Testing Quasi-independence for Truncation Data

*Takeshi Emura and **Weijing Wang[1]

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

**School of Pharmaceutical Sciences, Kitasato University, 5-9-1, Sirokane, Minato-ku, Tokyo, Japan*

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

Abstract

Quasi-independence is a common assumption for analyzing truncated data. To verify this condition, we consider a class of weighted log-rank type statistics that include existing tests proposed by Tsai (1990) and Martin and Betensky (2005) as special cases. To choose an appropriate weight function that may lead to a more power test, we derive a score test when the dependence structure under the alternative hypothesis is modeled via the odds ratio function proposed by Chaieb et al. (2006). Asymptotic analysis is established based on the functional delta method which can handle more general situations than results based on rank-statistics or U-statistics. Simulations are performed to examine finite-sample performances of the proposed method and its competitors. Two datasets are analyzed using different tests for illustrative purposes.

Key Words: Censoring; Conditional likelihood; Kendall's tau; Mantel-Heanszel test; Power; Survival data; Two-by-two table.

---

[1] Takeshi Emura is Assistant Professor, Division of Biostatistics, School of Pharmaceutical Sciences, Kitasato University, 5-9-1, Sirokane, Minato-ku, Japan, and Weijing Wang is Professor, Institute of Statistics, National Chiao Tung University, Hsin-Chu 300, Taiwan, R.O.C. The research was financially supported by the research grant 95-2118-M-009-005 funded by National Science Council of Taiwan, R.O.C.

1. INTRODUCTION

Truncated data are commonly seen in studies of biomedicine, epidemiology, astronomy and econometrics. Such data occur when the lifetime variables of interest can be observed if their values satisfy certain criteria. In this article, we discuss the situation that a pair of lifetime variables (X, Y) can be included in the sample only if $X \leq Y$. The variable Y is said to be left truncated by X and X is right truncated by Y.

For example, in the study of transfusion-related AIDS discussed in Lagakos, Barraj and Cruttola (1988), infected people could be included in the sample only if they developed AIDS within the study period. Accordingly the incubation time X was subject to right truncation by the lapse time Y measured from infection to the recruitment time. In this design, a subject with the incubation time exceeding the lapse time was completely missing. Consider another example that studied the survival time for residents in the Channing House retirement community in Palo Alto, California (Hyde 1977, 1980; Klein and Moeschberger 2003). To make fair comparison with the general population, statistical methods had to account for the fact that only those who had lived long enough to enter the retirement center could be observed. Hence the lifetime Y was left truncated by the entry age X. Notice that a truncated subject with $X > Y$ is completely missing and even its existence is unknown. Therefore truncated data are fundamentally different from censored data in which partial information of censored individuals is still available.

Inference methods for analyzing truncated data require making some assumption about the relationship between X and Y. Independence between the two variables was commonly assumed (Lynden-Bell 1971; Hyde 1977; Woodroofe 1985; Wang, Jewell and Tsai 1986; Lagakos et al. 1988). This assumption was later relaxed by Tsai (1990) to a weaker condition of quasi-independence which can be formulated as follows:

$$H_0 : \pi(x, y) = F_X(x) S_Y(y) / c_0 \quad (x \leq y), \tag{1}$$

where $\pi(x, y) = \Pr(X \le x, Y > y \mid X \le Y)$ and F_X and S_Y are right continuous distribution and survival functions of X and Y respectively, and c_0 is the constant satisfying $c_0 = -\iint_{x \le y} dF_X(x)dS_Y(y)$. Furthermore Tsai (1990) demonstrated that H_0 is a testable assumption.

In applications, external censoring may also arise. In the Channing House example, the lifetime Y was further subject to right censoring due to residents' withdrawal or the end-of-study effect. Let C be the censoring variable. With left-truncated and right-censored data, one observes X, $Z = Y \wedge C$ and $\delta = I(Y \le C)$ subject to $X \le Z$, where $a \wedge b = \min(a, b)$ and $I(\cdot)$ is the indicator function. Assume that C is independent of (X, Y). In this article we consider testing quasi-independence based on left-truncated and right-censored data of the form $\{(X_i, Z_i, \delta_i)\,(i = 1,...,n)\}$, a random sample of (X, Z, δ) .

Several methods for testing H_0 have been proposed. Tsai (1990) defined a conditional version of Kendall's tau and used its empirical estimator as the basis for constructing the testing procedure. Martin and Betensky (2005) extended the idea of Tsai (1990) to more complicated truncation structures and then applied the properties of U-statistics in variance estimation and large-sample analysis. Chen, Tsai and Chao (1996) constructed their test based on a conditional version of Pearson correlation coefficient.

Here we propose weighted log-rank type tests based on a series of 2×2 tables suitable for describing truncated data. By establishing a nice mathematical relationship between the proposed statistics and the conditional Kendall's tau statistics, the tests of Tsai (1990) and Martin and Betensky (2005) can be viewed as our special cases with different forms of weight. Accordingly choosing an appropriate weight function becomes the key issue for selecting a more powerful test. To attain this goal, we derive a score test when the dependence structure under the alternative hypothesis can be described by the odds ratio function proposed by Chaeib, Rivest and Abdous (2006). Specifically we suggest utilizing the likelihood information provided by the 2×2 tables to select a suitable model for the alternative hypothesis. For variance estimation and theoretical analysis, we adopt the functional delta method which is a more powerful tool than the techniques based on rank-statistics or U-statistics since it can handle flexible weight functions. Furthermore we provide both analytic and jackknife variance estimators and compared their finite-sample performances via simulations.

The paper is organized as follows. In Section 2, we propose the main methodology by temporarily ignoring censoring. In Section 3, we derive the score test and suggest a model selection method. Large sample properties are examined in Section 4. Modifications of all the results to account for the presence of right censoring are presented in Section 5. Section 6 contains numerical analysis including data analysis and simulation studies. Concluding remarks are given in Section 7.

2. THE PROPOSED METHOD WITHOUT CENSORING

To illustrate the main idea, we temporarily ignore external censoring by letting $C=\infty$. Observed data can be expressed as $\{(X_j,Y_j)\ (j=1,\ldots,n)\}$ subject to $X_j\le Y_j$.

2.1 Constructing the Test Statistics based on 2×2 Tables

Adapt to the nature of truncation, we can construct the following 2×2 table at an observed failure point (x,y) for $x\le y$.

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

Table 1: 2×2 table for truncated data without censoring

The cell counts and marginal counts in Table 1 are defined as

$$N_{11}(dx,dy)=\sum_j I(X_j=x,Y_j=y),\quad N_{\bullet 1}(x,dy)=\sum_j I(X_j\le x,Y_j=y),$$

$$N_{1\bullet}(dx,y)=\sum_j I(X_j = x, Y_j \ge y),\quad R(x,y)=\sum_j I(X_j \le x, Y_j \ge y),$$

Under H_0 and given the marginal counts, the conditional mean of $N_{11}(dx,dy)$ for $x \le y$ becomes

$$\mathrm{E}(N_{11}(dx,dy)\,|\,N_{1\bullet},N_{\bullet 1},R)=\frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)}. \quad (2)$$

To test quasi-independence, we propose the following weighted log-rank type statistics:

$$L_W=\iint_{x\le y} W(x,y)\left\{N_{11}(dx,dy)-\frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)}\right\}, \quad (3)$$

where $W(x,y)$ is a weight function. Motivated by the G^{ρ} class discussed in Harrington and Fleming (1982), we consider a sub-class of L_W with a particular form of $W(x,y)$ which can be written as

$$L_\rho=\iint_{x\le y} \hat{\pi}(x,y-)^{\rho}\left\{N_{11}(dx,dy)-\frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)}\right\}, \quad (4)$$

where $\hat{\pi}(x,y)=\sum_j I(X_j \le x, Y_j > y)/n$ and $\rho \in [0,\infty)$ is a pre-specified constant.

The statistics L_W is nonparametric in the sense that no distributional assumption about the joint distribution of (X,Y) is made. However such information would be helpful for choosing an appropriate weight or the value of ρ in (4) which may lead to a more powerful test. In Section 3, we derive a score test which utilizes the information of the underlying association structure provided by the 2×2 tables.

2.2 Relationship with Existing Tests

The tests proposed by Tsai (1990) and Martin and Betensky (2005) are both related to a conditional version of Kendall's tau defined as

$$\tau_a = E\{\mathrm{sgn}(X_i - X_j)(Y_i - Y_j)\,|\,A_{ij}\},$$

where $\mathrm{sgn}(x)$ is defined to be -1, 0, or 1 if $x<0$, $x=0$, or $x>0$ respectively and

$A_{ij} = I\{\breve{X}_{ij} \le \tilde{Y}_{ij}\}$, $\breve{X}_{ij} = X_i \vee X_j$ and $\tilde{Y}_{ij} = Y_i \wedge Y_j$. Note that when event A_{ij} occurs, the (i, j) pairs are both located in the observable region, $\{(x, y) : 0 < x \le y < \infty\}$ and hence τ_a is well-defined for truncated data. Under quasi-independence, Tsai (1990) showed that $\tau_a = 0$.

Based on the idea of using an empirical estimator of the conditional Kendall's tau measure for testing H_0, Tsai (1990) and Martin and Betensky (2005) both considered the statistics

$$K = \sum_{i<j} I\{A_{ij}\} \operatorname{sgn}\{(X_i - X_j)(Y_i - Y_j)\}. \tag{5}$$

Their tests differ in the way of calculating the variance of K. For example in absence of ties, by writing K as the sum of conditionally independent rank variables, Tsai (1990) was able to utilize rank-based results to derive the conditional variance of K explicitly. Martin and Betensky (2005) use the fact that K is a U-statistic to derive the asymptotic variance formula which can handle tied data. Another nice advantage of the statistic K is that it can be extended to account for external censoring (Tsai 1990) or even more complicated data structures (Martin and Betensky 2005).

Now we compare the proposed log-rank type test in (3) with the (conditional Kendall's tau) statistic K in (5). Consider the situation that the data have no ties so that the values of $X_1, \ldots X_n, Y_1, \ldots, Y_n$ are all distinct. Note that in such a case $N_{\bullet 1}(x, dy) = N_{1\bullet}(x, dy) = 1$ for all tables of interest and the expected value in (2) becomes $1/R(x, y)$. It can be shown that

$$L_W = -\sum_{i<j} I\{A_{ij}\} \frac{W(\breve{X}_{ij}, \tilde{Y}_{ij})}{R(\breve{X}_{ij}, \tilde{Y}_{ij})} \operatorname{sgn}\{(X_i - X_j)(Y_i - Y_j)\}. \tag{6}$$

The proof of the above equation is given in Appendix B under a more general setting that allows for right-censoring. By setting $W(x, y) = R(x, y)/n$, we get

$$L_{\rho=1} = \iint_{x \le y} \frac{R(x,y)}{n} \left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y) N_{\bullet 1}(x,dy)}{R(x,y)} \right\} = -\frac{K}{n} . \tag{7}$$

Equation (6) implies that L_W is also a U-statistic if $W(x,y)/R(x,y)$ is a deterministic function. However if we prefer a flexible weight function that can lead to a more powerful test, the technique of U-statistics is no longer applicable for variance estimation and large sample analysis. In Section 4, we will use the functional delta method to establish asymptotic properties of L_W .

