Statistica Sinica **21** (2011), 349-367

SEMI-PARAMETRIC INFERENCE FOR COPULA MODELS FOR TRUNCATED DATA

Takeshi Emura, Weijing Wang and Hui-Nien Hung

National Chiao-Tung University

Abstract: We investigate the dependent relationship between two failure time variables that truncate each other. Chaieb, Rivest, and Abdous (2006) proposed a semi-parametric model under the so-called "semi-survival" Archimedean-copula assumption and discussed estimation of the association parameter, the truncation probability, and the marginal functions. Here the same model assumption is adopted but different inference approaches are proposed. For estimating the association parameter, we extend the conditional likelihood approach (Clayton (1978)) and the two-by-two table approach (Wang (2003)) to dependent truncation data. We further show that the three estimators, including that proposed by Chaieb, Rivest, and Abdous (2006), differ in weights. The likelihood approach provides the formula for a good weight. Large sample properties of the proposed methods are established by applying the functional delta method, which can handle estimating functions that are not in the form of U-statistics. Analytic formulae for the asymptotic variance estimators are provided. Two competing methods are compared via simulations, and applied to the transfusion-related AIDS data.

Key words and phrases: Archimedean copula model, conditional likelihood, functional delta method, Kendall's tau, truncation data, two-by-two table.

1. Introduction

Consider a pair of failure times (X, Y) that can be included in a sample only if $X \leq Y$. The variable X is said to be right-truncated by Y, or Y is left-truncated by X. Most literature on truncated data considers marginal analysis by assuming that X and Y are quasi-independent (Tsai (1990)). This assumption might not hold in practice. For example, in the study of transfusion-related AIDS, the incubation time X is right-truncated by the lapse time Y measured from the time of infection to the end of the study (Lagakos, Barraj, and De Gruttola (1988)). Applying Tsai's test, the hypothesis of quasi-independence between X and Y is rejected. This surprising association might be attributed to changes in medical practice in different chronicled periods.

To assess the degree of association between truncated variables, Tsai (1990) modified Kendall's tau by conditioning on an event that guarantees that the chosen pairs are "comparable" under truncation. Tsai's idea was later extended to more complicated truncation structures by Martin and Betensky (2005). Recently Chaieb, Rivest, and Abdous (2006) proposed a "semi-survival" Archimedean copula (AC) model suitable for describing dependent truncation data. They also adopted the idea of pairwise comparison to estimate the association parameter. In addition estimators of the marginal functions and the truncation probability were proposed.

We consider the same model framework as in Chaieb, Rivest, and Abdous (2006), but propose different inference procedures. The paper is organized as follows. In Section 2, we review the related research. The proposed methods for estimating the association parameter and the marginal functions are presented in Section 3. Section 4 contains large sample analysis; proofs are given in the online supplement. In Section 5, the proposed procedure is modified to account for external censoring. Adjustment for analyzing datasets with small sample sizes is also discussed. Numerical results, including simulation studies and data analysis, are presented in Section 6. Concluding remarks are given in Section 7.

2. Preliminary

To simplify the analysis, it is assumed that both X and Y are continuous variables and hence the data contains no ties.

2.1. Association measures and models for typical failure time data

Robust measures are usually preferred in analysis of failure time data. For measuring association, Kendall's tau (τ) is commonly used because of its rank invariance property. Formally, $\tau = \Pr(\Delta_{ij} = 1) - \Pr(\Delta_{ij} = 0) = E(2\Delta_{ij} - 1)$, where $\Delta_{ij} = I\{X_i - X_j)(Y_i - Y_j) > 0\}$ is the concordance indicator for the two pairs (X_i, Y_i) and (X_j, Y_j) that are independent replications of (X, Y). Oakes (1989) proposed the following odds ratio function to describe local association:

$$\begin{aligned}\tilde{\theta}(x,y) &= \frac{\partial^2 \Pr(X > x, Y > y)/\partial x \partial y \cdot \Pr(X > x, Y > y)}{\partial \Pr(X > x, Y > y)/\partial x \cdot \partial \Pr(X > x, Y > y)/\partial y} \\ &= \frac{\Pr(\Delta_{ij} = 1 \mid \tilde{X}_{ij} = x, \tilde{Y}_{ij} = y)}{\Pr(\Delta_{ij} = 0 \mid \tilde{X}_{ij} = x, \tilde{Y}_{ij} = y)},\end{aligned}$$

where $\tilde{X}_{ij} = X_i \wedge X_j$ and $\tilde{Y}_{ij} = Y_i \wedge Y_j$, and where $a \wedge b = \min(a, b)$. Note that the sign of $\log\{\tilde{\theta}(x, y)\}$ indicates the direction of association at time (x, y).

Copula models form a class of bivariate distributions whose marginals are uniform on the unit interval (Genest and Mackay (1986)). In applications, the copula structure is usually imposed on the joint survival function of (X, Y) such that $\Pr(X > x, Y > y) = C_\alpha\{\Pr(X > x), \Pr(Y > y)\}$, where $C_\alpha(u, v)$:

$[0,1]^2 \rightarrow [0,1]$ and the parameter $\alpha \in R$ is related to Kendall's τ by $\tau(\alpha) = 4\int_0^1\int_0^1 C_\alpha(u,v)C_\alpha(du,dv) - 1$.

A useful sub-class of the copula family is the Archimedean copulas (AC) model in which $C_\alpha(u,v)$ is simplified to $C_\alpha(u,v) = \phi_\alpha^{-1}\{\phi_\alpha(u) + \phi_\alpha(v)\}$ for $u, v \in [0,1]$, where $\phi_\alpha(\cdot) : [0,1] \rightarrow [0,\infty]$ is a one-dimensional function satisfying $\phi_\alpha(1) = 0$, $\phi'_\alpha(t) = \partial\phi_\alpha(t)/\partial t < 0$ and $\phi''_\alpha(t) = \partial^2\phi_\alpha(t)/\partial t^2 > 0$. Oakes (1989) showed that, for an AC model indexed by the generating function $\phi_\alpha(\cdot)$, the odds ratio function can be written as $\tilde{\theta}(x,y) = \theta_\alpha\{\Pr(X > x, Y > y)\}$, where $\theta_\alpha(\cdot)$ is a univariate function satisfying

$$\theta_\alpha(v) = -v \cdot \frac{\phi''_\alpha(v)}{\phi'_\alpha(v)}. \tag{2.1}$$

Note that when $\phi_\alpha(t) = -\log(t)$, X and Y are independent and $\theta_\alpha(v) = 1$.

