Semi-parametric Inference for Copula Models for Dependently Truncated Data

*TAKESHI EMURA, **WEIJING WANG and **HUI-NIEN HUNG

emurat@pharm.kitasato-u.ac.jp, wjwang@stat.nctu.edu.tw, hhung@stat.nctu.edu.tw

**Division of Biostatistics, Department of Pharmaceutical Sciences,*

Kitasato University, Japan

***Institute of Statistics, National Chiao-Tung University,*

Hsin-Chu, Taiwan, R.O.C.

SUMMARY

In this article, we investigate the dependent relationship between two failure time variables which truncate each other. Chaieb et al. (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 estimation of the association parameter, we extend Clayton's conditional likelihood approach (1978) and Wang's two-by-two table approach (2003) to dependent truncation data and derive a relationship between the proposed methods and the method of Chaieb et al. (2006) based on U-statistics. For marginal estimation, we propose a novel algorithm and derive an explicit formula to obtain the estimator. Large sample properties are established on the functional delta method which can handles more general estimating functions than the U-statistic approach. Simulations are performed and the proposed methods are applied to the transfusion-related AIDS data for illustrative purposes.

Key words: Archimedean copula model, Conditional likelihood, Product-limit estimator, Kendall's tau, Truncation data.

1. Introduction

Consider the situation that a pair of failure times (X, Y) can be included in the sample only if $X \le Y$. The variable X is said to be right truncated by Y or Y is left truncated by X. Most literature for analyzing truncated data focus on marginal analysis under the assumption that X and Y are quasi-independent (Tsai, 1990). However this condition may 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 et al., 1988). Applying Tsai's test, the hypothesis of quasi-independence between X and Y was rejected. The surprising association may be attributed to the change of medical practice in different chronicle periods and hence sheds some light on the population dynamics of AIDS.

To assess the degree of association, Tsai (1990) modified Kendall's tau by conditioning on an event that guarantees 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 et al. (2006) proposed a "semi-survival" Archimedean copula (AC) model suitable for describing the relationship for two variables with the truncation relationship. Then they developed inference procedures to estimate the copula parameter, marginal functions and the truncation probability.

In this article, we adopt the same model framework as in Chaieb et al. (2006) but propose different inference procedures. The paper is organized as follows. Section 2 reviews related research development. The proposed methods for estimating the association parameter and estimation the marginal functions are presented in Section 3 and 4 respectively. Section 5 contains large sample analysis. The proofs are given in Appendix. Detailed derivations are provided in the attached technical report. In Section 6, 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 7. Concluding remarks are given in Section 8.

2. Preliminary

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

2.1 Association Measures and Models for Typical Failure Time Data

In analysis of failure time data, robust measures are usually preferred. Kendall's tau, known as the rank correlation coefficient, is defined as $\tau = E(2\Delta_{ij} - 1)$, 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) which are independent replications of any random variables (X, Y). Local association can be described by the odds ratio function proposed by Oakes (1989):

$$\widetilde{\theta}(x, y) = \frac{\partial^2 \Pr(X > x, Y > y) / \partial x \partial y \cdot \Pr(X > x, Y > y)}{\partial \Pr(X > x, Y > y) / \partial x \cdot \partial \Pr(X > x, Y > y) / \partial y}$$

$$= \frac{\Pr(\Delta_{ij} = 1 \mid \widetilde{X}_{ij} = x, \widetilde{Y}_{ij} = y)}{\Pr(\Delta_{ij} = 0 \mid \widetilde{X}_{ij} = x, \widetilde{Y}_{ij} = y)},$$

where $\widetilde{X}_{ij} = X_i \wedge X_j$ and $\widetilde{Y}_{ij} = Y_i \wedge Y_j$. Note that the sign of $\log \widetilde{\theta}(x, y)$ 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 survival functions of the failure time variables such that

$$\Pr(X > x, Y > y) = C_\alpha\{\Pr(X > x), \Pr(Y > y)\},$$

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

A useful sub-class of the copula family is the Archimedean copulas (AC) model with $C_\alpha(u,v)$ being simplified as $C_\alpha(u,v) = \phi_\alpha^{-1}\{\phi_\alpha(u) + \phi_\alpha(v)\}$ for $u,v \in [0,1]$, where $\phi_\alpha(\cdot):[0,1] \rightarrow [0,\infty]$ is a generating 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 $\phi_\alpha(\cdot)$, the odds ratio can be written as $\widetilde{\theta}(x,y) = \theta_\alpha\{\Pr(X > x, Y > y)\}$, where $\theta_\alpha(\cdot)$ is a univariate function satisfying

$$\theta_\alpha(v) = -v \cdot \phi''_\alpha(v)/\phi'_\alpha(v). \tag{1}$$

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 $X \le Y$, all the aforementioned descriptive measures are not identifiable. Tsai (1990) proposed a modified version of Kendall's tau given by $\tau_a = E(2\Delta_{ij} - 1 \mid A_{ij})$, where $A_{ij} = \{\breve{X}_{ij} \le \widetilde{Y}_{ij}\}$, $\breve{X}_{ij} = X_i \vee X_j$ and $\widetilde{Y}_{ij} = Y_i \wedge Y_j$. Notice that by conditioning on the event A_{ij}, two locations (X_i,Y_i) and (X_j,Y_j) on the plane become "comparable" under truncation since $(\breve{X}_{ij},\widetilde{Y}_{ij})$ is located in the identifiable region $\mathrm{R}_{\mathrm{U}} = \{(x,y): 0 \le x \le y < \infty\}$.

Chaieb et al. (2006) modified the local odds ratio function as follows:

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

$$= \frac{\Pr(\Delta_{ij} = 0 \mid \breve{X}_{ij} = x, \widetilde{Y}_{ij} = y)}{\Pr(\Delta_{ij} = 1 \mid \breve{X}_{ij} = x, \widetilde{Y}_{ij} = y)} \quad (x \le y). \tag{2.b}$$

The interpretation of $\theta^*(x,y)$ is similar to $1/\widetilde{\theta}(x,y)$ but only the former is identifiable for truncated data. In light of equation (2.a), Chaieb et al. (2006) suggested to impose the copula structure on the "semi-survival" function $\Pr(X \le x, Y > y)$ in the region of $\mathrm{R_U}$. They proposed the following semi-survival AC model:

$$\pi(x,y) = \phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\} + \phi_\alpha\{S_Y(y)\}]/c \quad (x \le y), \tag{3}$$

where $\pi(x,y) = \Pr(X \le x, Y > y \mid X \le Y)$, $F_X(\cdot)$ and $S_Y(\cdot)$ are arbitrary continuous distribution and survival function respectively and c is a normalizing constant satisfying $c = \iint_{x \le y} -\frac{\partial^2}{\partial x \partial y}\phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\} + \phi_\alpha\{S_Y(y)\}]dxdy$. A semi-survival AC model indexed by $\phi_\alpha(t)$ has the property, $\theta^*(x,y) = \theta_\alpha\{c\pi(x,y)\}$, where $\theta_\alpha(v)$ is defined in (1). Equations $F_X(x) = \Pr(X \le x)$, $S_Y(y) = \Pr(Y > y)$ and $c = \Pr(X < Y)$ do not hold in general since the marginal functions $F_X(\cdot)$ and $S_Y(\cdot)$ only stipulate the function $\pi(x,y)$ in the region $\mathrm{R_U}$. More detailed discussions for the interpretation of marginal functions can be found in Chaieb et al. (2006).

2.3 Previous Inference Results for Dependent Truncation Data

In truncation data, the sample consists of $\{(X_j, Y_j)\,(j = 1,\ldots,n)\}$ subject to $X_j \le Y_j$. Chaieb et al. (2006) considered an estimator of $\tau_a = E(2\Delta_{ij} - 1 \mid A_{ij})$ under semi-survival models in (3) and set their estimator to be equal to the nonparametric estimator;

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

This equation can be used as an estimating function for (α, c). Chaieb et al. (2006) rewrite the equation in the U-statistics form to study the asymptotic behavior;

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

Now we generalize the preceding formula in the following way;

$$\tilde{U}_w(\alpha,c)=\sum_{i<j}1\{A_{ij}\}\tilde{w}_{\alpha,c}(\breve{X}_{ij},\tilde{Y}_{ij})\left[\Delta_{ij}-\frac{1}{1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}}\right], \tag{5}$$

where $\tilde{w}_{\alpha,c}(x,y)$ is a weight function and $\hat{\pi}(x,y)=\sum_j I(X_j\le x, Y_j>y)/n$. When $\tilde{w}_{\alpha,c}(x,y)=1$, the above function reduces to the equation based on the conditional Kendall's tau. Notice that $\tilde{U}_w(\alpha,c)$ in (5) also involves the truncation proportion c. Note that in the special case of Clayton's model with $\phi_\alpha(t)=t^{-(\alpha-1)}-1$ $(\alpha>1)$ and $\theta_\alpha(v)=\alpha$, $\tilde{U}_w(\alpha,c)$ depends only on α. This implies additional estimation procedure for c is needed.