3. CONDITIONAL SCORE TEST

3.1 Construction of Conditional Likelihood

Without loss of generality, we assume that $\pi(x,y) = \Pr(X \le x, Y > y \mid X \le Y)$ is differentiable and hence the data have no ties. Oakes (1989) defined an odds ratio function for describing the dependent relationship between bivariate failure times. Chaieb et al. (2006) proposed a modified measure suitable for truncated data as follows:

$$\vartheta(x,y) = \frac{\pi(x,y) \cdot \partial^2 \pi(x,y) / \partial x \partial y}{\partial \pi(x,y)/\partial x \cdot \partial \pi(x,y)/\partial y} = \frac{\lambda_Y(y \mid X = x)}{\lambda_Y(y \mid X < x)},$$

where $x \le y$ and $\lambda_Y(y \mid A)$ is the hazard function of Y given that event A occurs. Note that $\vartheta(x,y)$ can be viewed as the hazard ratio of Y at time y based on different truncation ranges of x. Under quasi-independence, $\vartheta(x,y) = 1$ for all $0 < x \le y$. It should be noted that the case of $\vartheta(x,y) < 1$ implies positive association while $\vartheta(x,y) > 1$ implies negative association between the two truncated variables.

The information of $\vartheta(x,y)$ is contained in Table 1. Given the marginal counts, $N_{11}(dx,dy)$ follows a Bernoulli distribution with

$$\Pr(N_{11}(dx,dy) = 1 \mid N_{1\bullet} = N_{\bullet 1} = 1, R = r) = \frac{\vartheta(x,y)}{r - 1 + \vartheta(x,y)} .$$

This distributional result can be further utilized to construct a score test. Here we assume that $\vartheta(x,y)$ can be formulated as follows:

(i) The odds ratio function can be parameterized as $\vartheta(x,y) = \theta_\alpha\{\eta(x,y)\}$*, where* α

is a parameter and $\eta(x,y)$ *is an unspecified nuisance function which can be estimated separately from* α *.*

(ii) *For each fixed* η*,* $\theta_\alpha(\eta)$ *is a continuously differentiable function of* α *and* $\lim_{\alpha\to\alpha_0}\theta_\alpha(\eta)=1$*, where* α_0 *is the parameter value under quasi-independence.*

Suppose that $\eta(x,y)$ can be estimated by $\hat{\eta}(x,y)$. Under a working assumption of independence among different tables of (x,y) and ignoring the distributions of the marginal counts, we can construct the following conditional likelihood function:

$$L(\alpha)=\prod_{x\le y}\left[\frac{\theta_\alpha\{\hat{\eta}(x,y)\}}{R(x,y)-1+\theta_\alpha\{\hat{\eta}(x,y)\}}\right]^{N_{11}(dx,dy)}\left[\frac{R(x,y)-1}{R(x,y)-1+\theta_\alpha\{\hat{\eta}(x,y)\}}\right]^{1-N_{11}(dx,dy)}. \quad (8)$$

The corresponding score function becomes

$$\frac{\partial\log L(\alpha)}{\partial\alpha}=\iint_{x\le y}\frac{\dot{\theta}_\alpha\{\hat{\eta}(x,y)\}}{\theta_\alpha\{\hat{\eta}(x,y)\}}\left\{N_{11}(dx,dy)-\frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)\theta_\alpha\{\hat{\eta}(x,y)\}}{R(x,y)-1+\theta_\alpha\{\hat{\eta}(x,y)\}}\right\}, \quad (9)$$

where $\dot{\theta}_\alpha(v)=\partial\theta_\alpha(v)/\partial\alpha$. Note that equation (8) was motivated by Clayton (1978) and Oakes (1986) who considered the Clayton model for bivariate censored data.

By setting $\alpha\to\alpha_0$, the score test statistic can be obtained based on equation (9). Specifically since $\lim_{\alpha\to\alpha_0}\theta_\alpha\{\eta(x,y)\}=1$, the proposed score statistics has the form of L_W with the weight

$$W(x,y)=\lim_{\alpha\to\alpha_0}\dot{\theta}_\alpha\{\hat{\eta}(x,y)\}. \quad (10)$$

Equation (10) provides a clear guideline for choosing the weight function for L_W when the assumptions on $\vartheta(x,y)$ stated in (i) and (ii) are satisfied. The level of power improvement depends on whether $\theta_\alpha(\cdot)$ is correctly specified and how accurate $\eta(x,y)$ can be estimated. We will discuss these issues via specific examples in Section 3.2.

3.2 Semi-survival Archimedean Copula Models

For dependent truncation data, Chiaeb et al. (2006) proposed "semi-survival"

Archimedean copula (AC) models of the form

$$\pi(x,y)=\Pr(X\le x,Y>y\mid X\le Y)=\phi_\alpha^{-1}[\{\phi_\alpha\{F_X(x)\}+\phi_\alpha\{S_Y(y)\}]/c\,, \qquad (11)$$

where c is a normalizing constant satisfying

$$c=\iint_{x\le y}-\frac{\partial^2}{\partial x\partial y}(\phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\}+\phi_\alpha\{S_Y(y)\}])\,dxdy\,.$$

AC models are characterized by the generating function $\phi_\alpha(\cdot):[0,1]\to[0,\infty]$, where $\phi_\alpha(1)=0$, $\phi'_\alpha(t)=\partial\phi_\alpha(t)/\partial t<0$ and $\phi''_\alpha(t)=\partial^2\phi_\alpha(t)/\partial t^2>0$. For an AC model, one can write $\vartheta(x,y)=\theta_\alpha\{c\pi(x,y)\}$, where

$$\theta_\alpha(\eta)=-\eta\frac{\phi''_\alpha(\eta)}{\phi'_\alpha(\eta)}\,. \qquad (12)$$

Hence AC models satisfy assumption (i) such that $\eta(x,y)=c\pi(x,y)$. The case of quasi-independence corresponds to $\phi(t)=-\log(t)$ in (11). After appropriate parameterization, we may assume that $\phi_{\alpha_0}(t)=-\log(t)$ and $\theta_{\alpha_0}(\eta)=1$ for $\alpha_0=1$.

Estimation of $\eta(x,y)=c\pi(x,y)$ can be handled separately from α. The function $\pi(x,y)$ can be estimated by the empirical estimator $\hat{\pi}(x,y)=\sum_j I(X_j\le x,Y_j>y)/n$. Under H_0, the constant c can be estimated by

$$\hat{c}=\frac{n}{R(X_{(1)},X_{(1)})}\prod_{j:X_{(1)}<X_j}\left\{1-\frac{\sum_k I(X_k=X_j)}{R(X_j,X_j)}\right\},$$

where $X_{(1)}=\min_j X_j$ (He and Yang 1998).

In practice there may be several model choices under consideration. We suggest to choose the model that yields the highest value of $L(\hat{\alpha})$, where $\hat{\alpha}$ maximizes $L(\alpha)$ over the corresponding parameter space. Now we derive the suggested weight function in (10) for selected AC models. The influence of the weight function on power of the corresponding test will be evaluated later via simulations.

Example 1. Clayton copula

Clayton's model (1978) has the generating function $\phi_\alpha(t) = t^{-(\alpha-1)} - 1$ for $0 < \alpha < \infty$, $\alpha \neq 1$, and $\phi_{\alpha_0}(t) = -\log(t)$ when $\alpha_0 = 1$. It follows that $\theta_\alpha(\eta) = \alpha$ and hence it is easy to see that

$$\lim_{\alpha \to \alpha_0} \dot{\theta}_\alpha \{\eta(x, y)\} = 1,$$

which corresponds to $L_{\rho=0}$, a special case of L_ρ in (4). Notice that no nuisance parameter is involved.

Example 2. Frank copula

For Frank's model (Genest 1986), the generating function has the form $\phi_\alpha(t) = \log\{(1-\alpha)/(1-\alpha^t)\}$ for $0 < \alpha < \infty$, $\alpha \neq 1$ and $\phi_{\alpha_0}(t) = -\log(t)$ for $\alpha_0 = 1$. Since $\theta_\alpha(\eta) = \eta \log(\alpha)/\{e^{\eta \log(\alpha)} - 1\}$, we have

$$\lim_{\alpha \to \alpha_0} \dot{\theta}_\alpha \{\eta(x, y)\} \propto \pi(x, y).$$

If we estimate $\pi(x, y)$ by $\hat{\pi}(x, y-) = R(x, y)/n$, the resulting score test becomes $L_{\rho=1}$ in (7) which is equivalent to the test K considered by Tsai (1990) and Martin and Betensky (2005). This implies that these two tests are suitable for Frank's alternative.

Example 3. Gumbel copula

For the Gumbel model, the generating function equals $\phi_\alpha(t) = \{-\log(t)\}^\alpha$ for $\alpha > 1$ and $\phi_{\alpha_0}(t) = -\log(t)$ for $\alpha_0 = 1$. Under the Gumbel model, (X, Y) can only permit negative association. Since $\theta_\alpha(\eta) = 1 - (\alpha - 1)/\log(\eta)$, it follows that

$$\lim_{\alpha \to \alpha_0} \dot{\theta}_\alpha \{\eta(x, y)\} \propto -1/\log\{c\pi(x, y)\}.$$

By plugging in the estimators of $\pi(x, y)$ and c in the suggested weight, we denote the corresponding test as $L_{inv\log}$, which however is not a member of L_ρ in (4).

4. ASYMPTOTIC ANALYSIS

4.1 Asymptotic Normality

To simplify the analysis and without loss of generality, we assume that all the underlying distributions are absolutely continuous under the null hypothesis in (1). In this section, we state the main theoretical results. Sketch of the proofs are given in Appendix A and more details can be found in the technical report.

