	0.940		0.942		0.954		0.949	
$\tau = 0.50$	Bias	0.302	0.036	0.069	0.018	0.022	0.005	-0.009	-0.002
$\alpha = 2.0$	sd	1.360	0.141	0.455	0.096	0.288	0.068	0.190	0.047
	bse^*	8.808	0.132	0.811	0.093	0.288	0.068	0.196	0.047
	95% Cover.P	0.965		0.951		0.939		0.952	
$\tau = 0.75$	Bias	7.136	0.160	0.539	0.030	0.164	0.010	0.014	0.001
$\alpha = 4.0$	sd	41.24	0.104	3.830	0.073	0.975	0.050	0.635	0.038
	bse^*	44.55	0.089	12.353	0.067	1.590	0.049	0.659	0.036
	95% Cover.P	0.971		0.963		0.960		0.940	

bse^* : Bootstrap standard error.

follow ups. Since the visits were scheduled every 6 months, the seroconversion time is known to be within a six-month window. Seroconversion time is imputed as the midpoint between the last seronegative visit and the first seropositive visit. All the 271 subjects were evaluated at 12-18 months after HIV seroconversion for the evidence of GBV-C infection and a subgroup of 138 patients were re-examined 5-6 years after HIV seroconversion. The study only included data collected before Jan 1, 1996 to avoid the confounding due to the use of highly active antiretroviral therapy.

We consider the association between the duration of GBV-C infection and the HIV survival among people who were co-infected with both HIV and GBV-C at HIV onset. Therefore, the HIV survival is defined as the time from seroconversion to death, and the GBV-C infection time is defined as the time from seroconversion to GBV-C clearance.

In our analysis, we treat the GBV-C status evaluated at 12-18 months as the baseline GBV-C information to select a subsample of HIV patients who are assumed to be co-infected with GBV-C at baseline. The GBV-C status evaluated at 5-6 years after HIV seroconversion gives us the current status data for GBV-C infection time.

Gumbel copula is used for the bivariate distribution of HIV survival and GBV-C infection time. Table 3 gives the result for the pseudolikelihood estimator of association between HIV survival time and GBV-C infection time. Only 61 patients who were GBV-C positive at the first visit and GBV-C status at the late visit were known are included.

Boostrap standard deviation based on 1000 resamples with replacement was used to estimate the standard error and to construct the Wald confidence interval. The estimator

Table 3

Association between HIV survival time and GBV-C infection time; Only include patients who are GBV-C positive at early visit and GBV-C known at late visit (N=61).

	Estimate	Bootstrap standard error	95% Wald CI
$\hat{\alpha}_n$	2.0143	0.4558	$[1.1208, 2.9077]$
$\hat{\tau}_n$	0.5035	0.1450	$[0.2193, 0.7877]$

$\hat{\alpha}_n$ is about 2, with 95% Wald confidence interval ruled out 1. The corresponding Kendall's τ is 0.5, and its 95% Wald confidence interval did not include zero. Therefore, the GBV-C persistence appears to be associated with increased survival among HIV and GBV-C co-infected individuals. These results support recent articles suggesting increased survival in individuals co-infected with GBV-C.

7. Discussion

This paper proposes a way of assessing the association between two random variables which are subject to different censoring schemes, one is right censored and the other is current status data. The asymptotic properties of the estimator are established under mild technical assumptions. Although the asymptotic variance of the estimator is difficult to obtain analytically, the ordinal bootstrap method provides a practical and efficient way to estimate the variance.

Our simulation results suggest that the proposed estimator works well for moderate sample sizes and has the advantage of allowing for flexibility in the marginals. Moreover,

our numerical work shows that the proposed method is approximately as efficient as full maximum likelihood approach when the marginal distributions are known, with a moderate sample size. It suggests that the efficiency loss from the pseudolikelihood approach by estimating the marginals separately, if any, is minimal.

We assume Gumbel copula in the numerical examples (simulation and real data analysis). The selection of copula should not be an important issue here. Our main interest is whether there is or is not positive dependence, not the strength of the dependence. According to Wang & Wells (2000), when τ is small all the copula models behave similarly. In fact when $\tau \to 0$ all the copula models approach the independent copula; when there is positive dependence, any copula model can capture it, although the strength of dependency can be different.

Rather than considering an association model, another natural way of formulating the same problem is maybe through modeling the right censored data with a time-dependent covariate, the current status data. However, this approach is impossible for the data structure specified in (1), since current status data only provides information at one monitoring time and can not be specified at any failure time, the feature that is required for cox model with a time-dependent covariate (Kalbfleisch & Prentice (2002), page 200).