2.2. Association measures and models for truncated data

When (X,Y) are observable only if $X \le Y$, none of the aforementioned descriptive measures is identifiable. Tsai (1990) proposed a modified version of Kendall's tau given by $\tau_a = E(2\Delta_{ij} - 1 \mid A_{ij})$, where $A_{ij} = \{\breve{X}_{ij} \le \tilde{Y}_{ij}\}$, $\breve{X}_{ij} = X_i \vee X_j \equiv \max(X_i, X_j)$, and $\tilde{Y}_{ij} = Y_i \wedge Y_j$. Notice that the event A_{ij} requires $(\breve{X}_{ij}, \tilde{Y}_{ij})$ be located in the identifiable region $R_U = \{(x,y) : 0 \le x \le y < \infty\}$, and hence makes (X_i, Y_i) and (X_j, Y_j) be "comparable" in presence of truncation. Chaieb, Rivest, and Abdous (2006) applied a similar idea to modify the odds ratio function as:

$$\theta^*(x,y) = \frac{\Pr(\Delta_{ij} = 0 \mid \breve{X}_{ij} = x, \tilde{Y}_{ij} = y)}{\Pr(\Delta_{ij} = 1 \mid \breve{X}_{ij} = x, \tilde{Y}_{ij} = y)} \qquad (x \le y) \tag{2.2.a}$$

$$= \frac{\partial^2 \Pr(X \le x, Y > y)/\partial x \partial y \cdot \Pr(X \le x, Y > y)}{\partial \Pr(X \le x, Y > y)/\partial x \cdot \partial \Pr(X \le x, Y > y)/\partial y}. \tag{2.2.b}$$

The interpretation of $\theta^*(x,y)$ is similar to that of $1/\tilde{\theta}(x,y)$, but only the former is identifiable for truncated data. For example, $\theta^*(x,y) < 0$ means positive association.

In light of (2.2.b), Chaieb, Rivest, and Abdous (2006) suggested imposing the copula structure on the "semi-survival" function $\Pr(X \le x, Y > y)$ for $(x,y) \in R_U$. They proposed the semi-survival AC model

$$\pi(x,y) = \{\phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\} + \phi_\alpha\{S_Y(y)\}]\}/c \qquad (x \le y), \tag{2.3}$$

where $\pi(x,y) = \Pr(X \le x, Y > y \mid X \le Y)$, $F_X(\cdot)$, and $S_Y(\cdot)$ are arbitrary continuous and survival functions, respectively, and c is a normalizing constant.

A semi-survival AC model indexed by $\phi_\alpha(\cdot)$ also has the property $\theta^*(x,y) = \theta_\alpha\{c\pi(x,y)\}$, where $\theta_\alpha(v)$ is defined in (2.1).

2.3. Previous inference results for dependent truncation data

By temporarily ignoring the censoring effect, truncation data can be written as $\{(X_j, Y_j)(j = 1, \ldots, n)\}$ subject to $X_j \leq Y_j$. Tsai (1990) proposed to estimate τ_a by

$$\hat{\tau}_a = \frac{\sum_{i<j}(2\Delta_{ij} - 1)I\{A_{ij}\}}{\sum_{i<j} I\{A_{ij}\}}. \tag{2.4}$$

Let $\hat{\pi}(x,y) = \sum_j I(X_j \leq x, Y_j \geq y)/n$ be the empirical estimator of $\pi(x,y)$ for $x < y$. Under (2.3), Chaieb, Rivest, and Abdous (2006) proposed the estimating equation

$$\sum_{i<j} 1\{A_{ij}\}\Big[\Delta_{ij} - \frac{1}{1 + \theta_\alpha\{c\hat{\pi}(\breve{X}_{ij}, \tilde{Y}_{ij})\}}\Big] = 0. \tag{2.5}$$

Equation (2.5) is derived from the equivalent formula

$$\sum_{i<j}(2\Delta_{ij} - 1)I\{A_{ij}\} = \sum_{i<j} \frac{1 - \theta_\alpha\{c\hat{\pi}(\breve{X}_{ij}, \tilde{Y}_{ij})\}}{1 + \theta_\alpha\{c\hat{\pi}(\breve{X}_{ij}, \tilde{Y}_{ij})\}} I\{A_{ij}\},$$

where the two sides of the above equation are related to non- and semi-parametric estimators of τ_α. Notice that estimation of α under truncation also requires information on c. Only the Clayton model, with $\phi_\alpha(t) = (t^{-(\alpha-1)} - 1)/(\alpha - 1)$ $(\alpha > 0)$ and $\theta_\alpha(v) = \alpha$, does (2.5) not depend on c. Under dependent truncation, existing methods for estimating the truncation probability (He and Yang (1998)) are not applicable.

Based on the model at (2.3), Chaieb, Rivest, and Abdous (2006) proposed their second estimating equation for the marginal functions and c:

$$\phi_\alpha\{c\hat{\pi}(t, t-)\} = \phi_\alpha\Big\{c\frac{R(t,t)}{n}\Big\} = \phi_\alpha\{F_X(t)\} + \phi_\alpha\{S_Y(t-)\}, \tag{2.6.a}$$

where t is an observed failure point for X or Y, and $R(x,y) = \sum_j I(X_j \leq x, Y_j \geq y)$. To solve (2.6.a), they made the additional constraint that, for some $x_0 > y_0 > 0$,

$$F_X(x_0) = 1, \; S_Y(x_0) > 0, \; F_X(y_0) > 0, \; S_Y(y_0) = 1. \tag{2.6.b}$$

An idea of Rivest and Wells (2001) was then adopted to solve the equations. Specifically, the jumps $\phi_\alpha\{S_Y(Y_i)\} - \phi_\alpha\{S_Y(Y_i-)\}$ and $\phi_\alpha\{F_X(X_i)\} - \phi_\alpha\{F_X(X_i+)\}$ are calculated first, then summed up to time t, which leads to the estimators of $\phi_\alpha\{F_X(t)\}$ and $\phi_\alpha\{S_Y(t)\}$, respectively. By plugging these marginal estimators back into (2.6.a), an estimating function for c can be obtained.

In light of condition (2.6.b), a convenient choice for (x_0, y_0) is $x_0 = t_{2n-1}$ and $y_0 = t_1$. Accordingly the estimating equation for c can be written as

$$U_c(\alpha, c) = \sum_{j:t_1<x_j} \Big[\phi_\alpha\Big\{c\frac{\tilde{R}(x_j)}{n}\Big\} - \phi_\alpha\Big\{c\frac{\tilde{R}(x_j)-1}{n}\Big\}\Big] + \phi_\alpha\Big(\frac{c}{n}\Big), \tag{2.7}$$

which is obtained by setting $t = y_0 = t_1$ in equation (5.1) of Chaieb, Rivest, and Abdous (2006). The resulting marginal estimators are

$$\phi_\alpha\{\hat{S}_Y(t)\} = -\sum_{j:y_j\le t} \Big[\phi_\alpha\Big\{c\frac{\tilde{R}(y_j)}{n}\Big\} - \phi_\alpha\Big\{c\frac{\tilde{R}(y_j)-1}{n}\Big\}\Big],$$

$$\phi_\alpha\{\hat{F}_X(t)\} = -\sum_{j:t\le x_j} \Big[\phi_\alpha\Big\{c\frac{\tilde{R}(x_j)}{n}\Big\} - \phi_\alpha\Big\{c\frac{\tilde{R}(x_j)-1}{n}\Big\}\Big],$$

where $\tilde{R}(t) = \sum_j I(X_j \le t, Y_j \ge t)$.

3. The Proposed Method for Estimating the Association Parameter

3.1. Motivation

Before introducing the proposed methods, we briefly review the development of semiparametric inference methods for copula models imposed on the survival functions. Early work focused on the Clayton copula based on bivariate right-censored data. In his landmark paper, Clayton (1978) obtained an estimator for the association parameter by maximizing a conditional likelihood function. This estimator was later re-expressed by Clayton and Cuzick (1985) as a weighted version of Oakes' concordance estimator (Oakes (1982)). The new representation is related to a U-statistics and is useful for deriving asymptotic properties (Oakes (1986)).