We will examine asymptotic properties of the following two statistics:

$$L_w = \iint_{x \le y} w\{\hat{\pi}(x, y-)\}\left\{ N_{11}(dx, dy) - \frac{N_{1\bullet}(dx, y) N_{\bullet 1}(x, dy)}{R(x, y)} \right\}, \tag{13}$$

$$L_w^* = \iint_{x \le y} w\{\hat{c}\hat{\pi}(x, y-)\}\left\{ N_{11}(dx, dy) - \frac{N_{1\bullet}(dx, y) N_{\bullet 1}(x, dy)}{R(x, y)} \right\}, \tag{14}$$

where $w(v)$ is a known continuously differentiable function on $v \in (0,1)$. Notice that L_w and L_w^* differ in whether $\hat{c}$, the estimator of c_0, is involved. The formula in (13) and (14) can be re-expressed as the following functional forms:

$$L_w = -\frac{n}{2} \iint\iint_{x \vee x^* \le y \wedge y^*} \frac{w\{\hat{\pi}(x \vee x^*, y \wedge y^* -)\}}{\hat{\pi}(x \vee x^*, y \wedge y^* -)} \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\hat{\pi}(x, y) d\hat{\pi}(x^*, y^*),$$

$$L_w^* = -\frac{n}{2} \iint\iint_{x \vee x^* \le y \wedge y^*} \frac{w\{g(\hat{\pi})\hat{\pi}(x \vee x^*, y \wedge y^* -)\}}{\hat{\pi}(x \vee x^*, y \wedge y^* -)} \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\hat{\pi}(x, y) d\hat{\pi}(x^*, y^*)$$

where $g(\cdot)$ is a functional such that $\hat{c} = g(\hat{\pi})$, defined in Appendix A.3. The functionals L_w and L_w^* can be shown to be Hadamard differentiable functions of $\hat{\pi}$ given the differentiability of $w(\cdot)$. By applying the functional delta method (Van Der Vaart 1998, p. 297), we obtain the following asymptotic expression:

$$n^{-1/2} L_w = -n^{-1/2} \sum_j U(X_j, Y_j) + o_P(1),$$

where the random variable $U(X_j, Y_j)$ is defined in Appendix A.1.

Theorem 1: Under H_0, $n^{-1/2} L_w$ converges weakly to a mean-zero normal distribution with the variance $\sigma^2 = E[U(X_j, Y_j)^2]$.

Corollary 1: Under H_0 $n^{-1/2}L_\rho$, *the special case of* $n^{-1/2}L_w$ *with* $w(v)=v^\rho$, *converges weakly to a mean-zero normal distribution with the variance* $E[U_\rho(X_j,Y_j)^2]$, *where* $U_\rho(X_j,Y_j)$ *is defined in Appendix A.2*

Similarly it follows that

$$n^{-1/2}L_w^* = -n^{-1/2}\sum_j U^*(X_j,Y_j)+o_P(1),$$

where $U^*(X_j,Y_j)$ is defined in Appendix A.3. Note that L_w^* involves the estimator of the constant $c_0=-\iint_{x\le y}dF_X(x)dS_Y(y)$, which is closely related to the marginal estimators of $F_X(t)$ and $S_Y(t)$. To establish asymptotic normality, we need the following condition:

Identifiability Assumption (I): There exists two positive numbers $y_L<x_U$ *such that*

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

The above statement 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). Assumption (I) also guarantees the condition that $R(t,t)/n$ is away from zero asymptotically (Chaieb et al. 2006) so that $\hat{c}$ is a consistent estimator.

Theorem 2: Under H_0 *and Assumption (I),* $n^{-1/2}L_w^*$ *converges weakly to a mean-zero normal distribution with the variance* $\sigma_*^2=E[U^*(X_j,Y_j)^2]$.

4.2 Variance Estimation

Equation (7) shows that, in absence of ties, $L_{\rho=1}$ is equivalent to K. Variance estimation of K has been proposed by Tsai (1990) and Martin and Betensky (2005). Here we discuss variance estimation when a general form of $W(x,y)$ is employed. The

L_ρ statistics forms a nice subset of L_W since its asymptotic variance defined by $E[U_\rho(X_j, Y_j)^2]$ has a tractable form. Based on the method of moment and applying the plug-in principle, we obtain the following estimator of the asymptotic variance for L_ρ:

$$\sum_j \Bigg[\frac{1}{n}\sum_k I\{A_{jk}\}\hat{\pi}(\breve{X}_{jk}, \tilde{Y}_{jk}-)^{\rho-1} \operatorname{sgn}\{(X_j - X_k)(Y_j - Y_k)\} + \frac{(\rho+1)L_\rho}{n}$$
$$+\frac{\rho-1}{n^2}\sum_{k<l} I\{A_{kl}\}\hat{\pi}(\breve{X}_{kl}, \tilde{Y}_{kl}-)^{\rho-2} \operatorname{sgn}\{(X_k - X_l)(Y_k - Y_l)\} I(X_j \le \breve{X}_{kl}, Y_j \ge \tilde{Y}_{kl}) \Bigg]^2 .$$

The derivation of the above formula is given in Appendix A.2.

However in general analytic expressions of σ^2 and σ_*^2 are quite complicated especially when censoring is involved. In such a case the jackknife method provides a convenient tool for variance estimation. The variance of L_W can be estimated by the following jackknife estimator:

$$\frac{n-1}{n}\sum_j (L_W^{(-j)} - L_W^{(\cdot)})^2 ,$$

where $L_W^{(-j)}$ is the statistics L_W ignoring the jth observation and $L_W^{(\cdot)} = (1/n)\sum_j L_W^{(-j)}$. In the technical report, we provide simulation results which compare the two variance estimators. It is found that although the analytic estimator sometimes has better performance (with smaller MSE) in variance estimation, it tends to yield less accurate type-I probability compared with the jackknife estimator. It seems that the higher-order terms omitted in the linear expression of L_ρ still play some role in estimating the variance for finite samples.

The validity of the jackknife estimator is closely related to the smoothness of the corresponding functional expression. Unfortunately, Hadamard differentiability of L_w and L_w^* alone does not ensure consistency of the jackknife estimator. This property requires a more stringent smoothness condition on the corresponding statistical functional. The following theorems provide theoretical justification for the jackknife variance estimator:

Theorem 3: Under H_0, the asymptotic variances σ^2 of L_w can be consistently estimated by the jackknife method.

Theorem 4: Under H_0 and Assumption (I), the asymptotic variances σ_*^2 of L_w^* can be consistently estimated by the jackknife method.

The above theorems can be proved by checking a sufficient condition of continuous Gateaux differentiability (Shao 1993) for consistency of the jackknife method. The detailed proof, in which continuous differentiability of the function $w(\cdot)$ plays an essential role, is given in the technical report.

5. MODIFICATION FOR RIGHT CENSORING

In presence of censoring, observed data can be written as $\{(X_i, Z_i, \delta_i)\ (i = 1, ..., n)\}$.

5.1 The Proposed Test Statistic under Censoring

Table 2 is a modification of Table 1 such that (x, y) denotes an uncensored failure point satisfying $x \le y$. To simplify the presentation, we use the same notations as before but modify their definitions as follows.

$$N_{11}(dx, dy) = \sum_j I(X_j = x, Z_j = y, \delta_j = 1),\ N_{1\bullet}(dx, y) = \sum_j I(X_j = x, Z_j \ge y),$$

$$N_{\bullet 1}(x, dy) = \sum_j I(X_j \le x, Z_j = y, \delta_j = 1) \text{ and } R(x, y) = \sum_j I(X_j \le x, Z_j \ge y).$$

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

Table 2: 2×2 table for truncated data subject to right censoring

After the modification, the proposed log-rank statistics still has the same form,

$$L_W = \iint_{x \le y} W(x,y)\left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)} \right\},$$

but now the L_ρ statistic under censoring becomes

$$L_\rho = \iint_{x \le y} \hat{v}(x,y-)^\rho \left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)} \right\}, \tag{15}$$

where $\rho \in [0,\infty)$ is a constant, $\hat{v}(x,y) = (1/n)\sum_j I(X_j \le x, Z_j > y) / \hat{S}_C(y)$ is a nonparametric estimator of $\pi(x,y)$ and $\hat{S}_C(y)$ is the Lynden-Bell's estimator for $\Pr(C > y) = S_C(y)$ based on data $\{(X_i, Z_i, 1-\delta_i) \quad (i = 1,\ldots,n)\}$.

Now we re-express L_W in terms of the conditional Kendall's tau statistic based on pairwise comparison. In presence of censoring, the order of a pair is known for certain if the smaller one is observed. Martin and Betensky (2005) defines the event

$B_{ij} = \{\breve{X}_{ij} \le \widetilde{Z}_{ij}\} \cap \{(\delta_i = \delta_j = 1) \cup (Z_j - Z_i > 0 \,\&\, \delta_i = 1 \,\&\, \delta_j = 0) \cup (Z_i - Z_j > 0 \,\&\, \delta_i = 0 \,\&\, \delta_j = 1)\}$

which is a condition for the (i,j) pairs to be comparable and orderable. The conditional Kendall's tau under censoring, denoted as τ_b, has the same form as τ_a with A_{ij} being replaced by B_{ij}. Under quasi-independence, $E[\text{sgn}\{(X_i - X_j)(Z_i - Z_j)\} \mid B_{ij}] = 0$. In Appendix B, we show that

$$L_W = -\sum_{i<j} I\{B_{ij}\} \frac{W(\breve{X}_{ij}, \widetilde{Z}_{ij})}{R(\breve{X}_{ij}, \widetilde{Z}_{ij})} \text{sgn}\{(X_i - X_j)(Z_i - Z_j)\}. \tag{16}$$

By choosing $W(x,y) = R(x,y)/n$, L_W reduces to

$$-\frac{1}{n}\sum_{i<j} I\{B_{ij}\} \text{sgn}\{(X_i - X_j)(Z_i - Z_j)\} = -\frac{K}{n}, \tag{17}$$

where K is Tsai's statistics for censored case. Note that K in (17) no longer belongs to the class L_ρ in (15) when data are censored. For variance estimation, Tsai (1990) expressed K as the sum of independent rank variables, conditionally on the number at risk at observed failure points and then obtained an explicit variance formula for K. Martin and Betensky (2005) still apply properties of U-statistics to obtain the asymptotic variance of K in (17).