Throughout we did not provide the hypothesis testing on the independence, due to a potential boundary problem. If a copula function, such as the Gumbel copula, is equivalent to the independent copula only when the association parameter takes its value on the boundary of the parameter space, some regularity conditions fail and the likelihood theory is broken

down. Donald (2001) investigated the testing problems which covers this sort. He specifies a set of high level conditions under which the asymptotic null distributions of quasi-likelihood ratio (QLR), rescaled quasi-likelihood ratio (RQLR), Wald and score tests are determined. A further exploration may be considered by following his approach. Another difficulty this boundary problem brings is the constructing of confidence interval. The lower bound of the Wald confidence interval could fall outside of the parameter space if the association is too weak. In this case a Likelihood Ratio type of confidence interval is preferable to the Wald confidence interval.

In this work, we did not account for the covariate effect when modeling the association between two event times. A further consideration would be given to using either the Cox Proportional hazard model or the accelerated hazard model to perform a regression analysis on each event time marginally. A more thorough investigation in this direction is needed.

Note that the dataset used for illustration in this paper is not an ideal example since we only had an available sample size of 61. The simulation studies suggest that the sample size of bigger than 200 is desired to obtain a well behaved estimator for a wide range of dependence. If the sample size is small, the validity of the inference is questionable. Including the HIV patients who were never infected with GBV-C at baseline, and performing a complete analysis based on a mixture distribution of GBV-C clearance time, is currently under development.

ACKNOWLEDGEMENTS

The authors wish to thank the Multicenter AIDS Cohort Study (MACS) for providing data. The MACS has centers located at: The Johns Hopkins Bloomberg School of Public Health

(Joseph Margolick); Howard Brown Health Center and Northwestern University Medical School (John Phair); University of California, Los Angeles (Roger Detels); University of Pittsburgh (Charles Rinaldo); and Data Analysis Center (Lisa Jacobson).

References

Birk, M., Lindback, S. & Lidman, C. (2002). No influence of GB virus C replication on the prognosis in a cohort of HIV-1-infected patients. *AIDS* 16 2482–2485.

Björkman, P., Flamholc, L., Naucler, A., Molnegren, V., Wallmark, E. & Widell, A. (2004). GB virus C viremia during the natural course of HIV-1 infection: viremia at diagnosis does not predict mortality. *AIDS* 18 877–886.

Dempster, A., Laird, N. & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B* 39 1–38.

Donald, W. K. A. (2001). Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica* 69 683–734.

Fleming, T. R. & Harrington, D. P. (1991). *Counting processes and survival analysis*. New York: Wiley.

Groeneboom, P. & Wellner, J. A. (1992). *Information Bounds and nonparametric maximum likelihood estimation*. Boston: Birkhäuser.

Hougaard, P. (1989). Fitting a multivariate failure time distribution. *IEEE Transactions on Reliability* 38 444–448.

Huang, J. & Wellner, A. (1995). Asymptotic normality of the NPMLE of linear functionals for interval censored data, case 1. *Statistica Neerlandica* 49 153–163.

Huang, J. & Wellner, J. A. (1997). Interval censored survival data: A review of recent progress. In *Proceedings of the first Seattle symposium in biostatistics: Survival analysis*. New York: Springer-Verlag, 123–169.

Kalbfleisch, J. D. & Prentice, R. L. (2002). *The statistical analysis of failure time data*. Wiley-Interscience, 2nd ed.

Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.* 53 457–481.

Liang, K. E., Self, S. G., Bandeen-Rocche, K. & Zeger, S. (1995). Some recent developments for regression analysis of multivariate failure time data. *Lifetime Data Analysis* 1 403–415.

Mises, R. V. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* 18 309–348.

Nelsen, R. B. (2006). *An Introduction to Copulas*. New York: Springer-Verlag, 2nd ed.

Oakes, D. (1989). Bivariate survival models induced by frailties. *J. Am. Statist. Assoc.* 84 487–493.

Oakes, D. (2000). Survival analysis. *J. Am. Statist. Assoc.* 95 282–285.

Pepe, M. S. (1991). Inference for events with dependent risks in multiple endpoint studies. *J. Am. Statist. Assoc.* 86 770–778.

Shih, J. H. & Louis, T. (1995). Inference on the association parameter in copula models for bivariate survival data. *Biometrics* 51 1384–99.

Tillmann, H., Heiken, H., Knapik-Botor, A., Heringlake, S., Ockenga, J. & et al (2001). Infection with GB virus C and reduced mortality among HIV-infected patients. *N Engl. J Med.* 345 715–724.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped censored and truncated data. *J. R. Statist. Soc. B* 38 290–295.

Van der Bij, A., Kloosterboer, N., Prins, M., Boeser-Nunnink, B., Lange, J. & et al (2005). GB virus C coinfection and HIV-1 disease progression: The amsterdam cohort study. *J Infect Dis* 191 678–685.

van der Vaart, A. W. (1998). *Asymptotic statistics*. Cambridge: Cambridge Univ. Press.