Inference of copula models has been extended to semi-competing risks data in which one variable is a competing risk for the other, but not versa. Log-rank type estimating functions have been proposed by Day, Bryant, and Lefkopolou (1997) under the Clayton model, and by Wang (2003) for general AC models respectively. Estimating functions using the concordance information of paired observations have been proposed by Fine, Jiang, and Chappell (2001) and Lakhal, Rivest, and Abdous (2008).

3.2. Estimation based on conditional likelihood

Here we generalize the idea of Clayton (1978) to truncation data. Define the set of grid points as follows:

$$\varphi = \Big\{(x, y) \mid x < y, \sum_{j=1}^{n} I(X_j \le x, Y_j = y) = 1, \sum_{j=1}^{n} I(X_j = x, Y_j \ge y) = 1\Big\}.$$

For a point (x,y) in φ, write $\Delta(x,y)=\sum_j I(X_j=x,Y_j=y)$ for the binary variable indicating whether failure occurs at (x,y). Given $R(x,y)=r$ for $(x,y)\in\varphi$, and under model (2.3), $\Delta(x,y)$ is Bernoulli with

$$\Pr\{\Delta(x,y)=1 \mid R(x,y)=r,(x,y)\in\varphi\}=\frac{\theta_\alpha\{c\pi(x,y)\}}{r-1+\theta_\alpha\{c\pi(x,y)\}}.$$

Since $\Pr\{R(x,y)=r \mid (x,y)\in\varphi\}$ may contain little information about α, we suggest using the conditional probability to construct the likelihood function:

$$L(\alpha,c,\pi)=\prod_{(x,y)\in\varphi}\Big[\frac{\theta_\alpha\{c\pi(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\pi(x,y)\}}\Big]^{\Delta(x,y)} \times\Big[\frac{R(x,y)-1}{R(x,y)-1+\theta_\alpha\{c\pi(x,y)\}}\Big]^{1-\Delta(x,y)}.$$

Differentiating $\log\{L(\alpha,c,\hat\pi)\}$ with respect to α, we obtain the estimating function

$$U_L(\alpha,c)=\iint_{(x,y)\in\varphi}\frac{\dot\theta_\alpha\{c\hat\pi(x,y)\}}{\theta_\alpha\{c\hat\pi(x,y)\}}\Big[\Delta(x,y)-\frac{\theta_\alpha\{c\hat\pi(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\pi(x,y)\}}\Big]=0,$$

where $\dot\theta_\alpha=\partial\theta_\alpha(v)/\partial\alpha$. For Clayton's model, $U_L(\alpha,c)$ reduces to

$$U_L(\alpha)=\iint_{(x,y)\in\varphi}\frac{1}{\alpha}\Big[\Delta(x,y)-\frac{\alpha}{R(x,y)-1+\alpha}\Big],$$

which depends only on α. However for other members in the AC family, estimation of α requires the information of c as well. Specifically $\partial\log L(\alpha,c,\pi)/\partial c$ yields the same estimating function as $U_L(\alpha,c)$, since $\theta_\alpha(cv)$ depends on (α,c) only through a single parameter. For example the Frank copula (Genest (1987)), defined as $\phi_\alpha(v)=\log\{(1-\alpha^{-1})/(1-\alpha^{-v})\}$, has the form $\theta_\alpha(cv)=cv\log(\alpha)/(1-e^{-cv\log(\alpha)})$, and it is a function of the single parameter $\gamma=c\log(\alpha)$. This implies that the likelihood function cannot identify α and c simultaneously.

3.3. Estimation based on two-by-two tables

Motivated by ideas of Day, Bryant, and Lefkopolou (1997) and Wang (2003), we can construct a series of 2×2 tables at observed failure points (x,y) with $x\le y$:

	$Y=y$	$Y>y$	
$X=x$	$N_{11}(dx,dy)$		$N_{1\bullet}(dx,y)$
$X<x$			
	$N_{\bullet 1}(x,dy)$		$R(x,y)$

Here cell counts are $N_{11}(dx,dy)=\sum_j I(X_j=x,Y_j=y)$, $N_{\bullet 1}(x,dy)=\sum_j I(X_j\le x,Y_j=y)$, and $N_{1\bullet}(x,dy)=\sum_j I(X_j=x,Y_j\ge y)$. The odds ratio of the above table is related to $\theta^*(x,y)$ defined in (2.2). In absence of ties, all tables of interest have $N_{1\bullet}=N_{\bullet 1}=1$ and N_{11} is Bernoulli. Given the marginal counts, the conditional mean of $N_{11}(dx,dy)$ under model (2.3) is

$$\begin{aligned}&E\{N_{11}(dx,dy)\mid N_{1\bullet},N_{\bullet 1},R)\\ &\quad=\Pr(N_{11}(dx,dy)=1\mid N_{1\bullet}=N_{\bullet 1}=1,R)N_{\bullet 1}(x,dy)N_{1\bullet}(dx,y)\\ &\quad=\frac{\theta_\alpha\{c\pi(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\pi(x,y)\}}N_{\bullet 1}(x,dy)N_{1\bullet}(dx,y).\end{aligned}$$

The resulting log-rank type estimating function can be written as

$$U_w(\alpha,c)=\iint_{x\le y}w_{\alpha,c}(x,y)\Big[N_{11}(dx,dy)-\frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\hat{\pi}(x,y)\}}\Big],\tag{3.1.a}$$

where $w_{\alpha,c}(x,y)$ is a weight function. Note that we have $N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)=1$ if and only if $(x,y)\in\varphi$ and $\Delta(x,y)=N_{11}(dx,dy)$. Accordingly, $U_w(\alpha,c)$ can be written as

$$U_w(\alpha,c)=\iint_{(x,y)\in\varphi}w_{\alpha,c}(x,y)\Big[\Delta(x,y)-\frac{\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\hat{\pi}(x,y)\}}\Big].\tag{3.1.b}$$

3.4. Concordance-type expression

For right-censored data, Clayton's likelihood estimator can be expressed in terms of concordance indicators (Clayton and Cuzick (1985)). We now establish a similar relationship for truncation data. By some algebraic calculations,

$$\begin{aligned}U_w(\alpha,c)&=\iint_{(x,y)\in\varphi}w_{\alpha,c}(x,y)\Big[\Delta(x,y)-\frac{\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\hat{\pi}(x,y)\}}\Big]\\ &=-\sum_{i<j}1\{A_{ij}\}\frac{w_{a,c}(\breve{X}_{ij},\tilde{Y}_{ij})[1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}]}{R(\breve{X}_{ij},\tilde{Y}_{ij})-1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}}\Big[\Delta_{ij}-\frac{1}{1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}}\Big].\end{aligned}\tag{3.2.a}$$

The proof is given in Appendix B of the on-line supplement in which right-censoring is also considered. From (3.2.a), the estimating function proposed by Chaieb, Rivest, and Abdous (2006) can be generalized to

$$\tilde{U}_w(\alpha,c)=\sum_{i<j}1\{A_{ij}\}\tilde{w}_{\alpha,c}(\breve{X}_{ij},\tilde{Y}_{ij})[\Delta_{ij}-\frac{1}{1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}}\Big].\tag{3.2.b}$$