Since X and Y are dependent, existing methods for estimating the truncation probability are not applicable (He and Yang, 1998) since they assume quasi-independence. Chaieb et al. (2006) proposed their second estimating procedure based on the modification of equation (3):

$$\phi_\alpha\left\{c\frac{R(t,t)}{n}\right\}=\phi_\alpha\{F_X(t)\}+\phi_\alpha\{S_Y(t-)\}. \tag{6}$$

They adopted the idea of Rivest and Wells (2001) by first estimating the jumps, $\phi_\alpha\{S_Y(Y_i)\}-\phi_\alpha\{S_Y(Y_i-)\}$ and $\phi_\alpha\{F_X(X_i)\}-\phi_\alpha\{F_X(X_i+)\}$ based on (6), and then summing them up over all the failure times prior to t to obtain the estimators for $\phi_\alpha\{F_X(t)\}$ and $\phi_\alpha\{S_Y(t)\}$. By plugging in all the marginal estimators into equation (6), an estimating function for c can be obtained.

3. Estimating the Association Parameter for Semi-survival Copula Models

3.1 Motivation

Before we introduce the proposed methods, it is worthy to review the development of inference methods for copula models imposed on the survival functions. Early work focused on the Clayton copula (Clayton, 1978) based on right censored data. In his landmark paper, Clayton (1978) obtained an estimator for the association parameter by maximizing a likelihood function which can be expressed as the product of conditional probabilities. 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 which 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 et al. (1997) and Wang (2003) for Clayton's model and general AC models respectively. Estimating functions using the concordance information for paired observations have been proposed by Fine et al. (2001).

3.2 Estimation based on Conditional Likelihood

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

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

For a point (x,y) in φ, we define the number at-risk $R(x,y) = \sum_j I(X_j \le x, Y_j \ge y)$. Let $\Delta(x,y) = \sum_j I(X_j = x, Y_j = y)$ be a binary variable indicating whether failure occurs at (x,y). Given $R(x,y) = r$ for $(x,y) \in \varphi$ and under model (3), the variable $\Delta(x,y)$ follows a Bernoulli distribution 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 only little information about α, we can only use the conditional probability to construct the following likelihood function:

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

The nuisance parameter $\pi(x,y)$ can be estimated nonparametrically by $\hat{\pi}(x,y) = R(x,y)/n$. Differentiating $\log L(\alpha, c, \hat{\pi})$ with respect to α, we get the following estimating function

$$U_L(\alpha, c) = \iint_{(x,y)\in\varphi} \frac{\dot{\theta}_\alpha\{c\hat{\pi}(x,y)\}}{\theta_\alpha\{c\hat{\pi}(x,y)\}} \left[\Delta(x,y) - \frac{\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y) - 1 + \theta_\alpha\{c\hat{\pi}(x,y)\}} \right] = 0, \qquad (7)$$

where $\dot{\theta}_\alpha(v) = \partial\theta_\alpha(v)/\partial\alpha$. For the case of Clayton's model, $U_L(\alpha, c)$ reduces to

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

which depends only on α. However for the other members in the AC family, estimation of α requires the information of c. For many models, $\partial \log L(\alpha, c, \pi)/\partial c$ 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, Frank copula has the form

$\theta_\alpha(c\pi) = c\pi\log(\alpha)/(e^{c\pi\log(\alpha)} - 1)$ and it is a function of single parameter $\gamma = c\log(\alpha)$. This implies that the likelihood function can not identify α and c simultaneously.

3.2 Estimation based on Two-by-two Tables

Motivated by the ideas of Day et al. (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 the cell counts can be denoted as $N_{11}(dx,dy) = \sum_j I(X_j = x, Y_j = y)$, $N_{\bullet 1}(x,dy) = \sum_j I(X_j \le x, Y_j = y)$ 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). Given the marginal counts, the conditional mean of $N_{11}(dx,dy)$ under model (3) is

$$E\{N_{11}(dx,dy) \mid N_{1\bullet}, N_{\bullet 1}, R\} = \frac{N_{\bullet 1}(x,dy)N_{1\bullet}(dx,y)\theta_\alpha\{c\pi(x,y)\}}{R(x,y) - 1 + \theta_\alpha\{c\pi(x,y)\}}$$

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

$$U_w(\alpha,c) = \iint_{x \le y} w_{\alpha,c}(x,y)\left[N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y) - 1 + \theta_\alpha\{c\hat{\pi}(x,y)\}}\right], \quad \text{(8a)}$$

where $w_{\alpha,c}(x,y)$ is a weight function. When the data have no ties, 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,

$$U_w(\alpha,c) = \iint_{(x,y)\in\varphi} w_{\alpha,c}(x,y)\left[\Delta(x,y) - \frac{\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y) - 1 + \theta_\alpha\{c\hat{\pi}(x,y)\}}\right], \quad \text{(8b)}$$

which implies that $U_L(\alpha,c)$ is a special case of $U_w(\alpha,c)$ with the weight

$$w_{\alpha,c}(x,y)=\dot{\theta}_\alpha\{c\hat{\pi}(x,y)\}/\theta_\alpha\{c\hat{\pi}(x,y)\}$$

which utilizes the conditional likelihood information. The preceding weight function can also be obtained as the derivative of the conditional mean of $N_{11}(dx,dy)$ with respect to α, divided by the conditional variance of $N_{11}(dx,dy)$, This is a guideline for the optimal weight function under the independence assumption of all tables (Godambe 1991).

3.3 Concordance-type Expression

Clayton & Cuzick (1985) have expressed Clayton's likelihood estimator in terms of concordance indicators for censored data. We now establish similar relationship for truncation data. Some algebraic calculations yield the following identity:

$$\iint_{(x,y)\in\varphi} w_{\alpha,c}(x,y)\left[\Delta(x,y)-\frac{\theta_\alpha\{c\hat{\pi}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c\hat{\pi}(x,y)\}}\right]$$

$$=-\sum_{i<j}1\{A_{ij}\}\frac{w_{\alpha,c}(\breve{X}_{ij},\tilde{Y}_{ij})[1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}]}{R(\breve{X}_{ij},\tilde{Y}_{ij})-1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}}\left[\Delta_{ij}-\frac{1}{1+\theta_\alpha\{c\hat{\pi}(\breve{X}_{ij},\tilde{Y}_{ij})\}}\right]. \quad (9)$$

The proof is given in Appendix B which includes external censoring in the analysis. Hence the proposed estimating function $U_L(\alpha,c)$ is a special case of $\tilde{U}_w(\alpha,c)$ in (5) with

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

The estimating function proposed by Chaieb et al. (2006) corresponds to

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

The above analysis shows that the three different inference methods can be unified. What really matters is the choice of weight. Although some guidelines for choosing the weight function have been suggested (Fine et al., 2001), no theoretical justification is provided. It is reasonable to expect that $U_L(\alpha,c)$, which utilizes some likelihood information in derivation of the weight, may produce more efficient result than estimating functions with ad-hoc weights. We will examine this conjecture via simulations.

4. Estimation of Marginal Functions and Truncation Probability

Now we propose a new algorithm for a joint estimator of (α,c,F_X,S_Y). Let $t_1 < \cdots < t_{2n}$ be ordered observed points of $(X_1,\ldots,X_n,Y_1,\ldots,Y_n)$ and $t_0 = 0$. Replacing $\pi(t,t)$ by $R(t,t+)/n$ in equation (3) , we obtain

$$\phi_\alpha\left\{c\frac{R(t_i,t_i+)}{n}\right\} = \phi_\alpha\{F_X(t_i)\} + \phi_\alpha\{S_Y(t_i)\} \qquad (i = 1,\ldots,2n-1). \tag{10}$$

This is a minor modification of (6), but a key to perform the subsequent algorithm. Denote the estimators of $F_X(t)$ and $S_Y(t)$ as $\hat{F}_X(t)$ and $\hat{S}_Y(t)$ respectively and assume that they are step functions with jumps only at observed failure points. Under model (3), the unknown parameters consist of:

$$\{\alpha,c,F_X(X_1),\ldots,F_X(X_n),S_Y(Y_1),\ldots,S_Y(Y_n)\} \in R^{2n+2}$$

To produce a unique solution, $2n+2$ equations are needed. There are only $2n-1$ equations in (10) which permit numerous solutions. Without prior information, it is reasonable to add two boundary constraints $\hat{F}_X(t_{2n-1}) = 1$ and $\hat{S}_Y(t_1) = 1$. Together with the equation $U_L(\alpha,c) = 0$, there are $2n+2$ constraints which would produce a unique solution. Fixing an arbitrary value for (α,c), equation (10) can be regarded as an estimating function for $\{F_X(t_i),S_Y(t_i)\}$. Solution for $\{F_X(t_i),S_Y(t_i)\}$ can only be obtained for the boundary points. For instance, the initial constraint $\hat{S}_Y(t_1) = 1$ immediately gives the solution $\hat{F}_X(t_1) = cR(t_1,t_1+)/n$. The procedure is successively defined for $j = 1,\ldots,2n-1$.