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

Wang, W. & Ding, A. A. (2000). On assessing the association for bivariate current status data. *Biometrika* 87 879–893.

Wang, W. & Wells, M. T. (2000). Model selection and semiparametric inference for bivariate failure time data. *J. Am. Statist. Assoc.* 95 62–72.

Williams, C., Klinzman, D., Yamashita, T., Xiang, J. & et al (2004). Persistent GB virus C infection and survival in HIV-infected men. *N Engl. J Med.* 350 981–990.

Xiang, J., Wunschmann, W., Diekema, D., Klinzman, D., Patrick, K. & et al (2001). Effect of coinfection with GB virus C on survival among among patients with HIV infection. *N Engl. J Med.* 345 707–714.

Appendix

This section provides a sketch of proofs for the lemmas and theorems stated in section 4. We use the modern empirical process theory justifying our proof. We denote $\int f dP$ by Pf and $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$ by $\mathbb{P}_n f$.

Proof of Lemma 1:

Since $\mathcal{F}_j$ consists of uniformly bounded monotone functions on the real line, by the Theorem 2.7.5 of van der Vaart & Wellner (1996), for any $\epsilon > 0$, for $j = 1, 2$, there exists a set of brackets:

$$[f_{j1}^L, f_{j1}^U], [f_{j2}^L, f_{j2}^U], \cdots, [f_{jN_j}^L, f_{jN_j}^U],$$

with $N_j \leq \exp(D/\epsilon)$ and $\int |f_{ji}^U - f_{ji}^L| dP \leq \epsilon$ for any $1 \leq i \leq N_j$, such that for any $f_j \in \mathcal{F}_j$ and any $t_j \in [0, t_{0j}]$, $f_{jq_j}^L(t_j) \leq f_j(t_j) \leq f_{jq_j}^U(t_j)$ for some $1 \leq q_j \leq N_j$.

By condition $(A2)$, $V_{\alpha,1}(\alpha, f_1(t_1), f_2(t_2), \delta_1, \delta_2)$ is continuous. We can then construct a set of bracket as following: For any $i = 1, 2, \cdots, N_1, s = 1, 2, \cdots, N_2$ and for any $t_j \in [0, t_{0j}]$, we can find the unique maximum and minimum of $V_{\alpha,1}(\alpha, f_1(t_1), f_2(t_2), \delta_1, \delta_2)$ on the product

set $[f_{1i}^L, f_{1i}^U] \times [f_{2s}^L, f_{2s}^U]$. Let

$$(f_1^{L,(i,s)}(t_1), f_2^{L,(i,s)}(t_2)) = \underset{\substack{f_1 \in [f_{1i}^L, f_{1i}^U] \\ f_2 \in [f_{2s}^L, f_{2s}^U]}}{\operatorname{argmin}} V_{\alpha,1}(\alpha, f_1(t_1), f_2(t_2), \delta_1, \delta_2)$$

$$(f_2^{U,(i,s)}(t_1), f_2^{U,(i,s)}(t_2)) = \underset{\substack{f_1 \in [f_{1i}^L, f_{1i}^U] \\ f_2 \in [f_{2s}^L, f_{2s}^U]}}{\operatorname{argmax}} V_{\alpha,1}(\alpha, f_1(t_1), f_2(t_2), \delta_1, \delta_2)$$

and let

$$V_{\alpha,1}^{L,(i,s)}(t_1, t_2, \delta_1, \delta_2) = V_{\alpha,1}(\alpha, f_1^{L,(i,s)}(t_1), f_2^{L,(i,s)}(t_2), \delta_1, \delta_2)$$

$$V_{\alpha,1}^{U,(i,s)}(t_1, t_2, \delta_1, \delta_2) = V_{\alpha,1}(\alpha, f_1^{U,(i,s)}(t_1), f_2^{U,(i,s)}(t_2), \delta_1, \delta_2).$$

The class $\mathcal{G}_{\mathcal{F}}$ is then covered by a set of $N_1 \times N_2$ brackets:

$$\{[V_{\alpha,1}^{L,(i,s)}(t_1, t_2, \delta_1, \delta_2), V_{\alpha,1}^{U,(i,s)}(t_1, t_2, \delta_1, \delta_2)] : i = 1, 2, \cdots, N_1, s = 1, 2, \cdots, N_2\}.$$