5.2 Conditional Score Test under Censoring

Under censoring, derivations of the score test are the same as those discussed in Section 3.1 and 3.2 based on the modified counts defined in Section 5.1. The key change is the definition of η which appears in the formula $\theta_\alpha(\eta) = -\eta\phi''_\alpha(\eta)/\phi'_\alpha(\eta)$. In presence of censoring, we define $\eta(x,y) = c^* v(x,y)$, where $v(x,y) = \Pr(X \le x, Z > y \mid X \le Z)/S_C(y)$ and $c^* = \Pr(X \le Z)$. One can estimate $v(x,y)$ by $\hat{v}(x,y)$ and c^* by the following standard estimator:

$$\hat{c}^* = \frac{n}{R(X_{(1)}, X_{(1)})} \prod_{j:X_{(1)}<X_j} \left\{ 1 - \frac{\sum_k I(X_k = X_j)}{R(X_j, X_j)} \right\}, \tag{18}$$

where $X_{(1)} = \min_j X_j$. The proposed score test is still a special case of L_W with weight given in (10) and $\hat{\eta}(x,y) = \hat{c}^* \hat{v}(x,y)$.

5.3 Asymptotic Analysis under Censoring

Consider the two statistics:

$$L_w = \iint_{x \le y} w\{\hat{v}(x,y-)\} \left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y) N_{\bullet 1}(x,dy)}{R(x,y)} \right\}, \tag{19}$$

$$L_w^* = \iint_{x \le y} w\{\hat{c}^* \hat{v}(x,y-)\} \left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y) N_{\bullet 1}(x,dy)}{R(x,y)} \right\}. \tag{20}$$

It can be shown that both L_w and L_w^* are Hadamard differentiable functionals of the empirical process, $\hat{H}(x,y,c) = \sum_j I(X_j \le x, Y_j > y, C_j > c)/n$. Asymptotic normality of the two statistics can be established by applying the functional delta method and the fact that $n^{1/2}(\hat{H} - H)$ converges weakly to a mean-zero Gaussian process, Specifically in

Appendix A.3, we can show that

$$L_w = -\frac{n}{2}\iint\iint\iint_{x\vee x^*\le y\wedge y^*<c\wedge c^*} \frac{w\{\varphi(\hat{H}; x\vee x^*, y\wedge y^*\wedge c\wedge c^*)\}}{\hat{H}(x\vee x^*, y\wedge y^*\wedge c\wedge c^*-, y\wedge y^*\wedge c\wedge c^*-)}, \\ \times \operatorname{sgn}\{(x-x^*)(y-y^*)\} d\hat{H}(x,y,c) d\hat{H}(x^*,y^*,c^*)$$

$$L_w^* = -\frac{n}{2}\iint\iint\iint_{x\vee x^*\le y\wedge y^*<c\wedge c^*} \frac{w\{g^*(\hat{H})\varphi(\hat{H}; x\vee x^*, y\wedge y^*\wedge c\wedge c^*)\}}{\hat{H}(x\vee x^*, y\wedge y^*\wedge c\wedge c^*-, y\wedge y^*\wedge c\wedge c^*-)}, \\ \times \operatorname{sgn}\{(x-x^*)(y-y^*)\} d\hat{H}(x,y,c) d\hat{H}(x^*,y^*,c^*)$$

where $\varphi(\cdot; x, y)$ and $g^*(\cdot)$ are functions such that $\hat{v}(x, y-) = \varphi(\hat{H}; x, y)$ and $\hat{c}^* = g^*(\hat{H})$, each of which is defined in Appendix A.3. Due to the complexity of the related formulas under censoring, we only provide a brief sketch of the proof.

Similar to the uncensored case, consistency of the jackknife estimator can be proved by checking the continuous Gateaux differentiability of the corresponding functional. The proof follows the same lines as that for Theorem 3 and hence is omitted.

6. NUMERICAL ANALYSIS

6.1 Data Analysis

We apply the proposed methods to the aforementioned AIDS data and Channing House data and compare our results with existing analyses.

Lagakos et al. (1988) divided the AIDS data into two age groups of “children” (37 subjects) and adults (258 subjects) and assumed independence between the incubation time X and the lapse time Y. Five tests for testing quasi-independence are applied in absence of censoring. Specifically the proposed log-rank statistics based on $L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$ utilize the jackknife method for variance estimation. The tests proposed by Tsai (1990) and Martin and Betensky (2005) have the form of $L_{\rho=1}$ or K but use their own variance estimators in the standardization. Note that, to handle ties for computing the maximized log-likelihoods and Tsai’s variance estimator, we adopt the Breslow-Peto correction (Breslow 1972; Peto 1972). The Z-values and p-values of the five tests are reported in Table 3. All the results show significant deviation from quasi-independence in the adult group. The sign of the Z-values indicates positive association between X and Y ($\tau_a = 0.111$). This implies that people infected in earlier chronicle time tended to have longer length of incubation. Although similar pattern of association was also discovered in the children group ($\tau_a = 0.117$), it did not reach 5% level of statistical significance since the sample size is much smaller. Nevertheless H_0 is still rejected at 10% significance level by the tests of $L_{\rho=0}$ and Martin and Betensky (2005).

To determine which form of weight is the best choice for L_w, we compare the values of $\log L(\hat{\alpha})$ under the three model alternatives, namely the Clayton ($L_{\rho=0}$), Frank ($L_{\rho=1}$) and Gumbel ($L_{inv\log}$) copula. Recall that $\log L(\hat{\alpha})$ denotes the log-likelihood when $\hat{\alpha}$ is the maximized value of α over the corresponding parameter space. For both groups, the Clayton model seems to be the best fitted one and hence $L_{\rho=0}$ is the best test. In summary, our analysis indicates that there is some positive association between the lapse time and the incubation time. The pattern of association can be considered as a constant over the observable region, $\{(x, y): 0 < x \le y < \infty\}$.

For the Channing House example, we ran two analyses for testing quasi-independence between a resident's lifetime (Y) and his/her entry age to the community (X).The results of five tests are given in Table 4. The first analysis uses the data provided in Hyde (1980) which contains 462 (97 men and 365 women) subjects. Among them, 286 people withdrew from the community yielding the censoring proportion 0.62. Based on the first half of Table 4, the Z-value of each test indicates slightly positive association between X and Y (τ_b = 0.088). The three tests, namely $L_{\rho=1}$, Tsai and Martin and Betensky tests, reach the 10% significance level. In fact, the likelihood analysis favors the Frank model under which the score test is $L_{\rho=1}$. Recall that in presence of censoring, Tsai and Martin and Betensky's tests use the weight $R(x,y)/n$ while $L_{\rho=1}$ adopt the weight $\hat{v}(x,y) = R(x,y)/\{n\hat{S}_C(y)\}$. Hence they are no longer equivalent.

The second analysis uses the data in Hyde (1977), where only the 97 men were studied with 51 subjects being censored. This subset also reveals positive association between X and Y (τ_b=0.199). Based on the second half of Table 4, the three log-rank type tests fail to reject quasi-independence. The values of maximized log-likelihood still favor the Frank alternative in which the score test is $L_{\rho=1}$ with the p-value 0.168. In contrast, both Tsai test and Martin and Betensky test suggest rejecting quasi-independence at 5% level (p-values: 0.043 and 0.040 respectively).

Now we examine the results in more detail. The methods of variance estimation seem to have not much effect. In fact, if we tested the second dataset using L_W with $W(x,y) = R(x,y)/n$ and the jackknife variance estimator, the corresponding Z-value becomes -2.033 (p-value: 0.042) which is very close to the results based on the two competing tests. In an analysis not shown here, we find that both weight functions reveal similar patterns under the two samples. Therefore the test result seems to be mostly affected by the chosen weight function. Note that the function $R(x,y)/n$ assigns higher weight to early failure time y than $\hat{v}(x,y)$. We suspect that the association at earlier time period is higher for the subset of men than it is for the whole sample of 462 subjects.

6.2 Simulation Studies

Finite-sample performances of the proposed test and their competitors are evaluated via simulations. Random pairs of (X,Y) were generated from three semi-survival AC models, namely the Clayton, Frank and Gumbel families discussed in (11). The level of association for an AC model can be described in terms of (pre-truncated) Kendall's tau defined as $\tau = E[\text{sgn}(X_i - X_j)(Y_i - Y_j)]$ which is independent of the marginal distributions. We also provide the value of conditional Kendall's tau τ_a or τ_b which depend on both the copula structure and the marginal distributions. The censoring variable C was generated independently form (X,Y). The marginals of (X,Y,C) follow exponential distributions with the hazard rate $(\lambda_X,\lambda_Y,\lambda_C) = (1,0.5,0)$, $(1,1,0)$, $(0.5,1,0)$ and $(1.5,1,0.5)$ respectively. Under the four settings, the corresponding truncation proportions become $c = P(X \le Y) = 0.67$, 0.5, 0.33 and $c^* = P(X \le Z) = 0.5$ respectively.

We consider three proposed tests, namely $L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$, which use the weights calculated in the three examples of Section 3.2 and the jackknife method for variance estimation. For the Clayton, Frank and Gumbel alternatives, the score tests correspond to $L_{\rho=0}$, $L_{\rho=1}$ (or the two competing tests) and $L_{inv\log}$ respectively. The tests proposed by Tsai (1990) and Martin and Betensky (2005) are also evaluated. Recall that, in absence of censoring, these two tests constructed based on K are equivalent to $L_{\rho=1}$ but use their own variance estimators. Performances of the five tests at $n = 100$ and 200 are studied.

Table 5 and 6 list the results based on 500 replications when (X,Y) follow the Clayton model with the special property of $\tau = \tau_a = \tau_b$. Under the null hypothesis of quasi-independence, the rejection probability for most tests are close to the nominal level with $\alpha = 0.05$. In all the cases, the score test $L_{\rho=0}$ is uniformly more powerful than the other tests. When censoring is present and $n = 100$, the gain in power by using the score test is more obvious. The test $L_{\rho=1}$ and two related tests proposed by Tsai and Martin and Betensky (2005) have similar and sometimes unsatisfactory performances.

The results for the Frank model under different levels of association are summarized in Table 7 ($n = 100$) and Table 8 ($n = 200$). As mentioned earlier, the score test based on $L_{\rho=1}$ and the tests proposed by Tsai (1990) and Martin and Betensky (2005) are closely related. Under the Frank model, they have shown higher power than $L_{\rho=0}$ and $L_{inv\log}$ as expected but a clear-cut dominance among the three is not found. However compared with the Clayton's case, the magnitude of power improvement reduces a little bit. This may be due to extra variation of $\pi(x, y)$ in the suggested weight for the Frank model.

Table 9 contains the results under the Gumbel model with $\tau = -0.2$ and $\tau = -0.4$ since the semi-survival Gumbel model only permits negative association. In contrast to the Clayton and Frank models, the discrepancy for the power curves of different tests becomes less clear at $n = 100$. Nevertheless when $n = 200$, the proposed score test based on $L_{inv\log}$ still performs slightly better than the competing tests. We suspect that the gain by using the suggested form of weight $w\{\eta(x, y)\} = 1/\log\{c\pi(x, y)\}$ may be somewhat offset by estimating two nuisance parameters c and $\pi(x, y)$.