(Step 1) If t_j corresponds to an observed value of X, set

$$\phi_\alpha\{\hat{S}_Y(t_j)\} = \phi_\alpha\{\hat{S}_Y(t_{j-1})\} \quad \text{and} \quad \phi_\alpha\{\hat{F}_X(t_j)\} = \phi_\alpha\left\{c\frac{R(t_j,t_j+)}{n}\right\} - \phi_\alpha\{\hat{S}_Y(t_{j-1})\} \ ;$$

and if t_j corresponds to an observed value of Y, set

$$\phi_\alpha\{\hat{F}_X(t_j)\} = \phi_\alpha\{\hat{F}_X(t_{j-1})\} \ \text{ and } \ \phi_\alpha\{\hat{S}_Y(t_j)\} = \phi_\alpha\left\{c\frac{R(t_j,t_j+)}{n}\right\} - \phi_\alpha\{\hat{F}_X(t_{j-1})\} .$$

(Step 2) Set $U_c(\alpha,c) = \phi_\alpha\{\hat{F}_X(x_{(n)})\} = 0$ for $x_{(n)} = \max_j(X_j)$ to meet the constraint $\hat{F}_X(t_{2n-1}) = 1$. Jointly solving this equation and $U_L(\alpha,c) = 0$ gives the estimators of (α,c), denoted as $(\hat{\alpha},\hat{c})$.

(Step 3) Redo (Step 1) by setting $(\alpha,c) = (\hat{\alpha},\hat{c})$ obtained in (Step 2) and then update $(\phi_\alpha\{\hat{F}_X(t_j)\},\phi_\alpha\{\hat{S}_Y(t_j)\})$.

It is easy to see that the above algorithm can be written in explicit formulas:

$$\phi_\alpha\{\hat{F}_X(t)\} = \sum_{j;x_{(1)}<x_j\le t}\left[\phi_\alpha\left\{c\frac{\widetilde{R}(x_j)}{n}\right\} - \phi_\alpha\left\{c\frac{\widetilde{R}(x_j)-1}{n}\right\}\right] + \phi_\alpha\left(\frac{c}{n}\right), \tag{11}$$

$$\phi_\alpha\{\hat{S}_Y(t)\} = -\sum_{j;y_j\le t}\left[\phi_\alpha\left\{c\frac{\widetilde{R}(y_j)}{n}\right\} - \phi_\alpha\left\{c\frac{\widetilde{R}(y_j)-1}{n}\right\}\right], \tag{12}$$

where $\widetilde{R}(t) = R(t,t)$, $x_{(1)} = \min_j(X_j)$ and $y_{(1)} = \min_j(Y_j)$.

Also, the equation in (Step2) can be written as

$$U_c(\alpha,c) = \sum_{j;x_{(1)}<x_j}\left[\phi_\alpha\left\{c\frac{\widetilde{R}(x_j)}{n}\right\} - \phi_\alpha\left\{c\frac{\widetilde{R}(x_j)-1}{n}\right\}\right] + \phi_\alpha\left(\frac{c}{n}\right). \tag{13}$$

In the case of quasi-independence with $\phi_\alpha(t) = -\log(t)$, equations (11)- (13) reduce to the Lynden-Bell's estimators and the natural estimator of the truncation proportion (He and Yang, 1998). It is worthy to note that the representation of the Lynden-Bell's estimator as a solution to the moment equation in (10) with $\phi_\alpha(t) = -\log(t)$ is new in the literature. Compared with the product-limit expression, our approach provides a more general estimating scheme which allows for dependence between the two variables.

In principle, imposing other boundary constraints on $F_X(t_{2n-1})$ and $S_Y(t_1)$ produce a different solution to equations (10) and $U_L(\alpha,c)=0$. By choosing $\hat{F}_X(t_{2n-1})=1$ and $\hat{S}_Y(t_1)=1$, the explicit formula in (11)-(13) can be derived. Note that the proposed estimators in (11) and (13) are different from the results of Chaieb et al. (2006). However, equation (12) is identical to the proposal of Chaieb et al. (2006).

5. Asymptotic analysis

Under the regularity conditions (A-I)~(A-V) listed in Appendix A1, the estimators $(\hat{\alpha},\hat{c})$ which jointly solve $\widetilde{U}_w(\alpha,c)=0$ in (5) and $U_c(\alpha,c)=0$ in (13) are consistent and asymptotically normal. Note that $\widetilde{U}_w(\alpha,c)=0$ includes $U_L(\alpha,c)=0$ as a special case. Weak convergence of the marginal estimators is also established. The results are formally stated in the following theorems.

Theorem 1 *Random vector* $(\hat{\alpha},\hat{c})$ *is consistent.*

Theorem 2 *Random vector* $n^{1/2}(\hat{\alpha}-\alpha_0,\hat{c}-c_0)^T$ *converges in distribution to a bivairate normal distribution with mean 0 and covariance matirix* $A^{-1}B(A^{-1})^T$*, where* $A=E[\dot{U}_{\alpha_0,c_0}(X,Y)]$ *,* $B=E[U_{\alpha_0,c_0}(X,Y)U_{\alpha_0,c_0}(X,Y)^T]$ *and the definitions of* $U_{\alpha_0,c_0}(X,Y)$ *and* $\dot{U}_{\alpha_0,c_0}(X,Y)$ *are given in (A.4).*

Theorem 3 *The bivariate 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 in (A.4) for* $0 \le s,t < \infty$.

Note that Chaieb et al. (2006) established similar results for their estimator which solves $\widetilde{U}_w(\alpha, c)$ with $\widetilde{w}_{\alpha,c}(x, y) = 1$ by applying properties of U-statistics. However this approach may not be applicable when $\widetilde{w}_{\alpha,c}(x, y)$ involves the plugged-in estimator $\hat{\pi}(x, y)$ as in our case. Here we take a different approach which can handle more general weight functions. Specifically asymptotic linear representations of the proposed estimating functions are obtained. By applying the functional delta method (Van Der Vaart, 1998, theorem 20.8) and properties of empirical processes, large-sample properties of the proposed estimators can be established. The sketch of the proof is given in Appendix A2. For the details, please refer to the attached technical report. Since the analytic derivations involve complicated formula, we suggest to use the jackknife method. Specifically, to evaluate the variance of $\hat{\theta} = (\hat{\alpha}, \hat{c}, \hat{F}_X(x), S_Y(y))'$, the jackknife estimator is given by

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

where $\hat{\theta}^{(-j)}$ is the estimator ignoring j th observation and $\hat{\theta}^{(\cdot)} = \sum_j \hat{\theta}^{(-j)} / n$.

Given $\phi_\alpha(t) = -\log(t)$, the condition for quasi-independence, the asymptotic expression of $U_{\alpha,c}(X_i, Y_i) = 0$ in (A.5) of the Appendix reduces to the iid representation obtained in both Stute (1993) and He and Yang (1998). Specifically it follows that

$$n^{1/2}\begin{bmatrix} \hat{S}_Y(t) - S_Y(t) \\ \hat{F}_X(t) - F_X(t) \end{bmatrix} = \frac{1}{n^{1/2}}\sum_{i=1}^{n}\begin{bmatrix} -S_Y(t)L^Y(X_i, Y_i; t) \\ -F_X(t)L^X(X_i, Y_i; t) \end{bmatrix} + o_p(1),$$

where

$$L^Y(X, Y; t) = \int_{y_L}^{t} \frac{I(X \le u, Y \ge u)}{\pi(u,u)^2} d\pi(\infty, u) + \frac{I(Y \le t)}{\pi(Y, Y)}$$

and

$$L^X(X,Y;t) = -\int_t^{x_U} \frac{I(X \le u, Y \ge u)}{\pi(u,u)^2} d\pi(u,0) + \frac{I(X > t)}{\pi(X,X)}.$$

6. Extension and Modification

6.1. Extension under Right Censoring

Consider an extended situation that 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 $\{(X_i, Z_i, \delta_i)\,(i=1,...,n)\}$ 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) $(i=1,...,n)$ are random replications of (X,Y,C).