By condition (A2), $V_{\alpha,1^2}(\alpha, u, v, \delta_1, \delta_2)$ and $V_{\alpha,1,2}(\alpha, u, v, \delta_1, \delta_2)$ are bounded by some constant D on $[0, t_{01}] \times [0, t_{02}]$, then $V_{\alpha,1}(\alpha, u, v, \delta_1, \delta_2)$ satisfies the Lipchitz condition with respect to u and v. It follows that:

$$\begin{aligned}
& \int |V_{\alpha,1}^{U,(i,s)}(t_1, t_2, \delta_1, \delta_2) - V_{\alpha,1}^{L,(i,s)}(t_1, t_2, \delta_1, \delta_2)| dP \\
= & \int |V_{\alpha,1}(\alpha, f_1^{U,(i,s)}(t_1), f_2^{U,(i,s)}(t_2), \delta_1, \delta_2) - V_{\alpha,1}(\alpha, f_1^{L,(i,s)}(t_1), f_2^{L,(i,s)}(t_2), \delta_1, \delta_2)| dP \\
\leq & \int \left[D|f_1^{U,(i,s)}(t_1) - f_1^{L,(i,s)}(t_1)| + D|f_2^{U,(i,s)}(t_2) - f_2^{L,(i,s)}(t_2)| \right] dP \\
\leq & D\epsilon.
\end{aligned}$$

This indicates that the preceding $N_1 \times N_2$ brackets are $D\epsilon-$brackets. It follows that, for any $\epsilon > 0$, the bracketing number of class $\mathcal{G}_{\mathcal{F}}$ associated with $L_1(P)$ norm is bounded. By the Theorem 2.4.1 of van der Vaart & Wellner (1996), $\mathcal{G}_{\mathcal{F}}$ is a P-Glivenko-Cantelli class.

Proof of Lemma 2:

Based on the similar technique used in the proof of Lemma 1, we can construct a set of $N_1 \times N_2$ brackets:

$$\{[V_\alpha^{L,(i,s)}(t_1,t_2,\delta_1,\delta_2) - V_\alpha(\alpha,S_1(t_1),S_2(t_2),\delta_1,\delta_2), V_\alpha^{U,(i,s)}(t_1,t_2,\delta_1,\delta_2) - V_\alpha(\alpha,S_1(t_1),S_2(t_2),\delta_1,\delta_2)] : i = 1,2,\cdots,N_1, s = 1,2,\cdots,N_2\},$$

which covers $\mathcal{H}_\mathcal{F}$.

By condition (A2), $V_{\alpha,1}(\alpha,u,v,\delta_1,\delta_2)$ and $V_{\alpha,2}(\alpha,u,v,\delta_1,\delta_2)$ are bounded by some constant D on $[0,t_{01}]\times[0,t_{02}]$, then $V_\alpha(\alpha,u,v,\delta_1,\delta_2)$ satisfies the Lipchitz condition with respect to u and v. Also note that $(x+y)^2 = x^2+y^2+2xy \le 2x^2+2y^2$, it follows that

$$\begin{aligned}
& \int \left(V_\alpha^{U,(i,s)}(t_1,t_2,\delta_1,\delta_2) - V_\alpha^{L,(i,s)}(t_1,t_2,\delta_1,\delta_2)\right)^2 dP \\
= & \int |V_\alpha(\alpha, f_1^{U,(i,s)}(t_1), f_2^{U,(i,s)}(t_2),\delta_1,\delta_2) - V_\alpha(\alpha, f_1^{L,(i,s)}(t_1), f_2^{L,(i,s)}(t_2),\delta_1,\delta_2)|^2 dP \\
\le & \int \left[D|f_1^{U,(i,s)}(t_1) - f_1^{L,(i,s)}(t_1)| + D|f_2^{U,(i,s)}(t_2) - f_2^{L,(i,s)}(t_2)|\right]^2 dP \\
\le & \; 2D^2 \int |f_1^{U,(i,s)}(t_1) - f_1^{L,(i,s)}(t_1)|^2 dP + 2D^2 \int |f_2^{U,(i,s)}(t_2) - f_2^{L,(i,s)}(t_2)|^2 dP \\
\le & \; D\epsilon^2.
\end{aligned}$$

This indicates that the bracketing number of $\mathcal{H}_\mathcal{F}$ associated with $L_2(P)$ norm, denoted by $N_{[\,]}(\epsilon,\mathcal{H}_\mathcal{F},L_2(P))$, is bounded by $N_1 \times N_2$. It follows that $\log\left(N_{[\,]}(\epsilon,\mathcal{H}_\mathcal{F},L_2(P))\right) \le \log(N_1\times N_2) \le D/\epsilon$ for some constant D. Hence,

$$\int_0^1 \sqrt{\log N_{[\,]}(\epsilon,\mathcal{H}_\mathcal{F},L_2(P))}d\epsilon \le \int_0^1 D\epsilon^{-1/2}d\epsilon < \infty.$$

By the Theorem 19.5 of van der Vaart (1998) on page 270, $\mathcal{H}_\mathcal{F}$ is a P-Donsker Class.

