Technical Report
Asymptotic Analysis and Variance Estimation for Testing Quasi-independence under Truncation

*Takeshi Emura and **Weijing Wang

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

**Division of Biostatistics, 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

A class of weighted log-rank statistics for testing quasi-independence for truncated data is proposed. This manuscript contains asymptotic analysis for the proposed methods and simulation studies to evaluate their finite-sample performances. By temporarily ignoring censoring, we derive large-sample properties of the proposed tests in Section 1 and discuss variance estimation in Section 2. Section 3 considers the extended situation that accounts for right censoring. In Section 4, we compare two methods for estimating the variance of the proposed statistics. One estimator has an analytical formula and the other is the jackknife estimator. Our simulations indicate that the jackknife variance estimator, which is easier to implement, has reliable performance. Section 5 contains some concluding remarks. Detailed proofs are given in the Appendix.

Key Words: Functional delta method; Independence test; Jackknife method; Log-rank statistics; Mantel-Heanszel test; Two-by-two tables; Weak convergence.

1. LARGE SAMPLE PROPERITES IN ABSENCE OF CENSORING

We consider the truncation setting in which a pair of lifetime variables (X,Y) can be included in the sample only if $X \le Y$. Most methods for analyzing truncated data are based on the assumption of quasi-independence (Tsai 1990) which can be stated as

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

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, and c_0 is a constant satisfying $c_0 = -\iint_{x\le y} dF_X(x)dS_Y(y)$. In this section, we consider observed data of the form $\{(X_j,Y_j)\ (j=1,\ldots,n)\}$ subject to $X_j \le Y_j$ by temporarily ignoring external censoring.

1.1 The Proposed Test Statistics

For testing H_0, the proposed weighted log-rank type statistics can be written as

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

where

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

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

and $W(x,y)$ is a pre-specified weight function. It can be shown that, under H_0,

$$\mathrm{E}(N_{11}(dx,dy) \mid N_{1\bullet}, N_{\bullet 1}, R) = \frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)}. \tag{3}$$

A useful special case of L_W has the form,

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

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

When the observations have no ties such that the values of $X_1,\ldots X_n,Y_1,\ldots,Y_n$ are all distinct, we have $N_{\bullet 1}(x,dy) = N_{1\bullet}(x,dy) = 1$ and L_W can be expressed as

$$L_W = -\sum_{i<j} I\{A_{ij}\} \frac{W(\breve{X}_{ij}, \widetilde{Y}_{ij})}{R(\breve{X}_{ij}, \widetilde{Y}_{ij})} \mathrm{sgn}\{(X_i - X_j)(Y_i - Y_j)\}, \quad (5)$$

where $A_{ij} = \{\breve{X}_{ij} \le \widetilde{Y}_{ij}\}$ and $\breve{X}_{ij} = X_i \vee X_j$ and $\widetilde{Y}_{ij} = Y_i \wedge Y_j$, and $\mathrm{sgn}(x)$ is defined to be -1, 0, or 1 if $x < 0$, $x = 0$, or $x > 0$, respectively. If the form of $W(x,y)/R(x,y)$ is deterministic, the expression of L_W in (5) is a U-statistic. For example, if $W(x,y)/R(x,y) = 1$, the statistics leads the conditional Kendall's tau statistics

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

Properties of U-statistics have been utilized by Martin and Betensky (2005) to derive nice analytic results of K or its extended versions. However if the function $W(x,y)/R(x,y)$ is subject to random variation, this technique is not applicable.