3.5. Interpretation of the suggested weight in $U_L(\alpha, c)$

We have demonstrated that the two estimating functions $U_w(\alpha, c)$ and $\tilde{U}_w(\alpha, c)$ with flexible weight functions are equivalent. Generally speaking, the first-moment condition alone does not provide enough information if efficiency is also of concern. What matters is the choice of weights. Our proposal $U_L(\alpha, c)$, which is derived from the conditional likelihood, corresponds to the specific weights

$$w_{\alpha,c}(x,y) = \frac{\dot{\theta}_\alpha\{c\hat{\pi}(x,y)\}}{\theta_\alpha\{c\hat{\pi}(x,y)\}},$$
$$\tilde{w}_{\alpha,c}(x,y) = -\frac{\dot{\theta}_\alpha\{c\hat{\pi}(x,y)\}}{\theta_\alpha\{c\hat{\pi}(x,y)\}} \frac{1+\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\hat{\pi}(x,y)\}},$$

for (3.1.b) and (3.2.b), respectively.

Alternatively we may adopt the generalized estimating equations (GEE) for choosing the weights. Conditional on the marginal counts, the mean and variance of $N_{11}(dx, dy)$ are

$$\mu_{(\alpha,c)}(x,y) = \frac{\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\hat{\pi}(x,y)\}},$$
$$V_{(\alpha,c)}(x,y) = \frac{\theta_\alpha\{c\hat{\pi}(x,y)\}\{R(x,y)-1\}}{[R(x,y)-1+\theta_\alpha\{c\hat{\pi}(x,y)\}]^2}.$$

It can be shown that the recommended weights can be written as

$$\frac{\dot{\theta}_\alpha\{c\pi(x,y)\}}{\theta_\alpha\{c\pi(x,y)\}} = \Big\{\frac{\partial}{\partial\alpha}\mu_{(\alpha,c)}(x,y)\Big\} V_{(\alpha,c)}(x,y)^{-1},$$

where the right side is the suggestion by GEE under the independence working assumption (Liang and Zeger (1986)).

We examine two AC models for the sake of illustration. For the Clayton model, the measure of association $1/\theta_\alpha(v) = 1/\alpha$ does not depend on v. It follows that $\dot{\theta}_\alpha(v)/\theta_\alpha(v) = 1/\alpha$, which implies that equal weights are assigned to all observed points. For the Frank model, it follows that $1/\theta_\alpha(v) = (1 - \alpha^{-v})/\{v\log(\alpha)\}$ and $\dot{\theta}_\alpha(v)/\theta_\alpha(v) = 1/\{\alpha\log(\alpha)\} - v\alpha^{-v-1}/(1-\alpha^{-v})$, both of which depend on v. Figure 1 depicts the two functions under the Frank model with three patterns of association (positive, no, and negative). We see that the suggested weight $\dot{\theta}_\alpha(v)/\theta_\alpha(v)$ gets large as the level of association $1/\theta_\alpha(v)$ increases. Under the degenerate case of independence, the suggested weight for the Clayton copula is the un-weighted version of $U_w(\alpha, c)$, while that for the Frank copula becomes

$$\lim_{\alpha\to 1} \frac{\dot{\theta}_\alpha\{c\hat{\pi}(x,y)\}}{\theta_\alpha\{c\hat{\pi}(x,y)\}} = \frac{c\hat{\pi}(x,y)}{2} \propto R(x,y),$$

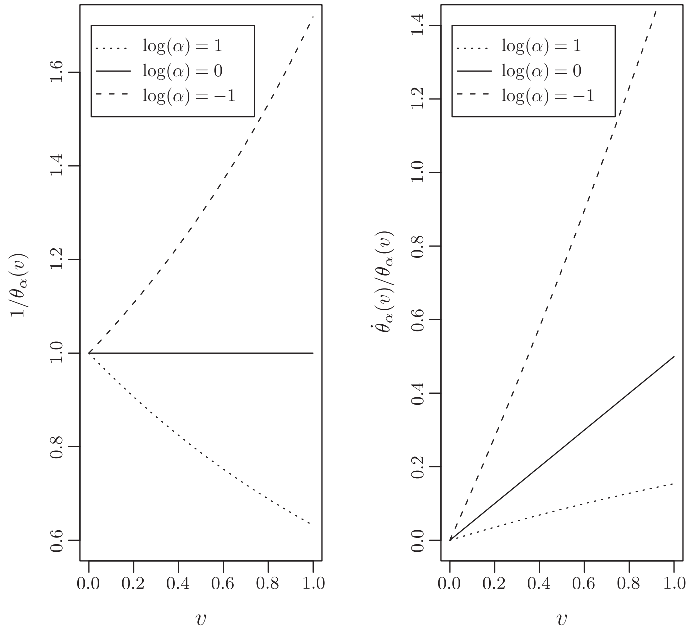

Figure 1. The odds ratio function $1/\theta_\alpha(v) = (1-\alpha^{-v})/\{v\log(\alpha)\}$ and the suggested weight function $\dot\theta_\alpha(v)/\theta_\alpha(v) = 1/(\alpha\log(\alpha)) - v\alpha^{-v-1}/(1-\alpha^{-v})$ under the Frank model. Three parameter values $\log(\alpha)=1$, $\log(\alpha)=0$, and $\log(\alpha)=-1$ correspond to negative association, no association, and positive association on (X,Y), respectively.

which is "the number at risk" for truncated data. This difference between the two weights is somewhat analogous to that between the log-rank and Gehan statistics in the classical problem of two-sample comparison.

4. Asymptotic Analysis

4.1. Asymptotic theory

Large sample analysis of $\tilde U_w(\alpha,c)$ can be conducted based on properties of U-statistics if $\tilde w_{\alpha,c}(x,y)$ is deterministic (Chaieb, Rivest, and Abdous (2006)), but this approach is not applicable here since the weight function involves the plugged-in estimator $\hat\pi(x,y)$. We apply the functional delta method (Van Der Vaart (1998, Thm. 20.8)) and properties of empirical processes for large-sample analysis. Under the regularity conditions stated in the on-line supplement, we obtain the asymptotic linear representation

$$\sqrt{n}\begin{bmatrix}\hat\alpha-\alpha_0\\ \hat c - c_0\end{bmatrix} = \frac{1}{n^{1/2}}\sum_{i=1}^{n} A^{-1}U_{\alpha_0,c_0}(X_i,Y_i)+o_p(1), \tag{4.1}$$

where A and $U_{\alpha_0,c_0}(\cdot,\cdot)$ are defined in the on-line supplement. Equation (4.1) also provides the basis for variance estimation, which is discussed in Section 4.2.

The first two theorems state that the proposed estimators $(\hat{\alpha},\hat{c})$, which jointly solve $\tilde{U}_w(\alpha,c)=0$ in (3.2.b) and $U_c(\alpha,c)=0$ in (2.7), are consistent and asymptotically normal. Theorem 3 establishes weak convergence of the proposed marginal estimators. The proofs are given in the on-line supplement. Let (α_0,c_0) be the true values of (α,c).