Proof of Theorem 1:

Let $\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) = \frac{1}{n} U_\alpha(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2)$, and define $\bar{\mathcal{L}}_n(\alpha, S_1, S_2, \delta_1, \delta_2)$ in a similar way. First we show that $\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) \xrightarrow{p} \mathrm{E}_{\alpha_0} l(\alpha, S_1, S_2, \delta_1, \delta_2)$ for any $\alpha \in A$.

Consider the Taylor series expansion:

$$\begin{aligned} \bar{\mathcal{L}}_n(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) &= \bar{\mathcal{L}}_n(\alpha, S_1, S_2, \delta_1, \delta_2) + (\hat{S}_1 - S_1)\frac{\partial}{\partial u}\bar{\mathcal{L}}_n(\alpha, u, \hat{S}_2, \delta_1, \delta_2) \mid_{u=\tilde{S}_1} \\ &\quad + (\hat{S}_2 - S_2)\frac{\partial}{\partial v}\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, v, \delta_1, \delta_2) \mid_{v=\tilde{S}_2} \end{aligned}$$

where $\tilde{S}_1$ is between S_1 and $\hat{S}_1$, and $\tilde{S}_2$ is between S_2 and $\hat{S}_2$. By the Weak Law of Large Number Theorem,

$$\bar{\mathcal{L}}_n(\alpha, S_1, S_2, \delta_1, \delta_2) \xrightarrow{p} \mathrm{E}_{\alpha_0} l(\alpha, S_1, S_2, \delta_1, \delta_2).$$

Since $\hat{S}_1(\cdot)$ converges in probability to $S_1(\cdot)$ uniformly in $[0, t_{01}]$ (Fleming & Harrington (1991), p.115) and $\hat{S}_2(\cdot)$ converges in probability to $S_2(\cdot)$ uniformly in $[0, t_{02}]$ (Groeneboom & Wellner (1992), 4.1), it suffices to show $\frac{\partial}{\partial u}\bar{\mathcal{L}}_n(\alpha, u, \hat{S}_2, \delta_1, \delta_2) \mid_{u=\tilde{S}_1}$ and $\frac{\partial}{\partial v}\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, v, \delta_1, \delta_2) \mid_{v=\tilde{S}_2}$ converge in probability. Note that:

$$\frac{\partial}{\partial u}\bar{\mathcal{L}}_n(\alpha, u, \hat{S}_2, \delta_1, \delta_2) \mid_{u=\tilde{S}_1} = \mathbb{P}_n V_{\alpha,1}(\alpha, \tilde{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2)$$

$\hat{S}_1(t_1) \xrightarrow{p} S_1(t_1)$, $\tilde{S}_1(t_1)$ is between S_1 and $\hat{S}_1$, then $\tilde{S}_1(t_1) \xrightarrow{p} S_1(t_1)$. $\hat{S}_2(t_2)$ converges to $S_2(t_2)$ as well. In addition, $V_{\alpha,1}$ is continuous by condition $(A2)$. Thus, $V_{\alpha,1}(\alpha, \tilde{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2)$ converges to $V_{\alpha,1}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$, followed by the Continuous Mapping Theorem. Then by the Dominated Convergence Theorem,

$$PV_{\alpha,1}(\alpha, \tilde{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) \xrightarrow{p} PV_{\alpha,1}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2).$$

Let $\mathcal{G}_{\mathcal{F}}$ be defined as in Lemma 1. Lemma 1 shows that under condition (A2), $\mathcal{G}_{\mathcal{F}}$ is a P-Glivenko-Cantelli class. by Van Der Vaart, Page 279, $|\mathbb{P}_n V_{\alpha,1}(\alpha, \tilde{S}_1, \hat{S}_2, \delta_1, \delta_2) - PV_{\alpha,1}(\alpha, \tilde{S}_1, \hat{S}_2, \delta_1, \delta_2)| \stackrel{a.s}{\to} 0$.

It follows that

$$\frac{\partial}{\partial u}\bar{\mathcal{L}}_n(\alpha, u, \hat{S}_2, \delta_1, \delta_2) \mid_{u=\tilde{S}_1} \stackrel{p}{\to} PV_{\alpha,1}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2).$$

Similarly, we can show that

$$\frac{\partial}{\partial v}\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, v, \delta_1, \delta_2) \mid_{v=\tilde{S}_2} \stackrel{p}{\to} PV_{\alpha,2}(\alpha, S_1(t_1), S_2(t_2), \delta_1, \delta_2).$$

This concludes $\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) \stackrel{p}{\to} \mathrm{E}_{\alpha_0} l(\alpha, S_1, S_2, \delta_1, \delta_2)$.