1.2 Large Sample Analysis

Here we consider a general class of L_W which can include flexible weight functions. It is expected that an appropriate weight function is helpful for increasing the power of the corresponding test. For large-sample analysis, we adopt the functional delta method, a powerful tool that can handle weight functions containing plugged-in estimators. In particular, we consider the following two statistics:

$$L_w = \iint_{x \le y} w\{\hat{\pi}(x, y-)\} \left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y) N_{\bullet 1}(x,dy)}{R(x,y)} \right\}, \quad (6)$$

$$L_w^* = \iint_{x \le y} w\{\hat{c}\hat{\pi}(x, y-)\} \left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y) N_{\bullet 1}(x,dy)}{R(x,y)} \right\}, \quad (7)$$

where $w(v)$ is a known function which is continuously differentiable and

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

Notice that L_w and L_w^* differ in whether the estimator $\hat{c}$ is involved. Also note that these two statistics can be considered as approximations of the unbiased statistics

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

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

respectively. The functional delta method can handle the extra estimation of the hypothesized weight function in a systematic way.

To simplify the analysis we assume that the distributions of (X,Y) under the null hypothesis in (1) are absolutely continuous. The formula in (6) and (7) 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 satisfying $\hat{c}=g(\hat{\pi})$ defined in Appendix A.3. The proofs of the above theoretical results are provided in the Appendices A.1 and A.3. Briefly speaking, these functionals 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 expressions:

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

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

where $U(X_j,Y_j)$ and $U^*(X_j,Y_j)$ are defined in Appendix A.1 and A.3.

Theorem 1: Under H_0, $n^{-1/2}L_w$ converges in distribution to a mean-zero normal random variable with variance $\sigma^2=E[U(X_j,Y_j)^2]$ where $U(X_j,Y_j)$ is defined in Appendix A.1.

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

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

An analytic variance estimator for the L_ρ class is presented in Section 4.

To study the statistics L_w^*, we need to examine the property of $\hat{c}$ which is closely related to the marginal estimators of $F_X(t)$ and $S_Y(t)$. Asymptotic normality of L_w^* can be established under 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 an identifiability condition for $(F_X(\cdot),S_Y(\cdot))$, 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 the denominator terms $R(X_{(1)},X_{(1)})$ and $R(X_j,X_j)$ for $\hat{c}$ is away from zero. Thus, under H_0 and Assumption (I), $\hat{c}$ is a reliable estimator of c_0.

Theorem 2: Under H_0 *and Assumption (I),* $n^{-1/2}L_w^*$ *converges in distribution to a mean-zero normal distribution with variance* $\sigma_*^2=E[U^*(X_j,Y_j)^2]$ *where* $U^*(X_j,Y_j)$ *is defined in Appendix A.3.*

2. JACKKNIFE ESTIMATOR OF VARIANCE

The asymptotic variance of L_W can be estimated by the jackknife estimator:

$$\frac{n-1}{n}\sum_i (L_W^{(-i)} - L_W^{(\cdot)})^2 ,$$

where $L_W^{(-i)}$ has the form of L_W without the i th observation and $L_W^{(\cdot)} = (1/n)\sum_i L_W^{(-i)}$. Asymptotic properties of the jackknife variance estimator are closely related to smoothness of the corresponding functional expression. Specifically the test statistics L_w in the class (6) can be expressed as $-n\Phi(\hat{\pi})$, where $\hat{\pi} = \hat{\pi}(x, y)$ is an empirical process and $\Phi(\cdot)$ is a functional defined on a space of function $\pi = \pi(x, y) = \Pr(X \le x, Y > y \mid X \le Y)$ (see Appendix A.1). In proving the asymptotic normality, we have shown that $\Phi(\cdot)$ is Hadamard differentiable with respect to the argument π. However, to show consistency of the jackknife variance estimator, we need a more strict smoothness condition on $\Phi(\cdot)$ called continuous Gateaux differentiability (Shao 1993).

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

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

The proofs of the above theorems are given in Appendix A.4.