Chaieb et al. (2006) expressed the semi-survival AC model as

$$\begin{aligned}\pi^*(x,y) &= \Pr(X \le x, Z > y \mid X \le Z) \\ &= S_C(y)\phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\} + \phi_\alpha\{S_Y(y)\}]/c^* \quad (x \le y),\end{aligned} \tag{14}$$

where $S_C(y) = \Pr(C > y)$ and c^* is a normalizing constant satisfying

$$c^* = \iint_{x \le y} -\frac{\partial^2}{\partial x \partial y}\left(S_C(y)\phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\} + \phi_\alpha\{S_Y(y)\}]\right) dxdy. \tag{15}$$

The objective is to estimate the unknown parameters $(\alpha, c^*, F_X(\cdot), S_Y(\cdot), S_C(\cdot))$. Hence 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 the conditional likelihood estimation. To simplify the presentation, we still use the same notations but change their definitions as follows. The set of grid points become:

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

Let

$$\Delta(x,y)=\sum_j I(X_j=x,Z_j=y,\delta_j=1)\,,\;\; R(x,y)=\sum_j I(X_j\le x,Z_j\ge y)\,.$$

Consequently the proposed estimating function is

$$U_L(\alpha,c^*)=\iint_{(x,y)\in\varphi}\frac{\dot\theta_\alpha\{c^*\hat v(x,y)\}}{\theta_\alpha\{c^*\hat v(x,y)\}}\left[\Delta(x,y)-\frac{\theta_\alpha\{c^*\hat v(x,y)\}}{R(x,y)-1+\theta_\alpha\{c^*\hat v(x,y)\}}\right],\tag{16}$$

where $\hat v(x,y)=R(x,y)/\{n\hat S_C(y)\}$ and $\hat S_C(y)$ is the Lynden-Bell's estimator given by

$$\hat S_C(y)=\prod_{z_j\le y,\delta_j=0}\{1-1/R(z_j)\}\,.$$

In presence of censoring, we can still express $U_L(\alpha,c^*)$ in terms of pairwise concordance indicators as in (9). The detailed expressions are given in Appendix B.

The procedure for marginal estimation can be modified as follows. Let $t_1<\cdots<t_{2n}$ be ordered observed points of $(X_1,\ldots,X_n,Z_1,\ldots,Z_n)$. The estimating functions become

$$\phi_\alpha\left\{c^*\frac{R(t_i,t_i+)}{nS_C(t_i)}\right\}=\phi_\alpha\{F_X(t_i)\}+\phi_\alpha\{S_Y(t_i)\}\quad(i=1,\ldots,2n-1)\,.\tag{17}$$

To solve the above equations, we impose additional constraints that the estimators of F_X, S_Y and S_C are step functions with jumps only at their observed values, and that $\hat F_X(t_{2n-1})=1$, $\hat S_Y(t_1)=1$ and $\hat S_C(t_1)=1$. The solution can be obtained by performing the following algorithm successively for $j=1,2,...,2n-1$.

(Step 1) If t_j corresponds to an observed value of X, set

$$\phi_\alpha\{\hat S_Y(t_j)\}=\phi_\alpha\{\hat S_Y(t_{j-1})\}\,,\;\;\phi_\alpha\{\hat F_X(t_j)\}=\phi_\alpha\left\{c^*\frac{R(t_j,t_j+)}{\hat S_C(t_{j-1})n}\right\}-\phi_\alpha\{\hat S_Y(t_{j-1})\}\,,$$

and $\hat S_C(t_j)=\hat S_C(t_{j-1})$; if t_j corresponds to an observed value of Y, set

$$\phi_\alpha\{\hat{F}_X(t_j)\} = \phi_\alpha\{\hat{F}_X(t_{j-1})\}\ ,\ \ \phi_\alpha\{\hat{S}_Y(t_j)\} = \phi_\alpha\left\{c^*\frac{R(t_j,t_j+)}{\hat{S}_C(t_{j-1})n}\right\} - \phi_\alpha\{\hat{F}_X(t_{j-1})\}$$

and $\hat{S}_C(t_j) = \hat{S}_C(t_{j-1})$; and if t_j corresponds to an observed value of C, set

$$\phi_\alpha\{\hat{F}_X(t_j)\} = \phi_\alpha\{\hat{F}_X(t_{j-1})\},\ \ \phi_\alpha\{\hat{S}_Y(t_j)\} = \phi_\alpha\{\hat{S}_Y(t_{j-1})\},$$

and $\hat{S}_C(t_j) = \{1 - 1/R(t_{j-1},t_{j-1}+)\}\hat{S}_C(t_{j-1})$.

(Step 2) Set $U_c(\alpha,c^*) = \phi_\alpha\{\hat{F}_X(x_{(n)})\} = 0$ to meet the constraint $\hat{F}_X(t_{2n-1}) = 1$. Jointly solving this equation and $U_L(\alpha,c^*) = 0$ in (16) produces the estimators of (α,c^*), denoted as $(\hat{\alpha},\hat{c}^*)$.

(Step 3) Redo (Step 1) by setting $(\alpha,c^*) = (\hat{\alpha},\hat{c}^*)$ obtained in (Step 2) and then obtain $(\phi_\alpha\{\hat{F}_X(t_j)\}, \phi_\alpha\{\hat{S}_Y(t_j)\}, \hat{S}_C(t_j))$.

Explicit formula of the proposed estimators are given by

$$\phi_\alpha\{\hat{S}_Y(t)\} = -\sum_{j;z_j\le t,\delta_j=1}\left[\phi_\alpha\left\{c^*\frac{R(z_j)}{n\hat{S}_C(z_j)}\right\} - \phi_\alpha\left\{c^*\frac{R(z_j)-1}{n\hat{S}_C(z_j)}\right\}\right], \quad (18)$$

$$\phi_\alpha\{\hat{F}_X(t)\} = \sum_{j;x_{(1)}<x_j\le t}\left[\phi_\alpha\left\{c^*\frac{R(x_j)}{n\hat{S}_C(x_j)}\right\} - \phi_\alpha\left\{c^*\frac{R(x_j)-1}{n\hat{S}_C(x_j)}\right\}\right] + \phi_\alpha\left(\frac{c^*}{n}\right). \quad (19)$$

The estimating function in (Step 2) is equivalent to

$$U_c(\alpha,c^*) = \sum_{j;x_{(1)}<x_j}\left[\phi_\alpha\left\{c^*\frac{R(x_j)}{\hat{S}_C(x_j)n}\right\} - \phi_\alpha\left\{c^*\frac{R(x_j)-1}{\hat{S}_C(x_j)n}\right\}\right] + \phi_\alpha\left(\frac{c^*}{n}\right). \quad (20)$$

Under the special case of $\phi_\alpha(t) = -\log(t)$ which is the condition of quasi-independence, the proposed estimators $\hat{S}_Y(t)$, $\hat{F}_X(t)$ and $\hat{S}_C(t)$ all reduce to the Lynden-Bell's estimators under right censoring. In Appendix C, we derive explicit formula of the proposed estimating functions for selected examples of $\phi_\alpha(t)$.

6.2. Modification for small risk sets

The proposed estimation procedure, as well as that proposed by Chaieb et al. (2006) are both based on the implicit assumption that $R(t_j, t_j+) \geq 1$ for all t_j. However it sometimes happens that an empty risk set may occur especially in the tail area. Several remedies have been proposed to handle this problem (Klein & Moeschberger, 2003, p. 122). Here we adopt the idea of Lai and Ying (1991) and propose the following modification:

$$\phi_\alpha\{\hat{S}_Y(t)\} = -\sum_{j;z_j \leq t, \delta_j = 1} \left[\phi_\alpha\left\{c * \frac{\widetilde{R}(z_j)}{n\hat{S}_c(z_j)}\right\} - \phi_\alpha\left\{c * \frac{\widetilde{R}(z_j) - 1}{n\hat{S}_c(z_j)}\right\}\right] I\{\widetilde{R}(z_j) \geq bn^a\}, \quad (21)$$

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 the simulations not shown here, we found that taking $b = 1$ and $a = 1/10$ would produce less biased results.

7 Numerical Analysis

7.1 Simulation Studies

The main purposes of the simulation studies are (i) to check the validity of the proposed estimators and (ii) to compare the performance of our method with its competitor proposed by Chaieb et al. (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 value of $\log(\alpha)$ were set to be 2.380 and 5.746. The former transformation corresponds to $\tau = 0.25$ and the latter corresponds to $\tau = 0.5$. The censoring variable C also followed an exponential distribution. Denote $c = \Pr(X \leq Y)$

and $c^* = \Pr(X \le Y \wedge C)$. For each setting, we report the bias and the MSE based on 500 replications.

Tables 1A and 1B summarize the results for comparing two estimators of α under the Clayton model and Frank model respectively. Recall that the proposed method solve $U_L(\alpha, c) = 0$ and the competing estimator proposed by Chaieb et al. (2006) solve $\widetilde{U}_w(\alpha, c) = 0$ with $\widetilde{w}_{\alpha,c}(x, y) = 1$. Explicit formulas for the Clayton and Frank models are derived in Appendix C. We see that both methods are approximately unbiased, and the MSE decreases as the sample size increases. Compared with the estimator proposed by Chaieb et al. (2006), the proposed estimator has significantly smaller MSE under Clayton's model. On the other hand, the gain in efficiency becomes modest under Frank's model. Notice that the two approaches produce similar results only under the Frank model in absence of external censoring $(c = c^*)$. In fact, via equation (10), we find that for Frank's model

$$\frac{\dot{\theta}_\alpha\{c\hat{\pi}(x,y)\}}{\theta_\alpha\{c\hat{\pi}(x,y)\}} \cong const. \times \frac{R(x,y) - 1 + \theta_\alpha\{c\hat{\pi}(x,y)\}}{1 + \theta_\alpha\{c\hat{\pi}(x,y)\}}$$

which explains why the numerical results are close. Nevertheless when the censoring proportion increases, the proposed estimator tends to have better performance.