Now, $\forall \alpha \in A$, using Jensen's inequality, it follows that

$$\begin{aligned}
& \bar{\mathcal{L}}_n(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) - \bar{\mathcal{L}}_n(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) \\
\stackrel{p}{\to} \; & \mathrm{E}_{\alpha_0} l(\alpha, S_1, S_2, \delta_1, \delta_2) - \mathrm{E}_{\alpha_0} l(\alpha_0, S_1, S_2, \delta_1, \delta_2) \\
= \; & \mathrm{E}_{\alpha_0} \log \frac{h(\alpha, t_1, t_2, \delta_1, \delta_2)}{h(\alpha_0, t_1, t_2, \delta_1, \delta_2)} < \log \mathrm{E}_{\alpha_0} \frac{h(\alpha, t_1, t_2, \delta_1, \delta_2)}{h(\alpha_0, t_1, t_2, \delta_1, \delta_2)} = 0.
\end{aligned}$$

Due to the convergence demonstrated above, $\forall \epsilon, \delta > 0$, for which $(\alpha_0 - \epsilon, \alpha_0 + \epsilon) \in A$, we may find an integer $N = N(\epsilon, \delta)$, such that, if $n > N$, for $\alpha = \alpha_0 \pm \epsilon$,

$$P(\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) < \bar{\mathcal{L}}_n(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2)) > 1 - \delta.$$

Thus for $n > N$,

$$P\big(\bar{\mathcal{L}}_n(\alpha, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) \text{ has a local maximum } \hat{\alpha}_n \in (\alpha_0 - \epsilon, \alpha_0 + \epsilon)\big) > 1 - 2\delta,$$

because of condition $(A1)$. This immediately shows that the sequence of random variables $\hat{\alpha}_n$ converge in probability to α_0 as $n \to \infty$.

Proof of Theorem 2:

Under condition $(A1)$, the Taylor expansion of the pseudo score function gives

$$0 = \mathbb{P}_n V_\alpha(\hat{\alpha}_n, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) = \mathbb{P}_n V_\alpha(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2)$$

$$+(\hat{\alpha}_n - \alpha_0)\mathbb{P}_n V_{\alpha^2}(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) + O_p(|\hat{\alpha}_n - \alpha_0|^2),$$

then we get

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0) = \frac{\sqrt{n}\mathbb{P}_n V_\alpha(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2)}{-\mathbb{P}_n V_{\alpha^2}(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) - O_p(|\hat{\alpha}_n - \alpha_0|)}.$$

First, we show that

$$\mathbb{P}_n V_{\alpha^2}(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) \xrightarrow{p} W(\alpha_0, S_1, S_2, \delta_1, \delta_2),$$

where

$$\begin{aligned} W(\alpha_0, S_1, S_2, \delta_1, \delta_2) &= P V_{\alpha^2}(\alpha_0, S_1, S_2, \delta_1, \delta_2) \\ &= -P\big[V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)\big]^2. \end{aligned}$$

We can rewrite $\mathbb{P}_n V_{\alpha^2}(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) = \mathbb{P}_n V_{\alpha^2}(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2) + R_n$. Under condition (A2), $V_{\alpha^2}(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$ satisfies the Lipchitz condition. Since $\sup_{t\in[0,t_{01}]} |\hat{S}_1(t_1) - S_1(t_1)| \xrightarrow{a.s} 0$ by Fleming & Harrington (1991), page 115 and $\sup_{t\in[0,t_{02}]} |\hat{S}_2(t_2) - S_2(t_2)| \xrightarrow{a.s} 0$

by Groeneboom & Wellner (1992), page 79, it follows that

$$\begin{aligned}
|R_n| &\leq \mathbb{P}_n|V_{\alpha^2}(\alpha_0, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) - V_{\alpha^2}(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)| \\
&\leq \mathbb{P}_n|V_{\alpha^2}(\alpha_0, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) - V_{\alpha^2}(\alpha_0, S_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2)| \\
&\quad + \mathbb{P}_n|V_{\alpha^2}(\alpha_0, S_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) - V_{\alpha^2}(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)| \\
&\leq D\mathbb{P}_n|\hat{S}_1(t_1) - S_1(t_1)| + D\mathbb{P}_n|\hat{S}_2(t_2) - S_2(t_2)| \\
&\leq D \sup_{0\leq t_1\leq t_{01}} |\hat{S}_1(t_1) - S_1(t_1)| + D \sup_{0\leq t_2\leq t_{02}} |\hat{S}_2(t_2) - S_2(t_2)| \\
&\stackrel{a.s}{\to} 0.
\end{aligned}$$