3. LARGE SAMPLE PROPERITES UNDER CENSORING

3.1 The Proposed Test Statistics

When Y_i is further subject to censoring by C_i, observed data become $\{(X_i, Z_i, \delta_i)$ $(i = 1,...,n)\}$ subject to $X_i \le Z_i$, where $Z_i = Y_i \wedge C_i$ and $\delta_i = I(Y_i \le C_i)$. Assume that C_i is independent of (X_i, Y_i). At an uncensored failure point (x, y) with $x \le y$, we use the same notations for the cell and marginal counts with the following modified definitions:

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

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

Accordingly the proposed log-rank statistics becomes

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

which can be expressed as

$$L_W = -\sum_{i<j} I\{B_{ij}\}\frac{W(\breve{X}_{ij},\tilde{Z}_{ij})}{R(\breve{X}_{ij},\tilde{Z}_{ij})}\operatorname{sgn}\{(X_i - X_j)(Z_i - Z_j)\}, \tag{8}$$

where the event

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

implies that the pair (i,j) is comparable and orderable (Martin and Betensky 2005). Under the quasi-independence assumption, it can be shown that

$$E[\operatorname{sgn}\{(X_i - X_j)(Z_i - Z_j)\} \mid B_{ij}] = 0.$$

For a constant $\rho \in [0,\infty)$, the L_ρ statistics becomes

$$L_\rho = \iint_{x\le y} \hat{v}(x,y-)^\rho \left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)} \right\}, \tag{9}$$

where $\hat{v}(x,y) = \sum_j I(X_j \le x, Z_j > y)/\{n\hat{S}_C(y)\}$ is an estimator of $\pi(x,y)$ and $\hat{S}_C(y)$ is the Lynden-Bell's (1971) estimator for $\Pr(C > y) = S_C(y)$ based on data $\{(X_i, Z_i, 1-\delta_i)$ $(i = 1,\dots,n)\}$.

3.2 Asymptotic Analysis under Censoring

Now we discuss the asymptotic normality of the two statistics in presence of censoring. For a continuously differentiable function $w(\cdot)$, we consider classes of statistics:

$$L_w = \iint_{x\le y} w\{\hat{v}(x,y-)\}\left\{ N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)} \right\}, \tag{10}$$

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

where $X_{(1)} = \min_j X_j$ and

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

Asymptotic normality of the modified statistics is only briefly sketched in Appendix A.5 since the related formula under censoring are too complicated and provide no new insight. Define $\hat{H}(x,y,c) = \sum_j I(X_j \le x, Y_j > y, C_j > c)/n$ which is the empirical estimator of $H(x,y,c) = \Pr(X \le x, Y > y, C > c \mid X \le Z)$. It follows that

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

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

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.5. Asymptotic normality of L_w and L_w^* can be established by applying the functional delta method based on the facts that both of them are Hadamard differentiable functions of $\hat{H}$ and the process $n^{1/2}(\hat{H} - H)$ converges weakly to a Gaussian process. Similar to the uncensored case, consistency of the jackknife variance estimator is built based on the continuous Gateaux differentiability of the corresponding functional expression. The proof is similar as that for Theorem 3 and 4 and hence is omitted.

4. EMPIRICAL VARIANCE AND JACKKNIFE VARIANCE ESTIMATOR

In absence of censoring, asymptotic variance of the L_ρ test defined in (4) has a tractable form. Based on the method of moment and applying the plug-in principle, we obtain the following analytic variance estimator of L_ρ:

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

The formula is derived in Appendix A.2. In presence of censoring, however, modification of the above formula involves complicated mathematical derivations. To examine whether the extra effort is worthy or not, we compare $\hat{V}(L_\rho)$ with the jackknife variance estimator via simulations. Note that the latter is computationally convenient even under the censoring situation and is theoretically justified in Theorem 3 and 4.

The first set of simulations evaluates the test L_ρ standardized by two variance formula, namely the analytic and jackknife estimators, in absence of censoring. The variables (X,Y) are generated independently from exponential distributions with hazards (λ_X,λ_Y) respectively. Five cases with $(\lambda_X,\lambda_Y)=(1.5,0.5)$, $(1,0.5)$, $(1,1)$, $(0.5,1)$ and $(0.5,1.5)$ ($c=\Pr(X\le Y)$=0.75, 0.67, 0.50, 0.33 and 0.25 respectively) are examined. At 5% level of significance, the null hypothesis is rejected if $|L_\rho/\hat{\sigma}(L_\rho)|$ is greater than 1.96, where $\hat{\sigma}(L_\rho)^2$ is either the analytic estimator or jackknife estimator.