Table 2A and Table 2B report the results for evaluating the proposed second inference procedure under the Clayton model and Frank model respectively. The proposed successive algorithm was carried out jointly with $U_L(\alpha, c) = 0$ to obtain the estimators of the marginal functions and c. The performances of $(\hat{F}_X(t), \hat{S}_Y(t))$ were evaluated at selected points of t with $F_X(t) = 0.2, 0.4, 0.6, 0.8$ and $S_Y(t) = 0.2, 0.4, 0.6, 0.8$. Denote $P_{CEN} = \Pr(C < Y \mid X \le Z)$ which measures the censoring proportion in the truncated sample. In all the cases, $(c^*, \hat{F}_X(\cdot), \hat{S}_Y(\cdot))$ are fairly unbiased. It is worthy to note that the estimated distribution/survival functions have nicer performance in the tail area but poorer performance in a middle time point, which behave differently from the usual Kaplan-Meier estimator. This indicates that the initial constraints $\hat{F}_X(t_{2n-1}) = 1$ and $\hat{S}_Y(t_1) = 1$ are good approximation of the true values $(F_X(t_{2n-1}), S_Y(t_1))$ with sufficiently large sample sizes considered here.

7.3 Data analysis

We apply the inference procedures to analyze the transfusion-related AIDS data discussed in Kalbfleisch and Lawless (1989). Let T be the infection time, measured form 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 is 102 months, individuals with $T + X \le 102$ were included in the sample which 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 analyze the data under two different model assumptions. The results are summarized in Table 3. Under the Clayton model, both estimators of α show positive correlation between X and Y which implies that patients who infected earlier tended to have longer incubation time. Both estimators rejected the null hypothesis of quasi-independence: $H_0 : \alpha = 1$. This conclusion coincides with Tsai's nonparametric test (1990). The confidence interval for $-\log(\alpha)$ based on the proposed likelihood estimator is narrower than that obtained by the estimator of Chaieb et al. (2006). The level of association between X and Y was even stronger under the Frank model assumption. As in the simulations (Table 1B with $c = c^*$), the two estimators produced similar results.

Figure 1 depicts the estimated incubation distributions under the two model assumptions by applying the proposed recursive algorithm. The estimated curve under the Clayton model is significantly lower than that under Frank's model. It implies that the marginal estimator is also sensitive to the model choice.

8. Conclusion

This article considers semiparametric inference for dependent truncation data based on semi-survival AC models proposed by Chaieb et al. (2006). In particular we generalize Clayton's conditional likelihood approach to estimating the copula parameter under truncation. We also generalize two-by-two table approaches of Wang (2003) to construct more general estimating functions. Furthermore we show that the three different inference methods can be unified under the same framework. The estimator proposed by Chaieb et al. (2006) is constructed based on the first moment condition which does not contain extra information for choosing the weight. On the other hand, the proposed conditional likelihood approach provides theoretical justification for the imposed weight and hence may produce more efficient results than other ad-hoc choices of weights. Another potential advantage of the conditional likelihood approach is that it can directly handle the situation when the dimension of α exceeds one. The proposed successive algorithm, which solves equations (10) and the two additional artificial constraints, yields simple explicit formula and is easy to implement. In derivations of the large-sample properties, we apply the functional delta method which can handle more general estimating functions than the U-statistic approach.

A series of simulations shows that the proposed conditional likelihood method decrease the MSE compared to the competing approach of Chaieb (2006). Nevertheless, it is still hard to derive any optimality property of the conditional likelihood estimator since it is constructed based on a working assumption of independence among the grids. It will be an interesting topic to further establish some theoretical foundation for this kind of conditional likelihood functions under Copula models.

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

Appendix

A1: *Regularity Conditions*

(A-I) A parameter space $\Theta \subset R^2$ for (α, c) is compact;

(A-II) Deterministic functions $\phi_\alpha(v)$, $\phi'_\alpha(v) \equiv \partial\phi_\alpha(v)/\partial v$, $\phi''_\alpha(v) \equiv \partial^2\phi_\alpha(v)/\partial v^2$, $\phi^{-1}{}_\alpha(v)$, $\dot{\phi}^{-1}_\alpha(v) = \partial\phi^{-1}_\alpha(v)/\partial\alpha$, $\theta_\alpha(v)$, $\theta'_\alpha(v) \equiv \partial\theta_\alpha(v)/\partial v$, $\theta''_\alpha(v) \equiv \partial^2\theta_\alpha(v)/\partial v^2$, $\dot{\theta}_\alpha(v) \equiv \partial\theta_\alpha(v)/\partial\alpha$, $\tilde{w}_\alpha(v)$, and $\tilde{w}'_\alpha(v) \equiv \partial\tilde{w}_\alpha(v)/\partial v$ are twice continuously differentiable and bounded function of (α, v);

(A-III) There exists a function $\tilde{w}_\alpha\{\cdot\}: R \to R$ such that

$$\sup_{x,y} \left| \tilde{w}_{\alpha,c}(x,y) - \tilde{w}_\alpha\{c\hat{\pi}(x,y)\} \right| = o_p(n^{-1/2});$$

(A-IV) There exists two positive numbers $y_L < x_U$ such that

$$F_X(y_L) > 0,\ S_Y(y_L) = 1,\ F_X(x_U) = 1 \text{ and } S_Y(x_U) > 0.$$

(A-V) The (2×2) matrix A is non-singular, whose definition is given later.

Note that (A-IV) is a condition for the identifiability of $(F_X(\cdot), S_Y(\cdot))$, which has been routinely used in theoretical analysis of truncation data. For example, the upper limit x_U plays the same role as the notation T^* in Wang et al. (1986).

A2: *A Sketch of Asymptotic Analysis*

To simplify the notations, we define the following quantities $g_\alpha(v) \equiv 1/\{1 + \theta_\alpha(v)\}$, $\pi(\infty, y) = \mathrm{pr}(Y > y \mid X \le Y)$ and $\pi(x, 0) = \Pr(X \le x \mid X \le Y)$. Also, let $\{D[0,\infty)\}^2$ be a space consisting of right-continuous function $(f_1(t), f_2(t))^T$ with left-side limits, where $f_k(t): [0,\infty) \mapsto R$ for $k = 1,2$. The metric is defined as $d(f,g) = \max\{\sup_{0 \le t < \infty} | f_k(t) - g_k(t) |; k = 1,2\}$ for $f, g \in \{D[0,\infty)\}^2$. Similarly, the space

$D\{[0,\infty)^2\}$ consists of right-continuous function $f(s,t)$ with left-side limits, where $f(s,t):[0,\infty)^2 \mapsto R$, equipped with the usual sup-norm. Let $\Theta \subset R^2$ be the parameter space for (α, c), and $(\alpha_0, c_0) \in \Theta$ is denoted as the true parameter value. Hereafter, expectation symbols represent the conditional expectation given $X \le Y$.

All the estimating functions and estimators in this paper can be approximately expressed as a Hadamard differentiable function of the empirical process $\hat{\pi}(x,y) = R(x,y)/n$. The functional delta method is applied based on the weak convergence result of $n^{1/2}(\hat{\pi}(x,y) - \pi(x,y))$ to a Gaussian process $W(x,y)$ on $D\{[0,\infty)^2\}$ with the covariance structure given by:

$$\operatorname{cov}\{W(x_1,y_1), W(x_2,y_2)\} = \pi(x_1 \wedge x_2, y_1 \vee y_2) - \pi(x_1,y_1)\pi(x_2,y_2)$$

for $x_1 \le y_1$ and $x_2 \le y_2$.