In general, the simulation results confirm that the suggested weight in (10) can improve the power when the alternative is correctly specified. On the other hand, a wrong choice of weight may result in loss of power.

7. CONCLUDING REMARKS

A related area of research is testing independence for bivariate failure times. Rank-based procedures were proposed by Cuzick (1982, 1985) and Dabrowska (1986). Oakes (1982) suggested a concordance test based on an estimate of Kendall's tau which keeps the information of ranks and has a nice expression as a U-statistic. Shih and Louis (1996) utilized the covariance process of martingale residuals to constructs test statistics. Hsu and Prentice (1996) generalized the idea of Mantel-Haenszel statistics to test independence for right censored data. Similar idea has been extended to bivariate current status data by Ding and Wang (2004) based on another formulation of 2×2 tables.

This article considers left truncated data in presence of right censoring. A modified version of Kendall's tau was proposed by Tsai (1990) and then used as the basis for testing quasi-independence by Tsai (1990) and Martin and Betensky (2005). Alternatively we apply the idea of log-rank type statistics based on 2×2 tables designed for describing truncation data. By permitting a flexible weight function, the proposed statistics forms a very general class of tests. A nice equivalence property between the log-rank type statistics and the Kendall's tau statistics has been established. This relationship allows us to compare different types of tests in a systematic way and it turns out that the weight function plays a crucial role. The 2×2 table construction sheds some light on the underlying likelihood structure. In particular, motivated by the papers of Clayton (1978) and Oakes (1986), we derive a score test when the dependence structure under the alternative hypothesis can be modeled via the odds ratio function $\vartheta(x,y)$. Compared with the conditional Kendall's tau measure, $\vartheta(x,y)$ is a better association measure since it is independent of the marginal distributions and can be accurately estimated in presence of censoring. By comparing the values of the conditional likelihood functions under different model choices, a heuristic model selection procedure is proposed. The score test has the log-rank expression with the weight function chosen to fit the alternative hypothesis and, as expected, it has nice power in the simulations. Nevertheless, it is still hard to derive any optimality property of the score test since the likelihood function in (6) is constructed based on a working assumption of independence among the tables. It will be an interesting topic to further establish some theoretical foundation for this kind of conditional likelihood functions. To establish large-sample properties for the weighted log-rank statistics, we have applied the functional delta method which can handle more general situations than the results based on rank statistics or U-statistics. In particular, by expressing the proposed test as a statistically differentiable functional, asymptotic normality of the weighted log-rank statistics and consistency of the jackknife variance estimator is justified.

For analyzing more complicated truncation and censoring structures, Martin and Betensky (2005) considered several extended versions of Kendall's tau and utilized properties of U-statistics in variance estimation and large-sample analysis. It would be interesting to apply the idea of log-rank tests to these data settings. This extension is not trivial since the formulation of appropriate "risk sets" in the construction of 2×2 tables over time is not straightforward. We will leave this as a future research topic.

To select a more powerful test, some investigation under the alternative hypothesis is needed. Here we propose a heuristic model selection procedure by comparing the values of the conditional likelihood functions under different choices of $\vartheta(x, y)$. When the association pattern is completely un-specified, we may consider combining several weighed log-rank statistics (Tarone 1981; Fleming and Harrington; Kosorok and Lin 1999). This approach is considered to be robust (Kosorok and Lin 1999) in that one may avoid using the worst weight choice for the data analysis. To implement this method, the joint distributions of several weighted log-rank statistics are needed.

APPENDIX A: ASYMPTOTIC ANALYSIS

Let $D\{[0,\infty)^2\}$ be the collection of all right-continuous functions with left-side limit defined on $[0,\infty)^2$, whose norm is defined by $\|f(x,y)\|_\infty = \sup_{x,y} |f(x,y)|$ for $f \in D\{[0,\infty)^2\}$. We assume that the function $\pi(x,y) = F_X(x)S_Y(y)/c$ is absolutely continuous. The empirical process on the plane is defined as:

$$\hat{\pi}(x,y) = \frac{1}{n}\sum_{j=1}^{n} I(X_j \le x, Y_j > y).$$

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 mean 0 Gaussian process $V(x,y)$ on $D\{[0,\infty)^2\}$ with the covariance structure given by

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

for any (x_1, y_1), $(x_2, y_2) \in [0,\infty)^2$.

A.1 Proof of Theorem 1

After some algebraic manipulations based on (6), we obtain

$$L_w = \iint_{x \le y} w\{\hat{\pi}(x, y-)\} \left\{ N_{11}(dx, dy) - \frac{N_{1\bullet}(dx, y) N_{\bullet 1}(x, dy)}{R(x, y)} \right\}$$

$$= -\frac{1}{2n} \sum_{i,j} I\{A_{ij}\} \frac{w\{\hat{\pi}(\breve{X}_{ij}, \tilde{Y}_{ij}-)\}}{\hat{\pi}(\breve{X}_{ij}, \tilde{Y}_{ij}-)} \operatorname{sgn}\{(X_i - X_j)(Y_i - Y_j)\}.$$

Here, the symmetry of each term between indices (i, j) and (j,i). The property

$$d\hat{\pi}(x, y) = \begin{cases} -1/n & X_i = x, Y_i = y \quad \text{for some } i \in \{1,...,n\} \\ 0 & \text{otherwise} \end{cases},$$

allows us to write the above expression as

$$L_w = -\frac{n}{2} \iint\iint_{x \vee x^* \le y \wedge y^*} \frac{w\{\hat{\pi}(x \vee x^*, y \wedge y^* -)\}}{\hat{\pi}(x \vee x^*, y \wedge y^* -)} \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\hat{\pi}(x, y) d\hat{\pi}(x^*, y^*)$$
$$\equiv -n\Phi(\hat{\pi}),$$

where the definition of the functional $\Phi(\cdot) : D\{[0,\infty)^2\} \to \mathbf{R}$ is

$$\Phi(\pi) = \iint\iint_{x \vee x^* \le y \wedge y^*} \frac{w\{\pi(x \vee x^*, y \wedge y^* -)\}}{2\pi(x \vee x^*, y \wedge y^* -)} \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\pi(x, y) d\pi(x^*, y^*).$$

By setting the argument π as $\pi(x, y) = \Pr(X \le x, Y > y \mid X \le Y)$ and viewing the above integral as an expectation, we have $\Phi(\pi) = 0$:

$$\Phi(\pi) = E\left[I\{A_{12}\} \frac{w\{\pi(\breve{X}_{12}, \tilde{Y}_{12}-)\}}{2\pi(\breve{X}_{12}, \tilde{Y}_{12}-)} \operatorname{sgn}\{(X_1 - X_2)(Y_1 - Y_2)\} \right]$$
$$= E\left[I\{A_{12}\} \frac{w\{\pi(\breve{X}_{12}, \tilde{Y}_{12}-)\}}{2\pi(\breve{X}_{12}, \tilde{Y}_{12}-)} E\{\operatorname{sgn}\{(X_1 - X_2)(Y_1 - Y_2)\} \mid \breve{X}_{12}, \tilde{Y}_{12}\} \right] = 0 \cdot$$

By direct calculations, we can show the Hadamard differentiability of $\Phi(\cdot)$. The differential map of $\Phi(\cdot)$ at $\pi \in D\{[0,\infty)^2\}$ with direction $h \in D\{[0,\infty)^2\}$ is

$$
\begin{aligned}
&\Phi'_\pi(h)\\
&=\iiiint_{x\vee x^*\le y\wedge y}\frac{w'\{\pi(x\vee x^*,y\wedge y^*-)\}}{2\pi(x\vee x^*,y\wedge y^*-)}h(x\vee x^*,y\wedge y^*-)\,\mathrm{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*)\\
&-\iiiint_{x\vee x^*\le y\wedge y}\frac{w\{\pi(x\vee x^*,y\wedge y^*-)\}}{2\pi(x\vee x^*,y\wedge y^*-)^2}h(x\vee x^*,y\wedge y^*-)\,\mathrm{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*)\\
&+\iiiint_{x\vee x^*\le y\wedge y}\frac{w\{\pi(x\vee x^*,y\wedge y^*-)\}}{\pi(x\vee x^*,y\wedge y^*-)}\mathrm{sgn}\{(x-x^*)(y-y^*)\}dh(x,y)d\pi(x^*,y^*).
\end{aligned}
$$

Applying the functional delta method (Van Der Vaart 1998, p. 297), we obtain the following asymptotic linear expression

$$
\begin{aligned}
n^{-1/2}L_w &= -n^{1/2}\Phi(\hat{\pi})\\
&= -n^{1/2}\{\Phi(\hat{\pi})-\Phi(\pi)\}\\
&= -n^{-1/2}\sum_j \Phi'_\pi(\delta_{(X_j,Y_j)}-\pi)+o_P(1),
\end{aligned}
$$

where $\delta_{(X_j,Y_j)}(x,y)=I(X_j\le x, Y_j>y)$. It is easy to see that the sequences,

$$
U(X_j,Y_j)\equiv\Phi'_\pi(\delta_{(X_j,Y_j)}-\pi)\quad\text{for}\quad j=1,\ldots,n,
$$

are *iid* random variables with mean-zero. From the central limit theorem, $n^{-1/2}L_w$ converges weakly to a mean 0 normal distribution with the variance $\sigma^2=E[U(X_j,Y_j)^2]$.