Table 1 reports the results of the $L_{\rho=0}$ test standardized by two variance estimators under $n=50,100$ and 200 based on 500 replications. The mean squared error (MSE) for estimating the sample variance decreases as the sample size gets large for both the analytic and jackknife variance estimator. However, the analytic variance estimator has smaller MSE than the jackknife estimator. On the other hands, the analytic method tends to systematically underestimate the variance so that the type I error rate of the corresponding test is inflated. It is clear that the test standardized by the jackknife variance has type I error rate closer to the nominal 5% level in all the cases. Table 2 summarizes the results of the $L_{\rho=1}$ test. Both variance estimators tend to overestimate the true variance but the jackknife estimator has slightly smaller MSE and still produces more accurate type I error probability than the analytic estimator. The results indicate that the test standardized by the jackknife variance estimator has more reliable performance in most situations.

The second set of simulations evaluates the test L_W standardized by the jackknife variance estimator in presence of censoring in which the analytic formula is not available. Three independent variables (X,Y,C) are generated from exponential distributions with hazard rates $(\lambda_X,\lambda_Y,\lambda_C)$=(1.5, 1.0, 0.5) respectively yielding $c^* = P(X \le Z)$=0.5. Table 3 summarizes the results for three tests $L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\log}$ under $n = 50$, 100 and 200 based on 500 replications. The jackknife variance method, on the average, is quite accurate. The MSE for estimating the variance gets small as the number of sample increases. Furthermore the type I error rates for the three statistics are all close to the nominal level.

5. CONCLUSION

To establish asymptotic normality, we apply the functional delta method which can handle more general forms of statistics than the approaches based on U-statistics or rank statistics. The expression of the proposed statistics as a statistically differentiable functional allows us to derive an analytic variance estimator and justify the use of the jackknife method in variance estimation. Simulation analysis indicates that the jackknife method is a better alternative because its convenience and reliable results. Nevertheless one may try to modify the analytic variance formula by including more omitted terms in the Taylor expansion to see if the bias can be reduced. However the derivations will be very tedious and may not be worthy from a practical point of view.

APPENDIX: ASYMPTOTIC ANALYSIS

Let $D\{[0,\infty)^2\}$ and $D\{[0,\infty)\}$ be the collection of all right-continuous functions with left-side limit defined on $[0,\infty)^2$ and $[0,\infty)$ respectively, under which the norms are defined by $\|f(x,y)\|_\infty = \sup_{x,y} |f(x,y)|$ for $f \in D\{[0,\infty)^2\}$ and $\|f(x)\|_\infty = \sup_x |f(x)|$ for $f \in D\{[0,\infty)\}$. We assume that the function $\pi(x,y) = F_X(x)S_Y(y)/c_0$ is absolutely continuous. Hereafter, expectation symbols represent the conditional expectation given $X \le Y$. The notation $o_P(1)$ is short of random variables that converge to zero in probability, under the conditional probability induced by $\pi(x,y) = F_X(x)S_Y(y)/c_0$. The empirical process on the plane is defined as:

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

It can be shown that $n^{1/2}(\hat{\pi}(x,y) - \pi(x,y))$ converges weakly to a mean 0 Gaussian process $V(x,y)$ on $D\{[0,\infty)^2\}$ with the covariance given by

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

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

A.1 Proof of Theorem 1

By some algebraic manipulations, the statistic L_w in (6) can be expressed in terms of the weighted sum of signs such that

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

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

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

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

where the last equation use the facts that each term is symmetric for indices (i,j) and (j,i) and that $\text{sgn}\{(X_j - X_j)(Y_j - Y_j)\} = 0$. Using the property that