The estimating function $\widetilde{U}_w(\alpha,c)$ can be expressed as:

$$\binom{n}{2}^{-1} \widetilde{U}_w(\alpha,c) = \Phi(\pi;\alpha,c) + \frac{1}{n}\sum_{i=1}^{n} \Phi'_\pi(\pi_{(X_i,Y_i)} - \pi;\alpha,c) + o_p(n^{-1/2}) \qquad \text{(A.1)}$$

where $\pi \in D\{[0,\infty)^2\}$, $\pi_{(X_i,Y_i)}(x,y) = I(X_i \le x, Y_i \ge y)$,

$$\begin{aligned}\Phi(\pi;\alpha,c) = &\iint\iint_{x\vee x^* \le y \wedge y^*} \widetilde{w}_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\} \\ &\times [I((x-x^*)(y-y^*)>0) - g_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\}]\, d\pi(x,y)\, d\pi(x^*,y^*),\end{aligned}$$

and

$$\begin{aligned}\Phi'_\pi(h;\alpha,c) = c&\iint\iint_{x\vee x^* \le y\wedge y^*} h(x\vee x^*, y\wedge y^*) \\ &\times \big(\widetilde{w}'_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\}[I\{(x-x^*)(y-y^*)>0\} - g_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\}] \\ &- \widetilde{w}_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\} g'_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\}\big) d\pi(x,y)\, d\pi(x^*,y^*) \\ &+ 2\iint\iint_{x\vee x^* \le y\wedge y^*} \widetilde{w}'_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\} \\ &\times [I\{(x-x^*)(y-y^*)>0\} - g_\alpha\{c\pi(x\vee x^*, y\wedge y^*)\}]\, dh(x,y)\, d\pi(x^*,y^*).\end{aligned}$$

Similarly $U_c(\alpha,c)$ can be expressed as

$$U_c(\alpha,c)=\Psi(\pi;\alpha,c)+\frac{1}{n}\sum_{i=1}^{n}\Psi'_\pi(\pi_{(X_i,Y_i)}-\pi;\alpha,c)+o_p(n^{-1/2}), \qquad \text{(A.2)}$$

where

$$\Psi(\pi;\alpha,c)=c\int_{y_L}^{x_U}\phi'_\alpha\{c\pi(u,u)\}d\pi(u,0)+\phi_\alpha\{c\pi(y_L,y_L)\}$$

and

$$\Psi'_\pi(\pi_{(X_i,Y_i)}-\pi;\alpha,c)=c^2\int_{y_L}^{x_U}\phi''_\alpha\{c\pi(u,u)\}\{I(X_i\le u,Y_i\ge u)-\pi(u,u)\}d\pi(u,0)$$

$$-c\int_{y_L}^{x_U}\{I(X_i<u)-\pi(u,0)\}d\phi'_\alpha(c\pi(u,u)).$$

Both terms of $\Phi'_\pi(\pi_{(X_i,Y_i)}-\pi;\alpha,c)$ and $\Psi'_\pi(\pi_{(X_i,Y_i)}-\pi;\alpha,c)$ have zero-means for any value of (α,c). Define the following notations:

$$U_{\alpha,c}(X_i,Y_i)=\begin{bmatrix}\Phi'_\pi(\pi_{(X_i,Y_i)}-\pi;\alpha,c)+\Phi(\pi;\alpha,c)\\ \Psi'_\pi(\pi_{(X_i,Y_i)}-\pi;\alpha,c)+\Psi(\pi;\alpha,c)\end{bmatrix}, \qquad \text{(A.4)}$$

$\dot{U}_{\alpha,c}(X_i,Y_i)=\partial U_{\alpha,c}(X_i,Y_i)/\partial(\alpha,c)$ and $A=E[\dot{U}_{\alpha_0,c_0}(X,Y)]$. From (A.1) and (A.2), we have $\sum_{i=1}^{n}U_{\hat{\alpha},\hat{c}}(X_i,Y_i)/n=o_p(n^{-1/2})$. This formula implies that $(\hat{\alpha},\hat{c})$ is an approximate Z-estimator (Van Der Vaart, 1998, p.46) for the criterion function $(\alpha,c)\mapsto U_{\alpha,c}(x,y)$. The consistency of $(\hat{\alpha},\hat{c})$ follows by checking the two conditions: (i) The point (α_0,c_0) is the unique zero for $E[U_{\alpha,c}(X,Y)]=0$ and (ii) the set of functions $\Im=\{U_{\alpha,c}(\cdot,\cdot);(\alpha,c)\in\Theta\}$ is Glivenko-Cantelli (Van Der Vaart, 1999, p.46). We also obtain

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

Thus, the statement of Theorem 2 holds by letting $B = E[U_{\alpha_0,c_0}(X,Y)U_{\alpha_0,c_0}(X,Y)^T]$.

For the marginal estimators, we derive the following asymptotic linear expression:

$$n^{1/2}\begin{bmatrix}\phi_{\hat{\alpha}}\{\hat{S}_Y(t)\}-\phi_{\alpha_0}\{S_Y(t)\}\\ \phi_{\hat{\alpha}}\{\hat{F}_X(t)\}-\phi_{\alpha_0}\{F_X(t)\}\end{bmatrix}=\frac{1}{n^{1/2}}\sum_{i=1}^{n}\left\{\begin{bmatrix}L^Y{}_{\alpha_0,c_0}(X_i,Y_i;t)\\ L^X{}_{\alpha_0,c_0}(X_i,Y_i;t)\end{bmatrix}+\begin{bmatrix}h_Y(t)^T\\ h_X(t)^T\end{bmatrix}A^{-1}U_{\alpha_0,c_0}(X_i,Y_i)\right\}+o_p(1),$$

where

$$H^Y{}_{\alpha,c}(\pi;t)=\begin{bmatrix}\int_{y_L}^{t}\frac{\partial}{\partial\alpha}\psi\{\alpha,c;\pi(u,u)\}d\pi(\infty,u) & \int_{y_L}^{t}\frac{\partial}{\partial c}\psi\{\alpha,c;\pi(u,u)\}d\pi(\infty,u)\end{bmatrix}^T,$$

$$H^X{}_{\alpha,c}(\pi;t)=-\begin{bmatrix}\int_{t}^{x_U}\frac{\partial}{\partial\alpha}\psi\{\alpha,c;\pi(u,u)\}d\pi(u,0) & \int_{t}^{x_U}\frac{\partial}{\partial c}\psi\{\alpha,c;\pi(u,u)\}d\pi(u,0)\end{bmatrix}^T$$

and $L^Y_{\alpha,c}(X_i,Y_i;t)$ and $L^X_{\alpha,c}(X_i,Y_i;t)$ equal

$$c^2\int_{y_L}^{t}\phi''_\alpha\{c\pi(u,u)\}\{I(X_i\le u,Y_i\ge u)-\pi(u,u)\}d\pi(\infty,u)+c\int_{y_L}^{t}\phi'_\alpha\{c\pi(u,u)\}d\{I(Y_i\ge u)-\pi(\infty,u)\},$$

$$-c^2\int_{t}^{x_U}\phi''_\alpha\{c\pi(u,u)\}\{I(X_i\le u,Y_i\ge u)-\pi(u,u)\}d\pi(u,0)-c\int_{t}^{x_U}\phi'_\alpha\{c\pi(u,u)\}d\{I(X_i\le u)-\pi(u,0)\}$$

respectively. The terms in the summation are i.i.d. mean-zero stochastic process and the summation is a tight process. The notation $o_p(1)$ holds uniformly for $t\in[0,\infty)$. Then we show that

$$n^{1/2}\begin{bmatrix}\hat{S}_Y(t)-S_Y(t)\\ \hat{F}_X(t)-F_X(t)\end{bmatrix}=\frac{1}{n^{1/2}}\sum_{i=1}^{n}\begin{bmatrix}V^Y{}_{\alpha_0,c_0}(X_i,Y_i;t)\\ V^X{}_{\alpha_0,c_0}(X_i,Y_i;t)\end{bmatrix}+o_p(1), \qquad (A.5)$$

where

$$V^Y{}_{\alpha_0,c_0}(X_i,Y_i;t)=\frac{L^Y{}_{\alpha_0,c_0}(X_i,Y_i;t)}{\phi'_{\alpha_0}\{S_Y(t)\}}+\left\{\frac{h_Y(t)}{\phi'_{\alpha_0}\{S_Y(t)\}}+\begin{bmatrix}\dot{\phi}_\alpha{}^{-1}\{\phi_{\alpha_0}S_Y(t)\}\\ 0\end{bmatrix}\right\}^T A^{-1}U_{\alpha_0,c_0}(X_i,Y_i)$$

and

$$V^X_{\alpha_0,c_0}(X_i,Y_i;t)=\frac{L^X_{\alpha_0,c_0}(X_i,Y_i;t)}{\phi'_{\alpha_0}\{F_X(t)\}}+\left\{\frac{h_X(t)}{\phi'_{\alpha_0}\{F_X(t)\}}+\begin{bmatrix}\dot{\phi}_\alpha^{-1}\{\phi_{\alpha_0}F_X(t)\}\\0\end{bmatrix}\right\}^T A^{-1}U_{\alpha_0,c_0}(X_i,Y_i)$$

are mean-zero i.i.d. stochastic processes and its summation is a tight process. Let $V_n(t)=n^{1/2}(\hat{S}_Y(t)-S_Y(t),\hat{F}_X(t)-F_X(t))^T$. Also, let $G(t)=(G_X(t),G_Y(t))^T$ be a zero-mean Gaussian random field, the covariance function being specified as $E[G_Y(s)G_Y(t)]=E[V^Y_{\alpha_0,c_0}(X,Y;s)V^Y_{\alpha_0,c_0}(X,Y;t)]$, $E[G_Y(s)G_X(t)]=E[V^Y_{\alpha_0,c_0}(X,Y;s)V^Y_{\alpha_0,c_0}(X,Y;t)]$ and $E[G_X(s)G_X(t)]=E[V^X_{\alpha_0,c_0}(X,Y;s)V^X_{\alpha_0,c_0}(X,Y;t)]$ for $0\le s,t<\infty$. Based on the central limit theorem, the finite dimensional distribution of $V_n(t)$ converges weakly to that of $G(t)$ and the tightness property of $V_n(t)$, we can prove *Theorem 3.*