Theorem 1. *$\hat{\alpha}\to\alpha_0$ and $\hat{c}\to c_0$ in probability.*

Theorem 2. *$n^{1/2}(\hat{\alpha}-\alpha_0,\hat{c}\to c_0)^T$ converges in distribution to a mean-zero bivairate normal with covariance matirix $A^{-1}B(A^{-1})^T$, where A is defined in* (4.1) *and $B=E[U_{\alpha_0,c_0}(X_i,Y_i)U_{\alpha_0,c_0}(X_i,Y_i)^T]$.*

Theorem 3. *The stochastic process $n^{1/2}(\hat{S}_Y(t)-S_Y(t),\hat{F}_X(t)-F_X(t))^T$ convergences weakly to the mean-zero Gaussian random field $G(t)=(G_X(t),G_Y(t))^T$ in the space $\{D[0,\infty)\}^2$, with the covariance function given at* (A.4) *of the on-line supplement.*

4.2. Asymptotic variance estimator

From Theorem 2, the asymptotic variance of $n^{1/2}(\hat{\alpha}-\alpha_0,\hat{c}\to c_0)^T$ is $A^{-1}B$ $(A^{-1})^T$, and can be estimated by plugging in $\hat{A}_{\hat{\alpha},\hat{c}}$ and $\hat{B}_{\hat{\alpha},\hat{c}}$ defined in Appendix A3 of the on-line supplement. Here we derive $\hat{A}_{\hat{\alpha},\hat{c}}$ and $\hat{B}_{\hat{\alpha},\hat{c}}$ under the Clayton model with $\phi_\alpha(t)=(t^{-(\alpha-1)}-1)/(\alpha-1)$ $(\alpha>1)$.

(1) For $\tilde{w}_{\alpha,c}(x,y)=1$ (the estimator of Chaieb, Rivest, and Abdous (2006)):

$$\hat{A}_\alpha=\frac{1}{(1+\alpha)^2n^2}\sum_{k,l}I\{A_{kl}\},$$

$$\hat{B}_\alpha=\frac{1}{n}\sum_i\left(\frac{2}{n}\sum_l I\{A_{il}\}[\Delta_{il}-\frac{1}{1+\alpha}]-\frac{2}{n^2}\sum_{k,l}I\{A_{kl}\}[\Delta_{kl}-\frac{1}{1+\alpha}]\right)^2.$$

(2) For $\tilde{w}_{\alpha,c}(x,y)=-(1+\alpha)/\{R(x,y)-1+\alpha\}$ (the proposed estimator):

$$\hat{A}_\alpha=-\frac{1}{\alpha^2n^2}\sum_{k,l}I\{A_{kl}\}\frac{\Delta_{kl}-1/(1+\alpha)}{\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl})}+\frac{1}{\alpha(1+\alpha)n^2}\sum_{k,l}\frac{I\{A_{kl}\}}{\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl})},$$

$$\hat{B}_\alpha=\frac{1}{n}\sum_i\Bigg(-\frac{1+\alpha}{\alpha n^2}\sum_{k,l}I\{A_{kl}\}\frac{[\Delta_{kl}-1/(1+\alpha)]}{\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl})^2}$$
$$\times\{I(X_i\le\breve{X}_{kl},Y_i\ge\tilde{Y}_{kl})-\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl})\}$$

$$+\frac{2(1+\alpha)}{\alpha n}\sum_{l} I\{A_{il}\}\frac{[\Delta_{il}-1/(1+\alpha)]}{\hat{\pi}(\breve{X}_{il},\tilde{Y}_{il})}$$
$$-\frac{2(1+\alpha)}{\alpha n^2}\sum_{k,l} I\{A_{kl}\}\frac{[\Delta_{kl}-1/(1+\alpha)]}{\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl})}\Bigg)^2.$$

The variance for the two estimators of α under the Clayton model is

$$\mathrm{Var}\,(n^{1/2}\hat{\alpha}) \approx \hat{A}_{\hat{\alpha}}^{-2}\hat{B}_{\hat{\alpha}}. \tag{4.2}$$

For other AC models, explicit derivation of the variance estimator becomes very complicated. We suggest using the jackknife method for variance estimation as in Chaieb, Rivest, and Abdous (2006). This is a convenient tool that can be applied to all members of AC models, even in the presence of censoring. Specifically, the variance of $\hat{\theta}=(\hat{\alpha},\hat{c})^T$ can be estimated by

$$\frac{n-1}{n}\sum_{j}(\hat{\theta}^{(-j)}-\overline{\theta}^{(\cdot)})(\hat{\theta}^{(-j)}-\overline{\theta}^{(\cdot)})^T,$$

where $\hat{\theta}^{(-j)}$ is the estimator after deleting the jth observation and $\bar{\theta}^{(\cdot)}=\sum_j \hat{\theta}^{(-j)}/n$.

In the simulation studies provided in the on-line supplement, we compare the performance of the estimator at (4.2) and the jackknife estimator under the Clayton model. Both variance estimators have reasonable performances in finite samples. The analytic estimator has smaller mean squared error (MSE) than the jackknife estimator in all cases, but the difference is modest. Confidence intervals using the two variances with the normal approximation have similar coverage probabilities that are close to the nominal level in all cases.

5. Extension and Modification

5.1. Extension under right censoring

Consider the situation in which Y is subject to left-truncation by X and right- censoring by another variable C. Assume that C is independent of (X,Y). The sample can be written as $\{(X_i,Z_i,\delta_i)(i=1,\ldots,n)\}$ satisfying $X_i\le Z_i$, where $Z_i=Y_i\wedge C_i$, $\delta_i=I(Y_i\le C_i)$, and $(X_i,Y_i,C_i)(i=1,\ldots,n)$ are random replications of (X,Y,C).

Chaieb, Rivest, and Abdous (2006) expressed the semi-survival AC model as

$$\pi^*(x,y)=\Pr(X\le x, Z>y \mid X\le Z)=S_C(y)\frac{\phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\}+\phi_\alpha\{S_Y(y)\}]}{c^*},$$

where $S_C(y)=\Pr(C>y)$, $x\le y$, and c^* is a normalizing constant. The objective is to estimate the unknown parameters $(\alpha,c^*,F_X(\cdot),S_Y(\cdot),S_C(\cdot))$. We re-parameterize $\theta_\alpha\{c\pi(x,y)\}$ as $\theta_\alpha\{c^*v(x,y)\}$, where $c\pi(x,y) = c^*v(x,y)$ and $v(x,y) = \pi^*(x,y)/S_C(y)$.