A.2 Analytic Variance Estimator for the G^ρ Class

Recall that the G^ρ class is a sub-family of L_w. For this class, one can obtain the explicit formula of $U(X_j,Y_j)$ as

$$
\begin{aligned}
&U_\rho(X_j,Y_j)\\
&=(\rho-1)/2\iiiint_{x\vee x^*\le y\wedge y^*}\pi(x\vee x^*,y\wedge y^*-)^{\rho-2}\\
&\quad\times\{I(X_j\le x\vee x^*,Y_j\ge y\wedge y^*)-\pi(x\vee x^*,y\wedge y^*-)\}\mathrm{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*)\\
&\quad-\iiiint_{x\vee x^*\le y\wedge y^*}\pi(x\vee x^*,y\wedge y^*-)^{\rho-1}\mathrm{sgn}\{(x-x^*)(y-y^*)\}\\
&\quad\times\{I(X_i=x,Y_i=y)+d\pi(x,y)\}d\pi(x^*,y^*).
\end{aligned}
$$

Accordingly it is relatively easier to obtain an analytic estimator of σ^2 based on asymptotic linear expressions. Specifically, the derivative map is given by

$$\Phi'_{\pi}(h)$$
$$=(\rho-1)/2\iiiint_{x\vee x^*\le y\wedge y^*}\pi(x\vee x^*,y\wedge y^*-)^{\rho-2}h(x\vee x^*,y\wedge y^*-)\operatorname{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*)$$
$$+\iiiint_{x\vee x^*\le y\wedge y^*}\pi(x\vee x^*,y\wedge y^*-)^{\rho-1}\operatorname{sgn}\{(x-x^*)(y-y^*)\}dh(x,y)d\pi(x^*,y^*).$$

Hence the asymptotic variance of L_ρ can be estimated by $\sum_j \Phi'_{\hat\pi}(\delta_{(X_j,Y_j)}-\hat\pi)^2$, where

$$\Phi'_{\hat\pi}(\delta_{(X_j,Y_j)}-\hat\pi)$$
$$=(\rho-1)/2\iiiint_{x\vee x^*\le y\wedge y^*}\hat\pi(x\vee x^*,y\wedge y^*-)^{\rho-2}$$
$$\times\{I(X_j\le x\vee x^*,Y_j\ge y\wedge y^*)-\hat\pi(x\vee x^*,y\wedge y^*-)\}\operatorname{sgn}\{(x-x^*)(y-y^*)\}d\hat\pi(x,y)d\hat\pi(x^*,y^*)$$
$$-\iiiint_{x\vee x^*\le y\wedge y^*}\hat\pi(x\vee x^*,y\wedge y^*-)^{\rho-1}\operatorname{sgn}\{(x-x^*)(y-y^*)\}$$
$$\times\{I(X_j=x,Y_j=y)+d\hat\pi(x,y)\}d\hat\pi(x^*,y^*)$$
$$=\frac{1}{n}\sum_k I\{A_{jk}\}\hat\pi(\breve X_{jk},\breve Y_{jk}-)^{\rho-1}\operatorname{sgn}\{(X_j-X_k)(Y_j-Y_k)\}+\frac{(\rho+1)L_\rho}{n}$$
$$+\frac{\rho-1}{n^2}\sum_{k<l}I\{A_{kl}\}\hat\pi(\breve X_{kl},\breve Y_{kl}-)^{\rho-2}\operatorname{sgn}\{(X_k-X_l)(Y_k-Y_l)\}I(X_j\le\breve X_{kl},Y_j\ge\breve Y_{kl}).$$

Based on the above expression, we can estimate the asymptotic variance $AVar(L_\rho)=n\sigma^2$

by the following empirical estimator:

$$n\hat\sigma^2=\sum_j\Bigg[\frac{1}{n}\sum_k I\{A_{jk}\}\hat\pi(\breve X_{jk},\breve Y_{jk}-)^{\rho-1}\operatorname{sgn}\{(X_j-X_k)(Y_j-Y_k)\}+\frac{(\rho+1)L_\rho}{n}$$
$$+\frac{\rho-1}{n^2}\sum_{k<l}I\{A_{kl}\}\hat\pi(\breve X_{kl},\breve Y_{kl}-)^{\rho-2}\operatorname{sgn}\{(X_k-X_l)(Y_k-Y_l)\}I(X_j\le\breve X_{kl},Y_j\ge\breve Y_{kl})\Bigg]^2.$$

A.3 Proof of Theorem 2

The statistic L_w^* involves estimation of $c=P(X\le Y)$. We suggest to adopt the estimator proposed by He and Yang (1998) which can be expressed as

$$\hat c=\int_0^\infty \hat S_Y(u)d\hat F_X(u),$$

where the marginal functions are estimated by the Lynden-Bell (1971) estimators,

$$\hat F_X(t)=\prod_{t<u}\left\{1-\frac{d\hat\pi(u,0)}{\hat\pi(u,u-)}\right\},\quad \hat S_Y(t)=\prod_{u\le t}\left\{1+\frac{d\hat\pi(\infty,u)}{\hat\pi(u,u-)}\right\}.$$

By writing $\hat c=g(\hat\pi)$, we will show that the map $g:\hat\pi\mapsto\hat c$ is the composition of two Hadamard differentiable maps:

$$\hat{\pi}(x,y) \quad \mapsto \quad (\hat{F}_X(x), \hat{S}_Y(y)) \quad \mapsto \quad \int_0^{\infty} \hat{S}_Y(u) d\hat{F}_X(u). \qquad \text{(A.1)}$$

It is well-known that the product limit estimator for right-censored data, in the first map of (A.1), is a Hadamard differentiable function of the empirical process. For truncation data, we apply the arguments of example 20.15 in Van Der Vaart (1998) to show the Hadamard differentiability of maps from $D\{[0,\infty)^2\}$ to $D\{[0,\infty)\}$:

$$\hat{\pi}(x,y) \mapsto \hat{F}_X(t), \quad \hat{\pi}(x,y) \mapsto \hat{S}_Y(t).$$

To prove the former statement, we decompose the map into three differentiable maps

$$\hat{\pi}(x,y) \quad \mapsto \quad (\hat{\pi}(x,0), 1/\hat{\pi}(x,x-)) \quad \mapsto \quad \hat{\Lambda}_X(t) = \int_0^t \frac{d\hat{\pi}(u,0)}{\hat{\pi}(u,u-)}$$
$$\mapsto \hat{F}_X(t) = \prod\nolimits_{t<u} \{1 - d\hat{\Lambda}_X(u)\},$$

where the Hadamard differentiability of the second map follows from Lemma 20.10 of Van Der Vaart (1998) and the last map follows from the Hadamard differentiability of the product integral. The Hadamard differentiability of the map $\hat{\pi}(x,y) \mapsto \hat{S}_Y(t)$ can be established by the same arguments. The Hadamard differentiability of the second map in (A.1) can be found in Lemma 20.10 of Van Der Vaart (1998).Using chain rules (Van Der Vaart 1998, Theorem 20.9), the map g is shown to be Hadamard differentiable. Let $g'_{\pi}(h) \in \mathbf{R}$ be the differential map of g at $\pi \in D\{[0,\infty)^2\}$ with direction $h \in D\{[0,\infty)^2\}$ such that

$$n^{1/2}(\hat{c} - c) = n^{1/2}(g(\hat{\pi}) - g(\pi)) = n^{-1/2} \sum\nolimits_j g'_{\pi}(h_{X_j,Y_j} - \pi) + o_P(1).$$

The statistics L_w^* can be expressed as

$$L_w^* = -\frac{n}{2} \iint\iint_{x \vee x^* \le y \wedge y^*} \frac{w\{g(\hat{\pi})\hat{\pi}(x \vee x^*, y \wedge y^* -)\}}{\hat{\pi}(x \vee x^*, y \wedge y^* -)} \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\hat{\pi}(x,y) d\hat{\pi}(x^*, y^*)$$
$$\equiv -n\Psi(\hat{\pi}).$$

Applying similar arguments in Section A.1, we can show $\Psi(\pi) = 0$. Now we show the Hadamard differentiability of the map $\Psi(\cdot): D\{[0,\infty)^2\} \to \mathbf{R}$. From the Hadamard differentiability of $g(\cdot)$,

$$g(\pi+th)=g(\pi)+g'_{\pi}(\pi)t+o(|t|),\quad t\to 0,$$

uniformly in h in compact subsets of $D\{[0,\infty)^2\}$. This leads to the following Taylor expansion

$$\begin{aligned}&w\{g(\pi+th)(\pi(x\vee x^*,y\wedge y^*-)+th(x\vee x^*,y\wedge y^*-)\}\\&=w\{c\pi(x\vee x^*,y\wedge y^*-)\}+t\{ch(x\vee x^*,y\wedge y^*-)+g'_{\pi}(h)\pi(x\vee x^*,y\wedge y^*-)\}+o(|t|).\end{aligned}$$

A little calculus shows the derivative map of $\Psi(\cdot)$ at $\pi\in D\{[0,\infty)^2\}$ with direction $h\in D\{[0,\infty)^2\}$:

$$\begin{aligned}&\Psi'_{\pi}(h)\\&=\iiiint_{x\vee x^*\le y\wedge y}\frac{cw'\{\pi(x\vee x^*,y\wedge y^*-)\}}{2\pi(x\vee x^*,y\wedge y^*-)}h(x\vee x^*,y\wedge y^*-)\operatorname{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*)\\&+\iiiint_{x\vee x^*\le y\wedge y}\frac{g'_{\pi}(h)w'\{\pi(x\vee x^*,y\wedge y^*-)\}}{2}\operatorname{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*)\\&-\iiiint_{x\vee x^*\le y\wedge y}\frac{w\{c\pi(x\vee x^*,y\wedge y^*-)\}}{2\pi(x\vee x^*,y\wedge y^*-)^2}h(x\vee x^*,y\wedge y^*-)\operatorname{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*)\\&+\iiiint_{x\vee x^*\le y\wedge y}\frac{w\{\pi(x\vee x^*,y\wedge y^*-)\}}{\pi(x\vee x^*,y\wedge y^*-)}\operatorname{sgn}\{(x-x^*)(y-y^*)\}dh(x,y)d\pi(x^*,y^*).\end{aligned}$$

By applying the functional delta method, we obtain the following asymptotic linear expression:

$$n^{-1/2}L_w^*=-n^{1/2}\Psi(\hat{\pi})=-n^{1/2}\{\Psi(\hat{\pi})-\Psi(\pi)\}$$

$$=-n^{-1/2}\sum_j\Psi'_{\pi}(h_{(X_j,Y_j)}-\pi)+o_P(1),$$

where the sequences,

$$U^*(X_j,Y_j)\equiv\Psi'_{\pi}(h_{(X_j,Y_j)}-\pi)\quad(j=1,\ldots,n),$$

are mean-zero *iid* random variables. From the central limit theorem, $n^{-1/2}L_w^*$ converges weakly to a mean-zero normal distribution with the variance σ_*^2.