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

the above expression can be written as

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

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

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

The argument $\pi \in D\{[0,\infty)^2\}$ in the preceding equation does not need to be $\pi(x,y) = F_X(x)S_Y(y)/c_0$, but if so, the above integral can be interpreted as an expectation under quasi-independence. It can be shown that $\Phi(\pi) = 0$ since

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

and

$$E\{\text{sgn}\{(X_1 - X_2)(Y_1 - Y_2)\} \mid \breve{X}_{12} = x, \tilde{Y}_{12} = y\} = 0 \ \text{ for } \ x \le y.$$

Here, the last equation follows from Chaieb et al. (2006). The basic idea of the functional delta method is to find the asymptotic behavior of $\Phi(\hat{\pi})$ through a differential analysis of $\Phi(\pi)$ in a neighborhood of $\pi(x,y) = F_X(x)S_Y(y)/c_0$. A first order expansion would have the form:

$$\Phi(\hat{\pi}) = \Phi(\hat{\pi}) - \Phi(\pi) \approx \Phi'_\pi(\hat{\pi} - \pi).$$

where $\Phi'_\pi(\cdot)$ will be rigorously defined later. The analysis turn the weak convergence of $\Phi(\hat{\pi})$ into the weak convergence of $\hat{\pi} - \pi$, which will next be further investigated in a rigorous fashion.

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

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

where $w'(u) = \partial w(u)/\partial u$. By applying the functional delta method (Van Der Vaart 1998, p. 297), we obtain the following asymptotic linear expression:

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

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 $E[U(X_j,Y_j)] = 0$ at $\pi(x,y) = F_X(x)S_Y(y)/c_0$ since

$$\begin{aligned}
E[\Phi'_\pi(\delta_{(X_j,Y_j)} - \pi)] &= \Phi'_\pi(E[\delta_{(X_j,Y_j)} - \pi]) \\
&= \Phi'_\pi(0) = 0
\end{aligned}.$$

Based on the central limit theorem, $n^{-1/2}L_w$ converges in distribution to a mean 0 normal random variable with variance $\sigma^2 = E[U(X_j,Y_j)^2]$.

A.2 Analytic Variance Estimator for the L_ρ Test

The L_ρ statistics forms a nice subset of L_w such that σ^2 can be expressed by analytic formula. The functional expression has the form $L_\rho = -n\Phi(\hat{\pi})$, where

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

for $\pi\in D\{[0,\infty)^2\}$ and $\Phi(\pi)=0$ for $\pi(x,y) = F_X(x)S_Y(y)/c_0$. Specifically, the corresponding derivative map is given by

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

The variance $\sigma^2 = E[U(X_j,Y_j)^2]$ can be estimated by $\sum_j \Phi'_{\hat{\pi}}(\delta_{(X_j,Y_j)}-\hat{\pi})^2/n$, where

$$\begin{aligned}&\Phi'_{\hat{\pi}}(\delta_{(X_j,Y_j)}-\hat{\pi})\\&=(\rho-1)/2\iiiint_{x\vee x^*\le y\wedge y^*}\hat{\pi}(x\vee x^*, y\wedge y^*-)^{\rho-2}\\&\quad\times\{I(X_j\le x\vee x^*, Y_j\ge y\wedge y^*)-\hat{\pi}(x\vee x^*, y\wedge y^*-)\}\operatorname{sgn}\{(x-x^*)(y-y^*)\}d\hat{\pi}(x,y)d\hat{\pi}(x^*,y^*)\\&-\iiiint_{x\vee x^*\le y\wedge y^*}\hat{\pi}(x\vee x^*, y\wedge y^*-)^{\rho-1}\operatorname{sgn}\{(x-x^*)(y-y^*)\}\\&\quad\times\{I(X_j=x, Y_j=y)+d\hat{\pi}(x,y)\}d\hat{\pi}(x^*,y^*)\end{aligned}$$