Now we modify the first estimating function based on conditional likelihood estimation. To simplify the presentation, we use the same notation but change definitions. To read

$$\varphi = \Big\{(x,y) \mid x \le y, \sum_j I(X_j \le x, Z_j = y, \delta_j = 1) = 1,$$
$$\sum_j I(X_j = x, Z_j \ge y) = 1\Big\},$$
$$\Delta(x,y) = \sum_j I(X_j = x, Z_j = y, \delta_j = 1),\ R(x,y) = \sum_j I(X_j \le x, Z_j \ge y).$$

The proposed estimating function is then given by

$$U_L(\alpha, c^*) = \iint_{(x,y)\in\varphi} \frac{\dot{\theta}_\alpha\{c^*\hat{v}(x,y)\}}{\theta_\alpha\{c^*\hat{v}(x,y)\}}\Big[\Delta(x,y) - \frac{\theta_\alpha\{c^*\hat{v}(x,y)\}}{R(x,y) - 1 + \theta_\alpha\{c^*\hat{v}(x,y)\}}\Big], \tag{5.1}$$

where $\hat{v}(x,y) = R(x,y)/\{n\hat{S}_C(y)\}$ and $\hat{S}_C(y)$ is an estimator of $S_C(y)$. In the presence of censoring, $U_L(\alpha, c^*)$ can also be expressed in terms of pairwise concordance indicators as in (3.2.a), refer to Appendix B of the on-line supplement for details. Under independent censorship, $S_C(y)$ can be estimated by the Lynden-Bell estimator

$$\hat{S}_C(y) = \prod_{j:Z_j \le y, \delta_j = 0} \{1 - \frac{1}{R(Z_j, Z_j)}\}.$$

The estimating equation $U_L(\alpha, c^*) = 0$ can be solved jointly with the second estimating equation proposed by Chaieb, Rivest, and Abdous (2006). In Appendix C of the on-line supplement, we derive explicit forms for these estimating functions for selected examples of $\phi_\alpha(t)$.

5.2. Modification for small risk sets

The proposed estimation procedure, as well as that proposed by Chaieb, Rivest, and Abdous (2006), are based on the implicit assumption that $R(t_j, t_j+) \ge 1$ for all t_j. However it sometimes happens that an empty risk set occurs especially in the tail area. Several remedies have been proposed (Klein and Moeschberger (2003, p.122)). Here we adopt the idea of Lai and Ying (1991) and propose the modification

$$\phi_\alpha\{\hat{S}_Y(t)\} = -\sum_{j:z_j \le t, \delta_j = 1} \Big[\phi_\alpha\Big\{c^*\frac{\tilde{R}(z_j)}{n\hat{S}_c(z_j)}\Big\} - \phi_\alpha\Big\{c^*\frac{\tilde{R}(z_j) - 1}{n\hat{S}_c(z_j)}\Big\}\Big] I\{\tilde{R}(z_j) \ge bn^a\},$$

Table 1A. Comparison of two estimators for the association parameter under the Clayton model.

(c, c^*)	$-\log(\alpha)$	$n = 250$		$n = 500$	
	(τ)	Proposed	Chaieb et al.	Proposed	Chaieb et al.
(0.55, 0.34)	0.5108	3.6 (0.37)	7.8 (0.98)	-2.2 (0.19)	-0.3 (0.51)
(0.55, 0.39)	(0.25)	0.5 (0.35)	5.2 (0.80)	0.3 (0.14)	1.2 (0.40)
(0.66, 0.45)		1.6 (0.44)	6.1 (1.04)	-0.6 (0.19)	-1.3 (0.49)
(0.66, 0.53)		0.7 (0.29)	0.5 (0.74)	-1.2 (0.12)	-1.5 (0.38)
(0.80, 0.63)		-0.9 (0.44)	1.5 (1.13)	-0.3 (0.18)	2.9 (0.53)
(0.80, 0.80)		0.3 (0.17)	5.4 (0.53)	2.5 (0.08)	3.7 (0.26)
(0.63, 0.36)	1.0986	3.0 (0.44)	7.1 (1.09)	-0.3 (0.20)	-0.5 (0.52)
(0.63, 0.42)	(0.5)	-3.5 (0.44)	-3.2 (0.95)	-1.5 (0.18)	2.7 (0.49)
(0.74, 0.48)		-5.4 (0.52)	-5.5 (1.27)	-0.1 (0.23)	0.2 (0.62)
(0.74, 0.58)		-0.2 (0.20)	4.0 (0.54)	-4.6 (0.18)	-2.9 (0.47)
(0.86, 0.66)		2.5 (0.56)	6.2 (1.44)	-0.7 (0.24)	0.2 (0.74)
(0.86, 0.86)		-6.7 (0.28)	-2.1 (0.86)	0.6 (0.13)	0.6 (0.38)

Each cell contains the bias ($\times 10^{-3}$) and MSE ($\times 10^{-2}$) (in parenthesis) of the corresponding estimator based on 500 replications.

where $0 < a < 1$ and $b > 0$ are arbitrary tuning parameters. Modifications for $\phi_\alpha\{\hat{F}_X(t)\}$ and $\hat{S}_C(t)$ are obtained in a similar way. In simulations not shown here, we found that taking $b = 1$ and $a = 1/10$ produced less biased results. Under the Clayton and Frank models with negative association ($\alpha > 1$), we found some simulated datasets for which the estimating equation for c^* does not have zero. This problem can be alleviated by the same type of modification to the estimating equation.

6. Numerical Analysis

6.1. Simulation studies

The main purpose of the simulation studies was (i) to check the validity of the proposed estimators, and (ii) to compare the performance of the proposed method with that of Chaieb, Rivest, and Abdous (2006). Random replications of (X, Y) were generated from the Clayton and Frank models with exponential marginal distributions subject to $X \leq Y$. For the Clayton model, the values of $-\log(\alpha)$ were chosen to be 0.511 and 1.099 and, for the Frank model, the values of $-\log(\alpha)$ were set to 2.380 and 5.746. The former setting corresponds to $\pi = 0.25$ and the latter corresponds to $\tau = 0.5$. The censoring variable C followed an exponential distribution. With $c = \Pr(X \leq Y)$ and $c^* = \Pr(X \leq Y \wedge C)$. We report the bias and MSE based on 500 replications.

Tables 1A and 1B summarize the results for the two competing estimators of α under the Clayton model and Frank model, respectively. The results in Table

Table 1B. Comparison of two estimators for the association parameter under the Frank model.

(c, c^*)	$-\log(\alpha)$	$n = 250$		$n = 500$	
	(τ)	Proposed	Chaieb et al.	Proposed	Chaieb et al.
(0.50, 0.31)	2.380	-371.9 (216.14)	-360.6 (241.92)	-243.3 (131.45)	-257.1 (151.72)
(0.50, 0.36)	(0.25)	-294.2 (201.13)	-342.5 (239.21)	-140.5 (94.01)	-141.8 (96.03)
(0.63, 0.43)		-102.2 (100.42)	-106.3 (116.84)	-99.2 (51.85)	-88.4 (59.76)
(0.63, 0.51)		-162.9 (95.06)	-156.4 (102.07)	-35.9 (44.27)	-34.6 (48.62)
(0.81, 0.63)		-53.5 (52.87)	-36.0 (55.31)	-19.4 (26.87)	-24.5 (28.49)
(0.81, 0.81)		-68.6 (37.55)	-68.1 (37.62)	-26.1 (20.53)	-25.8 (20.53)
(0.50, 0.29)	5.746	-429.6 (293.47)	-411.3 (332.07)	-373.9 (130.48)	-349.2 (146.83)
(0.50, 0.34)	(0.5)	-367.3 (223.11)	-368.2 (247.30)	-246.1 (115.74)	-252.5 (129.13)
(0.69, 0.44)		-182.2 (129.78)	-155.9 (147.13)	-0.1259 (68.20)	-96.1 (73.32)
(0.69, 0.53)		-136.9 (100.55)	-142.4 (104.41)	-114.0 (49.40)	-100.0 (51.47)
(0.88, 0.66)		-57.0 (64.71)	-21.4 (72.73)	-78.0 (33.97)	-78.0 (36.76)
(0.88, 0.88)		-128.3 (41.53)	-128.2 (41.56)	-27.8 (21.91)	-27.5 (22.01)

Each cell contains the bias ($\times 10^{-3}$) and MSE ($\times 10^{-2}$) (in parenthesis) of the corresponding estimator based on 500 replications.