So, $\mathbb{P}_n V_{\alpha^2}(\alpha_0, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) = \mathbb{P}_n V_{\alpha^2}(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2) + o_p(1)$. By the Weak Law of Large Number Theorem, $\mathbb{P}_n V_{\alpha^2}(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2) \stackrel{p}{\to} PV_{\alpha^2}(\alpha_0, S_1, S_2, \delta_1, \delta_2)$. Then we have

$$\mathbb{P}_n V_{\alpha^2}(\alpha_0, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) \stackrel{p}{\to} PV_{\alpha^2}(\alpha_0, S_1, S_2, \delta_1, \delta_2).$$

Second, we derive the asymptotic distribution of $\sqrt{n}\mathbb{P}_n V_\alpha(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2)$. Note that:

$$\begin{aligned}
&\mathbb{P}_n V_\alpha(\alpha_0, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) \\
= \; &(\mathbb{P}_n - P)\big(V_\alpha(\alpha_0, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) - V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)\big) \\
&+ \mathbb{P}_n V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2) \\
&+ P\big(V_\alpha(\alpha_0, \hat{S}_1(t_1), \hat{S}_2(t_2), \delta_1, \delta_2) - V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)\big) \\
= \; &u_{1n} + u_{2n} + u_{3n}.
\end{aligned}$$

Lemma 2 indicates that under Conditions (A1),(A2), $\mathcal{H}_{\mathcal{F}}$ is a P-Donsker class. Furthermore, since $\sup_{0\leq t_j\leq t_{0j}} |\hat{S}_j(t_j) - S_j(t_j)| \stackrel{p}{\to} 0,\ j = 1, 2$, by the Dominated Convergence Theorem,

$$\int (\hat{S}_j(t_j) - S_j(t_j))^2 dP(t_1, t_2, \delta_1, \delta_2) \stackrel{p}{\to} 0,\ j = 1, 2.$$

Therefore, $\sqrt{n}u_{1n} = o_p(1)$ by van der Vaart (1998), Lemma 19.24 on page 280.

u_{2n} is a sum of independent and identically distributed quantities. Where each quantity has mean:

$$\int V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2) dP(t_1, t_2, \delta_1, \delta_2) = 0$$

and variance:

$$\int \left[V_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)\right]^2 dP(t_1, t_2, \delta_1, \delta_2) = -W(\alpha_0, S_1, S_2, \delta_1, \delta_2).$$

By the Central Limit Theorem, $\sqrt{n}u_{2n}$ converges to a normal random variable with mean 0 and variance $-W(\alpha_0, S_1, S_2, \delta_1, \delta_2)$.

Applying Von Mises Expansion by Mises (1947) on u_{3n} around S_1, S_2, we get

$$u_{3n} \stackrel{d}{=} \int_0^{t_{01}} IC_1(t_1) d(\hat{S}_1 - S_1)(t_1) + \int_0^{t_{02}} IC_2(t_2) d(\hat{S}_2 - S_2)(t_2). \tag{5}$$

Where $\stackrel{d}{=}$ means that both sides have the same asymptotic distribution. (Mises (1947), on page 327.) $IC_j(t)$ is the influence curve of the functional $PV_\alpha(\alpha_0, S_1(t_1), S_2(t_2), \delta_1, \delta_2)$ and have the following form

$$\begin{aligned} IC_1(t_1) &= -\int_0^{t_1}\int_0^{t_{02}} V_{\alpha,u}(\alpha_0, S_1(\tau_1), S_2(\tau_2), \delta_1, \delta_2) dP(\tau_1, \tau_2, \delta_1, \delta_2) \\ &= \int_0^{t_1}\int_0^{t_{02}} M_{\alpha,u}(\alpha_0, S_1(\tau_1), S_2(\tau_2)) f(\tau_1, \tau_2) d\tau_1 d\tau_2 \end{aligned}$$

and

$$\begin{aligned} IC_2(t_2) &= -\int_0^{t_2}\int_0^{t_{01}} V_{\alpha,v}(\alpha_0, S_1(\tau_1), S_2(\tau_2), \delta_1, \delta_2) dP(\tau_1, \tau_2, \delta_1, \delta_2) \\ &= \int_0^{t_2}\int_0^{t_{01}} M_{\alpha,v}(\alpha_0, S_1(\tau_1), S_2(\tau_2)) f(\tau_1, \tau_2) d\tau_1 d\tau_2, \end{aligned}$$

where

$$M_{\alpha,u}(\alpha_0, S_1(\tau_1), S_2(\tau_2)) = -\mathrm{E}_{\delta_1\delta_2|\tau_1\tau_2} V_{\alpha,u}(\alpha_0, S_1(\tau_1), S_2(\tau_2), \delta_1, \delta_2)$$

$$M_{\alpha,v}(\alpha_0, S_1(\tau_1), S_2(\tau_2)) = -\mathrm{E}_{\delta_1\delta_2|\tau_1\tau_2} V_{\alpha,v}(\alpha_0, S_1(\tau_1), S_2(\tau_2), \delta_1, \delta_2).$$