Appendix B: *Equivalence of Different Estimating Functions*

Let

$$B_{ij}=\{I(X_{ij}\le Z_{ij})=1\}\cap$$

$$\cap\left[\{\delta_i=\delta_j=1\}\cup\{\delta_i=1,\delta_j=0,Z_j>Z_i\}\cup\{\delta_i=0,\delta_j=1,Z_i>Z_j\}\right]$$

be the event that the pair (i,j) is orderable and comparable (Martin and Betensky, 2005). We aim to establish the following identity:

$$I=\iint_{(x,y)\in\varphi} w_{\alpha,c^*}(x,y)\left[\Delta(x,y)-\frac{\theta_\alpha\{c^*\hat{v}(x,y)\}}{R(x,y)-1+\theta_\alpha\{c^*\hat{v}(x,y)\}}\right]$$

$$=-\sum_{i<j}I\{B_{ij}\}\frac{w_{\alpha,c^*}(\breve{X}_{ij},\tilde{Z}_{ij})[1+\theta_\alpha\{c^*\hat{v}(\breve{X}_{ij},\tilde{Z}_{ij})\}]}{R(\breve{X}_{ij},\tilde{Z}_{ij})-1+\theta_\alpha\{c^*\hat{v}(\breve{X}_{ij},\tilde{Z}_{ij})\}}$$

$$\times\left[\Delta_{ij}-\frac{1}{1+\theta_\alpha\{c^*\hat{v}(\breve{X}_{ij},\tilde{Z}_{ij})\}}\right].$$

As a special case with $C_i=\infty$, the above identity yields equation (9).

The following proof is for the general situation that permits external censoring. Let

$\hat{\theta}(x,y)=\theta_\alpha\{c^*\hat{v}(x,y)\}$ and $w(x,y)=w_{\alpha,c^*}(x,y)$. Writing the integral via the finite sum, we obtain

$$I=\sum_{i=1}^{n}\sum_{j:X_i<Z_j\le Z_i,X_j<X_i}\delta_j w(X_i,Z_j)\left[N_{11}(dX_i,dZ_j)-\frac{\hat{\theta}(X_i,Z_j)}{R(X_i,Z_j)-1+\hat{\theta}(X_i,Z_j)}\right]$$

$$=\sum_{i=1}^{n}\delta_i w(X_i,Z_i)\left[1-\frac{\hat{\theta}(X_i,Z_i)}{R(X_i,Z_i)-1+\hat{\theta}(X_i,Z_i)}\right]-\sum_{i=1}^{n}\sum_{j:X_i<Z_j<Z_i,X_j<X_i}\frac{\delta_j w(X_i,Z_j)\hat{\theta}(X_i,Z_j)}{R(X_i,Z_j)-1+\hat{\theta}(X_i,Z_j)}$$

$$\equiv I_1+I_2.$$

The first term I_1 can be written as

$$\sum_{i=1}^{n}\frac{\delta_i w(X_i,Z_i)\{R(X_i,Z_i)-1\}}{R(X_i,Z_i)-1+\hat{\theta}(X_i,Z_i)}=\sum_{i=1}^{n}\sum_{j:X_i<Z_j,X_j<X_i}\frac{\delta_i(1-\Delta_{ij})w(\breve{X}_{ij},\tilde{Z}_{ij})}{R(\breve{X}_{ij},\tilde{Z}_{ij})-1+\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}.$$

The above equation follows by noting that the number of j satisfying $Z_j>Z_i,X_j<X_i$ is $R(X_i,Z_j)-1$ and using the notation $\breve{X}_{ij}$ and $\tilde{Z}_{ij}$. It is easy to see that

$$I_2=-\sum_{i=1}^{n}\sum_{j:X_i<Z_j<Z_i,X_j<X_i}\frac{\delta_j w(\breve{X}_{ij},\tilde{Z}_{ij})\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}{R(\breve{X}_{ij},\tilde{Z}_{ij})-1+\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}.$$

$$=-\sum_{i=1}^{n}\sum_{j:X_i<Z_j,X_j<X_i}\frac{\delta_j\Delta_{ij} w(\breve{X}_{ij},\tilde{Z}_{ij})\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}{R(\breve{X}_{ij},\tilde{Z}_{ij})-1+\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}.$$

By combining these terms, we have

$$I=\sum_{i=1}^{n}\sum_{j:X_i<Z_j,X_j<X_i}w(\breve{X}_{ij},\tilde{Z}_{ij})\frac{\delta_i(1-\Delta_{ij})-\delta_j\Delta_{ij}\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}{R(\breve{X}_{ij},\tilde{Z}_{ij})-1+\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}.$$

$$=-\sum_{i<j}I\{B_{ij}\}\frac{w(\breve{X}_{ij},\tilde{Z}_{ij})\{-1+\Delta_{ij}+\Delta_{ij}\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})\}}{R(\breve{X}_{ij},\tilde{Z}_{ij})-1+\hat{\theta}(\breve{X}_{ij},\tilde{Z}_{ij})}$$

$$=-\sum_{i<j}I(B_{ij})\frac{w_{\alpha,c^*}(\breve{X}_{ij},\tilde{Z}_{ij})[1+\theta_\alpha\{c^*\hat{v}(\breve{X}_{ij},\tilde{Z}_{ij})\}]}{R(\breve{X}_{ij},\tilde{Z}_{ij})-1+\theta_\alpha\{c^*\hat{v}(\breve{X}_{ij},\tilde{Z}_{ij})\}}\left[\Delta_{ij}-\frac{1}{1+\theta_\alpha\{c^*\hat{v}(\breve{X}_{ij},\tilde{Z}_{ij})\}}\right].$$

Appendix C: *Examples*

For illustration, we derive explicit formulas for the Clayton and Frank models.

C1: *Clayton model (Clayton, 1978)*

The Clayton copula is indexed by $\phi_\alpha(t) = t^{-(\alpha-1)} - 1$ ($\alpha > 0$) and, by equation (1), $\theta_\alpha(v) = \alpha$. The semi-survival Clayton model follows that

$$\Pr(X \le x, Y > y \mid X \le Y) = (1/c)[\max\{F_X(x)^{-(\alpha-1)} + S_Y(y)^{-(\alpha-1)} - 1, 0\}]^{-\frac{1}{\alpha-1}}.$$

Note that the above expression also accommodates the case of $0 < \alpha < 1$, where $\phi_\alpha(0) < \infty$ (Nelsen, 1999, p.92). By equations (2) or (5), $\theta^*(x,y) = \alpha$ but its interpretation is the reciprocal of the usual odds ratio. Hence, when $0 < \alpha < 1$, we have $\theta^*(x,y) = \alpha < 1$ which implies positive association between X and Y.

The proposed estimating function is given by

$$U_L(\alpha) = \iint_{(x,y)\in\varphi} \frac{1}{\alpha}\left[\Delta(x,y) - \frac{\alpha}{R(x,y) - 1 + \alpha}\right] = 0.$$

and, by solving $U_L(\alpha) = 0$, $\hat{\alpha}$ can be obtained without estimating c^* or c. The second estimating function $U_c(\alpha, c^*) = 0$ reduces to the explicit formula

$$c^* = \left(\left(\frac{1}{n}\right)^{1-\alpha} + \sum_{j;x_{(1)}<x_j}\left[\left\{\frac{\widetilde{R}(x_j)}{n\hat{S}_C(x_j)}\right\}^{1-\alpha} - \left\{\frac{\widetilde{R}(x_j)-1}{n\hat{S}_C(x_j)}\right\}^{1-\alpha}\right]\right)^{-\frac{1}{1-\alpha}}.$$

Plugging in $\hat{\alpha}$ in the above equation, we obtain $\hat{c}^*$. The recursive algorithm yields the following marginal estimators:

$$\hat{S}_Y(t) = \left(1 - \sum_{j;z_j\le t,\delta_j=1}\left[\left\{\hat{c}^*\frac{\widetilde{R}(z_j)}{n\hat{S}_C(z_j)}\right\}^{1-\hat{\alpha}} - \left\{\hat{c}^*\frac{\widetilde{R}(z_j)-1}{n\hat{S}_C(z_j)}\right\}^{1-\hat{\alpha}}\right]\right)^{\frac{1}{1-\hat{\alpha}}},$$

$$\hat{F}_X(t) = \left(\left(\frac{\hat{c}^*}{n}\right)^{1-\hat{\alpha}} - \sum_{j;x_{(1)}<x_j\le t}\left[\left\{\hat{c}^*\frac{\widetilde{R}(x_j)}{n\hat{S}_C(x_j)}\right\}^{1-\hat{\alpha}} - \left\{\hat{c}^*\frac{\widetilde{R}(x_j)-1}{n\hat{S}_C(x_j)}\right\}^{1-\hat{\alpha}}\right]\right)^{\frac{1}{1-\hat{\alpha}}}.$$