1A show that the proposed estimator has significantly better performance under Clayton's model, reducing the MSE to half or one third in all cases. On the other hand, Table 1B shows that the gain in efficiency is modest under Frank's model. It is interesting to note that, under the Frank model in the absence of censoring ($c = c^*$), the two estimators produce similar results. In this case, we found that the numerical value of

$$\frac{\dot{\theta}_\alpha\{c\hat{\pi}(x,y)\}}{\theta_\alpha\{c\hat{\pi}(x,y)\}} \frac{1+\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\pi(x,y)\}} \tag{6.1}$$

did not vary much at different values of (x, y), which explains the similarity in MSE. Nevertheless when the censoring proportion increased, the value of at (6.1) depended more on (x, y) and the proposed estimator tended to have better performance. Notice that the MSE of both estimators was larger under the Frank model (Table 1B) than those under the Clayton model (Table 1A). From (5.1), we see that the suggested weight under the Clayton model (see $U_L(\alpha)$ in Appendix C1 of the on-line supplement) does not contain any nuisance parameter. On the other hand, the weight under the Frank model (see $U_L(\gamma)$ in Appendix C2 of the on-line supplement) involves estimation of the nuisance parameter $v(x, y)$.

Table 2 details the performance of the marginal estimators proposed by Chaieb, Rivest, and Abdous (2006) under the Clayton and Frank models, where $P_{CEN} = \Pr(C < Y \mid X \leq Z)$ denotes the proportion of censoring in the sample. Their estimating equations were calculated jointly with our proposal

Table 2. The proposed estimators of marginal functions and truncation proportion.

parameter	True	$n=250$ $P_{CEN}=0.00$	$n=250$ $P_{CEN}=0.41$	$n=500$ $P_{CEN}=0.00$	$n=500$ $P_{CEN}=0.41$
Clayton model					
c/c^*	0.86/0.66	2.0 (0.04)	1.3 (0.27)	2.0 (0.02)	0.4 (0.15)
$F_X(t_1)$	0.2	-1.6 (0.06)	-1.7 (0.05)	-0.2 (0.03)	-1.1 (0.03)
$F_X(t_2)$	0.4	-2.0 (0.08)	-2.1 (0.11)	-0.3 (0.04)	-1.1 (0.06)
$F_X(t_3)$	0.6	-0.2 (0.09)	-2.3 (0.15)	0.7 (0.05)	-1.9 (0.08)
$F_X(t_4)$	0.8	-1.2 (0.07)	0.7 (0.16)	1.9 (0.04)	-0.9 (0.07)
$S_Y(t_1)$	0.8	0.8 (0.09)	0.1 (0.08)	0.0 (0.04)	0.2 (0.04)
$S_Y(t_2)$	0.6	1.5 (0.10)	-0.3 (0.12)	-0.7 (0.05)	0.9 (0.06)
$S_Y(t_3)$	0.4	1.5 (0.08)	0.4 (0.14)	-1.1 (0.04)	0.2 (0.06)
$S_Y(t_4)$	0.2	-0.6 (0.06)	-1.3 (0.15)	-0.1 (0.03)	-0.2(0.07)
Frank model					
c/c^*	0.88/0.66	-0.9 (0.11)	-4.1 (0.34)	-1.7 (0.06)	1.0 (0.15)
$F_X(t_1)$	0.2	-0.1 (0.08)	-3.4 (0.07)	-1.8 (0.04)	-1.8 (0.04)
$F_X(t_2)$	0.4	-3.5 (0.11)	-3.6 (0.12)	-1.9 (0.05)	-2.1 (0.06)
$F_X(t_3)$	0.6	-3.7 (0.09)	-0.8 (0.14)	-0.8 (0.05)	-1.1 (0.07)
$F_X(t_4)$	0.8	-1.5 (0.06)	-1.9 (0.12)	-0.7 (0.03)	-0.9 (0.06)
$S_Y(t_1)$	0.8	-0.4 (0.11)	-4.7 (0.13)	-1.6 (0.07)	-2.4 (0.06)
$S_Y(t_2)$	0.6	-1.1 (0.11)	-3.9 (0.14)	-0.6 (0.06)	-1.7 (0.07)
$S_Y(t_3)$	0.4	-0.4 (0.10)	-2.6 (0.16)	-1.5 (0.05)	-2.1 (0.07)
$S_Y(t_4)$	0.2	-1.0 (0.06)	-1.9 (0.15)	-1.0 (0.03)	-1.8 (0.07)

Each cell contains the bias ($\times 10^{-3}$) and MSE ($\times 10^{-2}$) (in parenthesis) based on the recursive estimator using the likelihood method for the association parameter. The censoring proportion is denoted by $P_{CEN} = \Pr(C < Y \mid X \le Z)$.

$U_L(\alpha, c^*) = 0$. The estimators for distribution/survival functions were evaluated at selected points of t with $F_X(t) = 0.2, 0.4, 0.6, 0.8$ and $S_Y(t) = 0.2, 0.4, 0.6, 0.8$. In all cases, $(c^*, \hat{F}_X(\cdot), \hat{S}_Y(\cdot))$ were fairly unbiased. The MSE grew smaller as the sample size increased and the censoring rate (P_{CEN}) decreased. The performances under the Clayton model were very similar to, or slightly better than, those for the Frank models.

It is worth noting that the estimated distribution/survival functions had better performance in the tail area but poorer performance in the middle, which differs from that of the usual Kaplan-Meier estimator. The high accuracy in the tails may be due to the fact that the constraints $\hat{F}(t_{2n-1}) = 1$ and $\hat{S}_Y(t_1) = 1$ were close to the true condition in the simulation setting.

6.2. Data analysis

We apply the inference procedures to the transfusion-related AIDS data dis-

Table 3. Analysis of the transfusion-related AIDS data.

	Proposed		Chaieb et al.	
Assumption: Clayton copula	Estimates	95% jackknife interval	Estimates	95% jackknife interval
$-\log(\alpha)$	0.203 ($\tau = 0.101$)	(0.112, 0.295)	0.195 ($\tau = 0.097$)	(0.065, 0.326)
c	0.336	(0.201, 0.472)	0.329	(0.176, 0.483)
Wald's chi-squarefor	19.173		8.562	
for $H_0 : \log(\alpha) = 0$	(p-value<0.001)		(p-value<0.003)	
Assumption: Frank copula	Estimates	95% jackknife interval	Estimates	95% jackknife interval
$-\log(\alpha)$	3.752 ($\tau = 0.369$)	(2.272, 5.232)	3.736 ($\tau = 0.368$)	(2.256, 5.215)
c	0.543	(0.356, 0.729)	0.541	(0.354, 0.7271)
Wald's chi-squarefor	24.696		24.495	
for $H_0 : \log(\alpha) = 0$	(p-value<0.001)		(p-value<0.001)	

cussed in Kalbfleisch and Lawless (1989). Let T be the infection time, measured from January 1, 1978, and X be the incubation time from the time of infection to AIDS. Only individuals who developed AIDS by the starting date, July 1, 1986, could be observed. Since the total study period was 102 months, individuals with $T + X \leq 102$ were included in the sample that consisted of 293 subjects. Setting $Y = 102 - T$, the incubation time X is right truncated by Y. Note that there was no external censoring.