A.4 Asymptotic Analysis in Presence of Censoring

Based on the product integral form of the Lynden-Bell's estimator $\hat{S}_C(y)$, we obtain the expression

$$\hat{v}(x, y-) = \frac{\hat{H}(x, y-, y-)}{\prod_{u \le y} \left\{1 + \hat{H}(u,u,du) \big/ \hat{H}(u,u-,u-)\right\}} \equiv \varphi(\hat{H}; x, y). \qquad \text{(A.2)}$$

A little algebra shows that the event B_{ij} can be written as $I\{B_{ij}\} = I(\breve{X}_{ij} \le \widetilde{Y}_{ij} < \widetilde{C}_{ij})$, where $\breve{X}_{ij} = X_i \vee X_j$ and $\widetilde{Y}_{ij} = Y_i \wedge Y_j$. From the results in Appendix B, equations (A.1) and (A.2), we obtain the following functional expression:

$$L_w = \iint_{x \le y} w\{\hat{v}(x, y-)\}\left\{N_{11}(dx, dy) - \frac{N_{1\bullet}(dx, y)N_{\bullet 1}(x, dy)}{R(x, y)}\right\} = \sum_{i<j} I\{B_{ij}\}\frac{2w\{\varphi(\hat{H}; \breve{X}_{ij}, \widetilde{Z}_{ij})\}}{R(\breve{X}_{ij}, \widetilde{Z}_{ij})}\left(\Delta_{ij} - \frac{1}{2}\right)$$

$$= -\frac{n}{2} \times \frac{1}{n^2}\sum_{i,j} I\{\breve{X}_{ij} \le \widetilde{Y}_{ij} < \widetilde{C}_{ij}\}\frac{w\{\varphi(\hat{H}; \breve{X}_{ij}, \widetilde{Y}_{ij} \wedge \widetilde{C}_{ij})\}}{\hat{H}(\breve{X}_{ij}, \widetilde{Y}_{ij} \wedge \widetilde{C}_{ij}-, \widetilde{Y}_{ij} \wedge \widetilde{C}_{ij}-)}\operatorname{sgn}\{(X_i - X_j)(Y_i - Y_j)\}$$
$$= -\frac{n}{2}\iiiint\!\!\iint_{x \vee x^* \le y \wedge y^* < c \wedge c^*} \frac{w\{\varphi(\hat{H}; x \vee x^*, y \wedge y^* \wedge c \wedge c^*)\}}{\hat{H}(x \vee x^*, y \wedge y^* \wedge c \wedge c^*-, y \wedge y^* \wedge c \wedge c^*-)}$$
$$\times \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\hat{H}(x, y, c) d\hat{H}(x^*, y^*, c^*).$$

Based on similar arguments as in Section A.3, we can write $\hat{c}^* = g^*(\hat{H})$. By applying similar algebraic operations, we can obtain the functional expression for L_w^*.

APPENDIX B: PROOF OF (6) and (16)

In this section, we prove equations (6) and (16). Note that equation (6) is the uncensored case with $C_i = \infty$. For mathematical convenience, we define the discordant indicator $\Delta_{ij} = I\{(X_i - X_j)(Z_i - Z_j) < 0\}$ .To simplify the notations, let $W(\breve{X}_{ij}, \widetilde{Z}_{ij}) = \widetilde{W}_{ij}$ and $R(\breve{X}_{ij}, \widetilde{Z}_{ij}) = \widetilde{R}_{ij}$ .We can write

$$L_w = \sum_i \sum_{\substack{j: X_j \le X_i \\ X_i \le Z_j \le Z_i}} \delta_j W(X_i, Z_j)\left\{N_{11}(dX_i, dZ_j) - \frac{1}{R(X_i, Z_j)}\right\}$$

$$= \sum_i \delta_i W(X_i, Z_i)\frac{R(X_i, Z_i) - 1}{R(X_i, Z_i)} - \sum_i \sum_{\substack{j: X_j < X_i \\ X_i \le Z_j < Z_i}} \delta_j W(X_i, Z_j)\frac{1}{R(X_i, Z_j)}$$

$$\equiv I_1 - I_2.$$

Using the fact that $\sum_j I(X_j < X_i, Z_j > Z_i) = R(X_i, Z_i) - 1$, it follows that

$$I_1 = \sum_i \sum_{j:X_j<X_i, Z_j>Z_i} \delta_i \frac{W(X_i,Z_i)}{R(X_i,Z_i)} = \sum_i \sum_{j:X_j<X_i, Z_i<Z_j} \delta_i \frac{\widetilde{W}_{ij}}{\widetilde{R}_{ij}}.$$

The indicator Δ_{ij} equals zero for a pair (i,j) with $X_j < X_i, Z_i < Z_j$. Therefore

$$I_1 = \sum_i \sum_{j:X_j<X_i, Z_i<Z_j} \delta_i \Delta_{ij} \frac{\widetilde{W}_{ij}}{\widetilde{R}_{ij}} = \sum_i \sum_{j:X_j<X_i, X_i<Z_j} \delta_i \Delta_{ij} \frac{\widetilde{W}_{ij}}{\widetilde{R}_{ij}}.$$

By applying similar algebraic manipulations, it follows that

$$I_2 = \sum_i \sum_{\substack{j:X_j<X_i \\ X_i \le Z_j < Z_i}} \delta_j \frac{\widetilde{W}_{ij}}{\widetilde{R}_{ij}} = \sum_i \sum_{j:X_j<X_i, X_i \le Z_j} \delta_j (1-\Delta_{ij}) \frac{\widetilde{W}_{ij}}{\widetilde{R}_{ij}}.$$

Combining I_1 and I_2, we obtain

$$L_W = \sum_i \sum_{j:X_j<X_i, X_i \le Z_j} \widetilde{W}_{ij} \frac{\delta_i \Delta_{ij} - \delta_j(1-\Delta_{ij})}{\widetilde{R}_{ij}} = \sum_i \sum_{j:X_j<X_i} I\{\breve{X}_{ij} \le \widetilde{Z}_{ij}\} \widetilde{W}_{ij} \frac{\delta_i \Delta_{ij} - \delta_j(1-\Delta_{ij})}{\widetilde{R}_{ij}}.$$

For a pair (i,j) with $X_j < X_i$, the following equation holds:

$$\begin{aligned} &\delta_i \Delta_{ij} - \delta_j (1-\Delta_{ij}) \\ &= I\{(\delta_i = \delta_j = 1) \cup (Z_j - Z_i > 0 \,\&\, \delta_i = 1 \,\&\, \delta_j = 0) \cup (Z_i - Z_j > 0 \,\&\, \delta_i = 0 \,\&\, \delta_j = 1)\}(2\Delta_{ij} - 1). \end{aligned}$$

Thus, we obtain the equation (16) as follows:

$$\begin{aligned} L_W &= \sum_i \sum_{j:X_j<X_i} I\{B_{ij}\} \widetilde{W}_{ij} \frac{2\Delta_{ij} - 1}{\widetilde{R}_{ij}} = \sum_{i<j} I\{B_{ij}\} \widetilde{W}_{ij} \frac{2\Delta_{ij} - 1}{\widetilde{R}_{ij}} \\ &= -\sum_{i<j} I\{B_{ij}\} \frac{\widetilde{W}_{ij}}{\widetilde{R}_{ij}} \operatorname{sgn}\{(X_i - X_j)(Z_i - Z_j)\}. \end{aligned}$$

The second equation follows from the permutation symmetry of each term with respect to arguments (i,j).

REFERENCES

Breslow, N. E. (1972), "Discussion of Professor Cox's paper," *Journal of the Royal Statistical Society, Series B*, 34, 216-217.

Chaieb, L. Rivest, L. -P. and Abdous, B. (2006), "Estimating Survival Under a Dependent Truncation," *Biometrika*, 93, 655-669.

Chen, C. -H., Tsai, W. -Y. and Chao, W. -H. (1996), "The Product-moment Correlation Coefficient and Linear Regression for Truncated Data.", *Journal of the American Statistical Association*, 91, 1181-1186.

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.

Cuzick, J. (1982), "Rank Tests for Association With Right Censored Data," *Biometrika,* 69, 351-364.

Cuzick, J. (1985), "Asymptotic Properties of Censored Linear Rank Tests," *The Annals of Statistics*, 13, 133-141.

Ding A. A. and Wang. W. (2004), "Testing Independence for Bivariate Current Status Data," *Journal of the American Statistical Association*, 99**,** 145-155.

Dabroska, D. M. (1986), "Rank Tests for Independence for Bivariate Censored Data," *The Annals of Statistics*, 14, 250-264.

Fleming, T. R. and Harrington, D. P. (1991), "*Counting Process and Survival Analysis,*" New York: John Wiley and Sons.

Genest, C. (1987), "Frank's Family of Bivariate Distributions," *Biometrika*, 74, 549-555.

Harrington, D. P., and Fleming, T. R. (1982), "A Class of Rank Test Procedures for Censored Survival Data," *Biometrika*, 69, 133-143.

He, S. and Yang, G. L. (1998), "Estimation of the Truncation Probability in the Random Truncation Model," *Annals of Statistics*. 26, 1011-1027.

Hsu, L. and Prentice, R. L. (1996), "A Generalisation of the Mantel-Haenszel Test to Bivariate Failure Time Data," *Biometrika*, 83, 905-911.

Hyde, J. (1977), "Testing Survival Under Right Censoring and Left Truncation," *Biometrika,* 64, 225-230.

Hyde, J. (1980), "Survival Analysis with Incomplete Observations", *In Biostatistics*

Casebook, R. G. Miller, B. Efron, B. W. Brown, and L. E. Moses, eds. New York; John Wiley and Son, 31-46.

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

Kosorok, M. R. and Lin, C. (1999), "The Versatility of Functional-Indexed Weighted Log-rank Statistics," *Journal of the American Statistical Association*, 94, 320-332.

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. and DE Gruttola, V. (1988), "Non-parametric Analysis of Truncated Survival Data, With Application to AIDS," *Biometrika*, 75, 515-523.

Martin, E. C. and Betensky, R. A. (2005), "Testing Quasi-independence of Failure and Truncation via Conditional Kendall's Tau," *Journal of the American Statistical Association*, 100, 484-492.

Oakes, D. (1982), "A concordance Test for Independence in the Presence of Censoring," *Biometrics*, 38, 451-455.

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," *Journal of the American Statistical Association,* 84, 487-493.

Peto, R. (1972), "Discussion of Professor Cox's paper," *Journal of the Royal Statistical Society, Series B*, 34, 205-207.

Shao, J. (1993), "Differentiability of Statistical Functionals and Consistency of the Jackknife," *The Annals of Statistics,* 21, 61-75.

Shih, J. H. and Louis, T. A. (1996), "Tests of Independence for Bivariate Survival Data," *Biometrics,* 52, 1440-1449.

Tarone, R. E. (1981), "On the Distribution of the Maximum of the Log-rank Statistics and the Modified Wilcoxon Statistics," *Biometrics,* 37, 79-85.

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. and Tsai, W. Y. (1986), "Asymptotic Properties of the Product-limit Estimate and Right Censored Data," *Annals of Statistic*, 13, 1597-1605.

Woodroofe, M. (1985), "Estimating a Distribution Function With Truncated Data," *The Annals of Statistics,* 13, 163-177.