Using the martingale theory for counting process, Pepe (1991) showed that, for $t \in [0, t_{01}]$, $(\hat{S}_1(t_1) - S_1(t_1))$ is asymptotically equivalent to a sum of n i.i.d. random Variables $\sum_i I_1^0(t_{1i}, \delta_{1i})(t_1)/n$. It follows that

$$\int_0^{t_{01}} IC_1(t_1) d(\hat{S}_1 - S_1)(t_1) = \frac{1}{n}\sum_{i=1}^{n} I_1(t_{1i}, \delta_{1i}, \alpha_0), \tag{6}$$

where

$$I_1(t_{1i}, \delta_{1i}, \alpha_0) = \int_0^{t_{01}} \int_0^{t_{02}} M_{\alpha,u}(\alpha_0, S_1(\tau_1), S_2(\tau_2)) f(\tau_1, \tau_2) I_1^0(t_{1i}, \delta_{1i})(\tau_1) d\tau_1 d\tau_2$$

and I_1^0 is a martingale.

Since $M_{\alpha,u}(\alpha_0, S_1(t_1), S_2(t_2)) f(t_1, t_2)$ is a deterministic function, $\mathrm{E} I_1(t_{1i}, \delta_{1i}, \alpha_0) = 0$ for all $i = 1, 2, \cdots, n$. Then $\sqrt{n}\int_0^{t_{01}} IC_1(t_1) d(\hat{S}_1 - S_1)(t_1)$ converges to a normal random variable with mean 0.

On the other hand, $(\hat{S}_2 - S_2)(t_2)$ can not be written as sum of i.i.d random quantities. However, for current status data, the smooth functionals of NPMLE $\hat{F}$ can still be shown to be asymptotically normal at the $n^{1/2}$ rate (Huang & Wellner (1995)). Under conditions (A3)-(A6), Wang & Ding (2000) showed that

$$\begin{aligned} &\int_0^{t_{02}} IC_2(t_2) d(\hat{S}_2 - S_2)(t_2) \\ = \; &-[\frac{1}{n}\sum_{i=1}^{n} \tilde{l}(t_{2i}, \delta_{2i}, S_2, G_2, \psi_2) - E(\tilde{l})] + o_p(1), \end{aligned} \tag{7}$$

with $\tilde{l}(t_2, \delta_2, S_2, G_2, \psi_2) = -[\delta_2 - (1 - S_2(t_2))]\frac{\psi_2(t_2)}{g_2(t_2)} I[g_2(t_2) > 0]$ and thus $\sqrt{n}\int_0^{t_{02}} IC_2(t_2) d(\hat{S}_2 - S_2)(t_2)$ converges to a normal random variable with mean 0.

In summary, we obtain that,

$$\begin{aligned}
& \mathbb{P}_n V_\alpha(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2) \\
= \; & \frac{1}{n}\sum_{i=1}^{n} V_\alpha(\alpha_0, S_1(t_{1i}), S_2(t_{2i}), \delta_{1i}, \delta_{2i}) + \frac{1}{n}\sum_{i=1}^{n} I_1(\alpha_0, t_{1i}, \delta_{1i}) \\
& -\frac{1}{n}\sum_{i=1}^{n}\left[\tilde{l}(t_{2i}, \delta_{2i}, S_2, G_2, \psi_2) - E(\tilde{l})\right] + o_p(n^{-1/2}) \\
= \; & \frac{1}{n}\sum_{i=1}^{n}[Q(\alpha_0, t_{1i}, t_{2i}, \delta_{1i}, \delta_{2i}, S_1, S_2) + E(\tilde{l})] + o_p(n^{-1/2}).
\end{aligned}$$

Therefore, $\sqrt{n}\mathbb{P}_n V_\alpha(\alpha_0, \hat{S}_1, \hat{S}_2, \delta_1, \delta_2)$ is asymptotically normal with mean zero and variance $\mathrm{Var}(Q(\alpha_0, S_1, S_2, t_1, t_2, \delta_1, \delta_2))$. Hence,

$$\sqrt{n}(\hat{\alpha}_n - \alpha_0) \xrightarrow{d} N(0, \sigma^2),$$

where

$$\sigma^2 = \frac{\mathrm{Var}(Q(\alpha_0, S_1, S_2, t_1, t_2, \delta_1, \delta_2))}{W^2(\alpha_0, S_1, S_2, \delta_1, \delta_2)}.$$