C2: *Frank model (Frank, 1987)*

The Frank copula is indexed by $\phi_\alpha(t) = \log\{(1-\alpha)/(1-\alpha^t)\}$ $(\alpha > 0)$ with $\theta_\alpha(v) = v\log(\alpha)/(\alpha^v - 1)$. The semi-survival Frank's model can be written as

$$\Pr(X \le x, Y > y \mid X < Y) = (1/c)\log_\alpha[1-(1-\alpha^{F_X(x)})(1-\alpha^{S_Y(y)})/(1-\alpha)].$$

It follows that

$$\theta^*(x,y) = \theta\{c^* v(x,y)\} = \{c^* v(x,y)\}\cdot\log(\alpha)/(\alpha^{c^* v(x,y)} - 1).$$

Consider the transformation $\gamma = c^*\log(\alpha)$. The likelihood estimating function can be expressed in terms of γ, and the proposed estimating function of γ is given by

$$U_L(\gamma) \propto \iint_{(x,y)\in\varphi} \hat{w}_\gamma(x,y)\left[\Delta(x,y) - \frac{\gamma\,\hat{v}(x,y)}{\{e^{\gamma\,\hat{v}(x,y)} - 1\}\{R(x,y)-1\} + \gamma\,\hat{v}(x,y)}\right],$$

where $\hat{w}_\gamma(x,y) = 1 - \gamma\,\hat{v}(x,y)e^{\gamma\,\hat{v}(x,y)}/(1-e^{\gamma\,\hat{v}(x,y)})$. Let $\hat{\gamma}$ be the solution to $U_L(\gamma) = 0$. The association parameter α can be estimated by

$$\hat{\alpha} = 1 + \left(e^{\hat{\gamma}/n} - 1\right)\prod_{j;x_{(1)}<x_j}\left[\frac{e^{\hat{\gamma}\tilde{R}(x_j)/\{n\hat{S}_C(x_j)\}} - 1}{e^{\hat{\gamma}\{\tilde{R}(x_j)-1\}/\{n\hat{S}_C(x_j)\}} - 1}\right]$$

and hence $\hat{c}^* = \hat{\gamma}/\log(\hat{\alpha})$. Explicit formula for the marginal estimators are given by

$$\hat{S}_Y(t) = \log_\alpha\left(1 + (\alpha - 1)\prod_{j;z_j\le t,\delta_j=1}\left[\frac{\alpha^{c^*\{\tilde{R}(z_j)-1\}/\{n\hat{S}_C(z_j)\}} - 1}{\alpha^{c^*\tilde{R}(z_j)/\{n\hat{S}_C(z_j)\}} - 1}\right]\right),$$

$$\hat{F}_X(t) = \log_\alpha\left(1 + (\alpha^{c^*/n} - 1)\prod_{j;x_{(1)}<x_j\le t}\left[\frac{\alpha^{c^*\tilde{R}(x_j)/\{n\hat{S}_C(x_j)\}} - 1}{\alpha^{c^*\{\tilde{R}(x_j)-1\}/\{n\hat{S}_C(x_j)\}} - 1}\right]\right).$$

References

Chaieb, L. L., Rivest, L.-P. & Abdous, B. (2006). Estimating survival under a dependent truncation. *Biometrika*, **93**, 655-69.

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**, 141-51.

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

Day, R., Bryant, J. & 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. & Chappell, R. (2001). On semi-competing risks data. *Biometrika* **88**, 907-19.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549-55

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995). A semi-parametric estimation procedure for dependence parameters in multivariate families of distributions *Biometrika* **82**, 543-52.

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

Godambe, V. P. (1991). *Estimating Functions.* Oxford: Clarendon Press.

He, S. and Yang, G. L. (1998). Estimation of the truncation probability in the random truncation model. *Annals of Statistics* **26**, 1011-27.

Kalbfleisch, J. D. & Lawless, J. F. (1989). Inference based on retrospective ascertainment: an analysis of the data on transfusion-related AIDS. *J. Am. Statist. Assoc*. **84**, 360-72.

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

Lagakos, S. W., Barraj, L. M. & De Gruttola, V. (1998). Non-parametric analysis of truncated survival data, with application to AIDS. *Biometrika* **75**, 515-23.

Lai, T. L. & Ying, Z. (1991). Estimating a distribution function with truncated and censored data. *Annals of Statistics* **19**, 417-42.

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

Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3RC quasars. *Mon. Nat. R. Astr. Soc.* **155**, 95-118.

Lagakos, S. W., Barraj, L. M. & De Gruttola, V. (1998). Non-parametric analysis of truncated survival data, with application to AIDS. *Biometrika* **75**, 515-23.

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

Nelsen, R. B. (1999). *An Introduction to copulas.* New York: Springer-Verlag.

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

Oakes, D. (1986). Semi-parametric inference in a model for association in bivariate survival data. *Biometrika*, **73**, 353-61.

Oakes, D. (1989). Bivariate survival models induced by frailties. *J. Am. Statist. Assoc.* **84**, 487-93.

Rivest, L.-P. & Wells, M. T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *J. of Mult. Annal.* **79**, 138-55.

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

Stute, W. (1993). Almost sure representation of the product-limit estimator for truncated data. *Annals of Statistics.* **21**, 146-56.

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

Van Der Vaart. A. W. (1998). *Asymptotic statistics.* Cambridge Series in Statistics and Probabilistic Mathematics. Cambridge: Cambridge University Press.

Wang, M. C., Jewell, N. P. & Tsai, W. Y. (1986). Asymptotic properties of the product-limit estimate and right censored data. *Ann. Statist*. **13**, 1597-605.

Wang, W. & Ding, A. A. (2000). On assessing the association for bivariate current status data. *Biomertika* **87**, 897-93.

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

Woodroofe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13**, 163-77.

Zheng, M. & Klein, J. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika* **82**, 127-38.

Table 1 A: Comparison of Two Estimators for the Association Parameter under the Clayton Model

(c, c^*)	$-\log(\alpha)$ (τ)	$n = 250$		$n = 500$	
		Proposed	Chaieb	Proposed	Chaieb
(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 (.44)	1.5 (1.13)	-0.3 (0.18)	2.9 (0.53)
(0.80,0.80)		0.3 (.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.

Table 1 B: Comparison of Two Estimators for the Association Parameter

under the Frank Model

(c, c^*)	$\log(\alpha)$ (τ)	$n = 250$		$n = 500$	
		Proposed	Chaieb	Proposed	Chaieb
(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.

Table 2A: The proposed estimators of marginal functions and truncation proportion under Clayton Model ($\tau = 0.5$).

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
$c / c*$	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)

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 $P_{CEN} = \Pr(C < Y \mid X \le Z)$.

Table 2A: The proposed estimators of marginal functions and truncation proportion under Frank Model ($\tau = 0.5$).

parameter	True	$n = 250$, P_{CEN} =0.00	$n = 250$, P_{CEN} =0.39	$n = 500$, P_{CEN} =0.00	$n = 500$, P_{CEN} =0.39
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 $P_{CEN} = \Pr(C < Y \mid X \le Z)$.

Table 3. Analysis of the Transfusion-Related AIDS Data

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

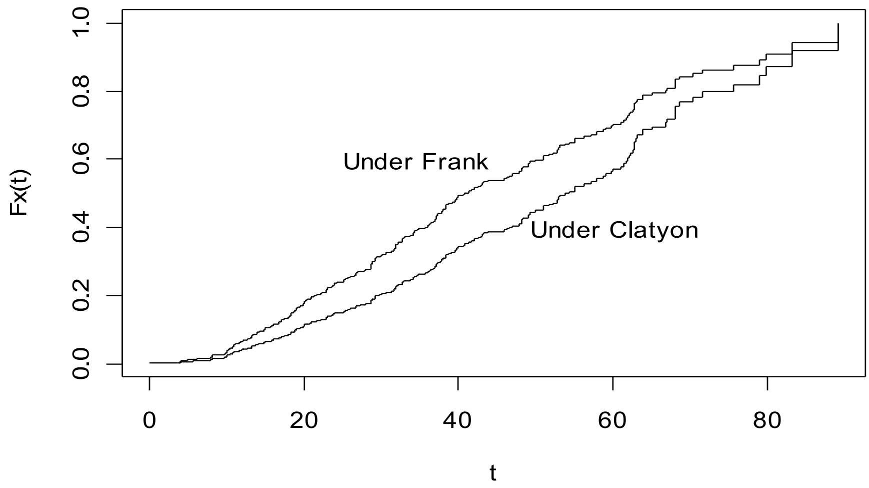

Fig. 1. The cumulative distribution functions of the incubation time of AIDS under two copula models.