Table 3. Tests of quasi-independence for the AIDS data

	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\log}$	Tsai	M & B
Adult					
Z-value	-5.012	-2.918	-3.795	2.567	2.833
P-value	5.398×10^{-7}	3.519×10^{-3}	1.475×10^{-4}	1.027×10^{-2}	4.610×10^{-3}
$\log L(\hat{\alpha})$	-1077.878	-1080.054	-1082.860	Undefined	Undefined
Children					
Z-value	-1.838	-1.379	-1.373	0.966	1.672
P-value	0.066	0.168	0.170	0.334	0.095
$\log L(\hat{\alpha})$	-95.225	-95.434	-95.859	Undefined	Undefined

Table 4. Tests of quasi-independence for the Channing House data.

	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\log}$	Tsai	M & B
462 subjects					
Z-value	-0.515	-1.669	-1.169	1.776	1.837
P-value	0.607	0.095	0.243	0.076	0.066
$\log L(\hat{\alpha})$	-809.207	-807.954	-809.316	Undefined	Undefined
97 men					
Z-value	-1.286	-1.379	-1.116	2.021	2.053
P-value	0.198	0.168	0.264	0.043	0.040
$\log L(\hat{\alpha})$	-139.297	-139.267	-140.268	Undefined	Undefined

Table 5. Empirical rejection probabilities of three proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under Clayton's model with sample size 100.

$c = \Pr(X \le Y)$ or $c^* = \Pr(X \le Z)$	τ (τ_a / τ_b)	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\log}$	**Tsai**	**M & B**
$c = 0.67$	-0.2 (-0.200)	0.908	0.832	0.860	0.856	0.800
	-0.1 (-0.100)	0.410	0.320	0.334	0.344	0.312
	0.0 (0.000)	0.052	0.044	0.042	0.046	0.046
	0.1 (0.100)	0.518	0.374	0.442	0.358	0.378
	0.2 (0.200)	0.998	0.914	0.962	0.900	0.910
$c = 0.50$	-0.2 (-0.200)	0.900	0.802	0.852	0.832	0.786
	-0.1 (-0.100)	0.404	0.290	0.344	0.334	0.280
	0.0 (0.000)	0.062	0.054	0.044	0.052	0.064
	0.1 (0.100)	0.456	0.354	0.376	0.322	0.372
	0.2 (0.200)	0.998	0.912	0.984	0.888	0.914
$c = 0.33$	-0.2 (-0.200)	0.900	0.794	0.838	0.846	0.786
	-0.1 (-0.100)	0.396	0.290	0.332	0.340	0.272
	0.0 (0.000)	0.046	0.036	0.032	0.036	0.038
	0.1 (0.100)	0.518	0.382	0.438	0.352	0.410
	0.2 (0.200)	0.990	0.896	0.978	0.900	0.920
$c^* = 0.50$	-0.2 (-0.200)	0.836	0.698	0.734	0.704	0.648
	-0.1 (-0.100)	0.342	0.236	0.258	0.240	0.230
	0.0 (0.000)	0.044	0.050	0.036	0.054	0.048
	0.1 (0.100)	0.360	0.292	0.296	0.236	0.276
	0.2 (0.200)	0.940	0.786	0.862	0.734	0.766

Table 6. Empirical rejection probabilities of three proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under Clayton's model with sample size 200.

$c = \Pr(X \le Y)$ or $c^* = \Pr(X \le Z)$	τ (τ_a / τ_b)	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\log}$	**Tsai**	**M & B**
$c = 0.67$	-0.2 (-0.200)	0.990	0.970	0.988	0.974	0.970
	-0.1 (-0.100)	0.706	0.534	0.626	0.590	0.522
	0.0 (0.000)	0.048	0.056	0.044	0.060	0.050
	0.1 (0.100)	0.872	0.646	0.798	0.622	0.658
	0.2 (0.200)	1.000	1.000	0.998	1.000	1.000
$c = 0.50$	-0.2 (-0.200)	1.000	0.984	0.992	0.984	0.978
	-0.1 (-0.100)	0.684	0.520	0.614	0.566	0.514
	0.0 (0.000)	0.044	0.056	0.044	0.060	0.054
	0.1 (0.100)	0.874	0.670	0.824	0.642	0.676
	0.2 (0.200)	1.000	0.998	1.000	0.998	0.998
$c = 0.33$	-0.2 (-0.200)	0.990	0.974	0.986	0.982	0.974
	-0.1 (-0.100)	0.684	0.510	0.628	0.570	0.504
	0.0 (0.000)	0.052	0.070	0.046	0.058	0.070
	0.1 (0.100)	0.886	0.678	0.822	0.642	0.696
	0.2 (0.200)	1.000	1.000	1.000	0.998	1.000
$c^* = 0.50$	-0.2 (-0.200)	0.980	0.940	0.962	0.932	0.918
	-0.1 (-0.100)	0.536	0.408	0.448	0.404	0.380
	0.0 (0.000)	0.054	0.036	0.044	0.042	0.044
	0.1 (0.100)	0.694	0.520	0.612	0.502	0.514
	0.2 (0.200)	0.998	0.980	0.996	0.974	0.980

Table 7. Empirical rejection probabilities of three proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under Frank's model with sample size 100.

$c = \Pr(X \le Y)$ or $c^* = \Pr(X \le Z)$	τ (τ_a / τ_b)	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\log}$	**Tsai**	**M & B**
$c = 0.67$	-0.4 (-0.242)	0.864	0.956	0.946	0.956	0.952
	-0.2 (-0.103)	0.292	0.348	0.322	0.366	0.330
	0.0 (0.000)	0.052	0.044	0.042	0.046	0.046
	0.2 (0.081)	0.214	0.256	0.236	0.234	0.270
	0.4 (0.163)	0.532	0.738	0.620	0.722	0.742
$c = 0.50$	-0.4 (-0.189)	0.664	0.852	0.806	0.848	0.830
	-0.2 (-0.075)	0.166	0.206	0.198	0.216	0.182
	0.0 (0.000)	0.062	0.054	0.044	0.052	0.064
	0.2 (0.047)	0.114	0.126	0.102	0.116	0.152
	0.4 (0.082)	0.216	0.234	0.208	0.244	0.286
$c = 0.33$	-0.4 (-0.135)	0.406	0.544	0.498	0.552	0.510
	-0.2 (-0.050)	0.100	0.098	0.082	0.112	0.098
	0.0 (0.000)	0.046	0.036	0.032	0.036	0.038
	0.2 (0.026)	0.060	0.084	0.046	0.076	0.090
	0.4 (0.034)	0.064	0.066	0.056	0.062	0.106
$c^* = 0.50$	-0.4 (-0.238)	0.684	0.786	0.760	0.792	0.762
	-0.2 (-0.100)	0.202	0.244	0.202	0.250	0.238
	0.0 (0.000)	0.044	0.050	0.036	0.054	0.048
	0.2 (0.073)	0.144	0.184	0.152	0.178	0.210
	0.4 (0.130)	0.294	0.398	0.302	0.404	0.432

Table 8. Empirical rejection probabilities of three proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X,Y) under Frank's model with sample size 200.

$c = \Pr(X \le Y)$ or $c^* = \Pr(X \le Z)$	τ (τ_a / τ_b)	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\log}$	**Tsai**	**M & B**
$c = 0.67$	-0.4 (-0.242)	0.990	1.000	0.996	1.000	1.000
	-0.2 (-0.103)	0.456	0.596	0.576	0.606	0.600
	0.0 (0.000)	0.048	0.056	0.044	0.060	0.050
	0.2 (0.081)	0.340	0.486	0.406	0.452	0.490
	0.4 (0.163)	0.820	0.970	0.906	0.968	0.968
$c = 0.50$	-0.4 (-0.189)	0.912	0.972	0.966	0.970	0.972
	-0.2 (-0.075)	0.260	0.410	0.370	0.428	0.418
	0.0 (0.000)	0.044	0.056	0.044	0.060	0.054
	0.2 (0.047)	0.188	0.242	0.208	0.226	0.236
	0.4 (0.082)	0.334	0.434	0.358	0.438	0.468
$c = 0.33$	-0.4 (-0.135)	0.654	0.812	0.792	0.824	0.810
	-0.2 (-0.050)	0.138	0.158	0.150	0.166	0.160
	0.0 (0.000)	0.052	0.070	0.046	0.058	0.070
	0.2 (0.026)	0.072	0.086	0.064	0.092	0.088
	0.4 (0.034)	0.104	0.148	0.102	0.148	0.180
$c^* = 0.50$	-0.4 (-0.238)	0.914	0.978	0.978	0.976	0.976
	-0.2 (-0.100)	0.342	0.402	0.390	0.398	0.406
	0.0 (0.000)	0.054	0.036	0.044	0.042	0.044
	0.2 (0.073)	0.246	0.306	0.258	0.288	0.300
	0.4 (0.130)	0.584	0.714	0.638	0.698	0.706

Table 9. Empirical rejection probabilities of three proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X,Y) under Gumbel's model with sample sizes 100 and 200.

$c = \Pr(X \le Y)$ or $c^* = \Pr(X \le Z)$	τ (τ_a / τ_b)	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\log}$	**Tsai**	**M & B**
$n = 100$						
$c = 0.67$	-0.4 (-0.200)	0.804	0.844	0.836	0.860	0.820
	-0.2 (-0.081)	0.226	0.224	0.220	0.234	0.206
$c = 0.50$	-0.4 (-0.169)	0.712	0.722	0.730	0.756	0.700
	-0.2 (-0.063)	0.196	0.174	0.168	0.186	0.168
$c = 0.33$	-0.4 (-0.138)	0.552	0.478	0.500	0.524	0.466
	-0.2 (-0.054)	0.142	0.124	0.116	0.126	0.116
$c^* = 0.50$	-0.4 (-0.197)	0.658	0.670	0.670	0.668	0.626
	-0.2 (-0.080)	0.134	0.146	0.126	0.144	0.134
$n = 200$						
$c = 0.67$	-0.4 (-0.200)	0.978	0.992	0.994	0.990	0.990
	-0.2 (-0.081)	0.360	0.376	0.392	0.386	0.364
$c = 0.50$	-0.4 (-0.169)	0.934	0.944	0.950	0.950	0.936
	-0.2 (-0.063)	0.308	0.302	0.312	0.310	0.306
$c = 0.33$	-0.4 (-0.138)	0.828	0.792	0.830	0.824	0.786
	-0.2 (-0.054)	0.226	0.200	0.208	0.212	0.204
$c^* = 0.50$	-0.4 (-0.197)	0.868	0.920	0.916	0.912	0.902
	-0.2 (-0.080)	0.278	0.236	0.248	0.238	0.208