We analyzed the data under two different model assumptions. The data contains many ties in both X and Y. We break these ties by adding uniform random variables on $[-0.4, 0.4]$ which does not change the original ordering. The results are summarized in Table 3. Under the Clayton model, both estimators of α showed positive correlation between X and Y: patients infected earlier tended to have longer incubation time. The dependence of the incubation period on the calendar time of infection may be a cohort effect due to changes in personal lifestyle or medical treatment. Both estimators rejected the null hypothesis of quasi-independence: $H_0 : \alpha = 1$. This conclusion also follows from Tsai's nonparametric test (1990). The confidence intervals for $-\log(\alpha)$ and c, based on the proposed likelihood estimator, are narrower than those obtained by the estimator of Chaieb, Rivest, and Abdous (2006). The level of association between X and Y appeared stronger when the Frank model was assumed. As in the simulations (Table 1B with $c = c^*$), the two estimators produced similar results under the Frank model.

Figure 2 depicts the estimated incubation distributions under the two model assumptions, applying the marginal estimators of Chaieb, Rivest, and Abdous (2006) together with the proposed estimators for the association parameter. The estimated curve under the Clayton model is significantly lower than that under Frank's model; the marginal estimators are also sensitive to the model choice.

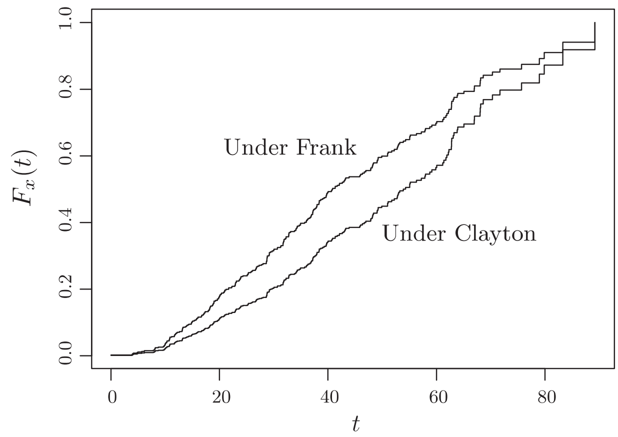

Figure 2. The cumulative distribution functions of the incubation time of AIDS under two copula models.

7. Conclusion

Chaieb, Rivest, and Abdous (2006) proposed semi-survival AC models for describing the joint behavior of two dependent variables that truncate each other, and suggested semiparametric inference methods. Here we proposed different approaches for parameter estimation and large-sample analysis.

For estimating the association parameter, our proposed log-rank statistics have an equivalent expression as the pairwise concordance statistics proposed by Chaieb, Rivest, and Abdous (2006). It is essential to choose a weight function that leads to a more efficient estimator. For semi-competing risks data, Fine, Jiang, and Chappell (2001) suggested some guidelines for selecting the weight function but did not provide any theoretical justification. Here the suggested weight function was originally derived from a conditional likelihood function, but can also be obtained by applying the principle of GEE. Simulation analysis confirmed that the suggested weight does produce more efficient results.

To establish large-sample properties, we used the functional delta method to handle estimating functions with flexible weights. The variance estimator at (4.2) can be viewed as a generalization of the results in Oakes (1982, 1986) under bivariate censored data. Alternatively, the jackknife estimator provides a convenient option for variance estimation.

More studies on the conditional likelihood approach are warranted. Compared to moment-type estimators, this approach has the advantage that it can handle the situation when the dimension of α exceeds one. Also, it may be interesting to investigate the efficiency loss in assuming the working assumption of independence among the different 2×2 tables. Extension of the approach under more complicated truncation settings, as in Martin and Betensky (2005), deserves further investigation.

Acknowledgements

This work was financially supported by the last author's research grant funded by the National Science Council of Taiwan.

References

Chaieb, L. L., Rivest, L.-P. and Abdous, B. (2006). Estimating survival under a dependent truncation. *Biometrika* **93**, 655-669.

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

Clayton, D. G. and Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model (with discussion). *J. Roy. Statist. Soc. Ser. A* **148**, 82-117.

Day, R., Bryant, J. and Lefkopolou, M. (1997). Adaptation of bivariate frailty models for prediction, with application to biological markers as prognostic indicators. *Biometrika* **84**, 45-56.

Fine, J. P., Jiang, H. and Chappell, R. (2001). On semi-competing risks data. *Biometrika* **88**, 907-919.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549-555.

Genest, C. and Mackay, R. J. (1986). The joy of Copulas: Bivariate distributions with uniform marginals. *Amer. Statist.* **40**, 280-283.

He, S. and Yang, G. L. (1998). Estimation of the truncation probability in the random truncation model. *Ann. Statist.* **26**, 1011-1027.

Kalbfleisch, J. D. and Lawless, J. F. (1989). Inference based on retrospective ascertainment: an analysis of the data on transfusion-related AIDS. *J. Amer. Statist. Assoc.* **84**, 360-372.

Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data.* Springer, New-York.

Lagakos, S. W., Barraj, L. M. and De Gruttola, V. (1988). Non-parametric analysis of truncated survival data, with application to AIDS. *Biometrika* **75**, 515-23.

Lai, T. L. and Ying, Z. (1991). Estimating a distribution function with truncated and censored data. *Ann. Statist.* **19**, 417-442.

Lakhal, L., Rivest, L.-P. and Abdous, B. (2008). Estimating survival and association in a semicompeting risks model. *Biometrics* **64**, 180-188.

Liang, K.-Y. and Zeger, S. L.(1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

Martin, E. C. and Betensky, R. A. (2005). Testing quasi-independence of failure and truncation via Conditional Kendall's Tau. *J. Amer. Statist. Assoc.* **100**, 484-92.

Oakes, D. (1982). A model for association in bivariate survival data. *J. Roy. Statist. Soc. Ser. B.* **44**, 414-22.

Oakes, D. (1986). Semi-parametric inference in a model for association in bivariate survival data. *Biometrika* **73**, 353-361.

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

Rivest, L.-P. and Wells, M. T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *J. Multivariate Anal.* **79**, 138-155.

Tsai, W. -Y. (1990). Testing the assumption of independence of truncation time and failure time. *Biometrika* **77**, 169-177.

Van Der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

Wang, W. (2003). Estimating the association parameter for copula models under dependent censoring. *J. Roy. Statist. Soc. Ser. B* **65**, 257-73.

Institute of Statistics, National Chiao-Tung University, Hsin-Chu, Taiwan, R.O.C.

E-mail: takeshi@stat.sinica.edu.tw

Institute of Statistics, National Chiao-Tung University, Hsin-Chu, Taiwan, R.O.C.

E-mail: wjwang@stat.nctu.edu.tw

Institute of Statistics, National Chiao-Tung University, Hsin-Chu, Taiwan, R.O.C.

E-mail: hhung@stat.nctu.edu.tw

(Received October 2008; accepted August 2009)