Turning the integral to the double summation, we obtain

$$\begin{aligned}\Phi'_{\hat{\pi}}(\delta_{(X_j,Y_j)}-\hat{\pi}) &= \frac{\rho-1}{n^2}\sum_{k<l} I\{A_{kl}\}\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl}-)^{\rho-2}\operatorname{sgn}\{(X_k-X_l)(Y_k-Y_l)\}I(X_j\le\breve{X}_{kl}, Y_j\ge\tilde{Y}_{kl})\\&-\frac{\rho-1}{n^2}\sum_{k<l} I\{A_{kl}\}\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl}-)^{\rho-1}\operatorname{sgn}\{(X_k-X_l)(Y_k-Y_l)\}\\&+\frac{1}{n}\sum_{k} I\{A_{jk}\}\hat{\pi}(\breve{X}_{jk},\tilde{Y}_{jk}-)^{\rho-1}\operatorname{sgn}\{(X_j-X_k)(Y_j-Y_k)\}\\&-\frac{2}{n^2}\sum_{k<l} I\{A_{kl}\}\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl}-)^{\rho-1}\operatorname{sgn}\{(X_k-X_l)(Y_k-Y_l)\}.\end{aligned}$$

The second and fourth terms combine to form $(\rho+1)L_\rho/n$. Thus we get

$$\begin{aligned}\Phi'_{\hat{\pi}}(\delta_{(X_j,Y_j)}-\hat{\pi}) &= \frac{1}{n}\sum_{k} I\{A_{jk}\}\hat{\pi}(\breve{X}_{jk},\tilde{Y}_{jk}-)^{\rho-1}\operatorname{sgn}\{(X_j-X_k)(Y_j-Y_k)\}+\frac{(\rho+1)L_\rho}{n}\\&+\frac{\rho-1}{n^2}\sum_{k<l} I\{A_{kl}\}\hat{\pi}(\breve{X}_{kl},\tilde{Y}_{kl}-)^{\rho-2}\operatorname{sgn}\{(X_k-X_l)(Y_k-Y_l)\}I(X_j\le\breve{X}_{kl}, Y_j\ge\tilde{Y}_{kl}).\end{aligned}$$

Based on the above expressions, we can estimate $AVar(L_\rho) = n\sigma^2$ by the following empirical estimator:

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

A.3 Proof of Theorem 2

The statistic L_w^* involves $\hat{c}$ which is the estimator of the normalizing constant c_0. From the result of He and Yang (1998), $\hat{c}$ can be written as

$$\hat{c} = \int_0^\infty \hat{S}_Y(u) d\hat{F}_X(u),$$

where $\hat{F}_X(t) = \prod_{t<u} \left\{ 1 - \frac{d\hat{\pi}(u,0)}{\hat{\pi}(u,u-)} \right\}$ and $\hat{S}_Y(t) = \prod_{u \le t} \left\{ 1 + \frac{d\hat{\pi}(\infty,u)}{\hat{\pi}(u,u-)} \right\}$ are the product limit estimators of $\Pr(X \le t)$ and $\Pr(Y > t)$ respectively (Lynden-Bell 1971). Define $\hat{c} = g(\hat{\pi})$ and 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 for right-censored data that the product limit estimator is Hadamard differentiable function of the corresponding empirical processes. For truncation data, we apply the arguments of Example 20.15 in Van Der Vaart (1998) to show the Hadamard differentiability of the 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 each map into three differentiable maps. For example, one can write

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

where the Hadamard differentiability of the second map follows from Lemma 20.10 in 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 applying similar arguments. The Hadamard differentiability of the second map in (A.1) can be found in Lemma 20.10 in Van Der Vaart (1998). Using the chain rule (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

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

The statistics L_w^* can be expressed as

$$\begin{aligned} L_w^* &= -\frac{n}{2} \iiiint_{x \vee x^* \leq y \wedge y^*} \frac{w\{g(\hat{\pi})\hat{\pi}(x \vee x^*, y \wedge y^*-)\}}{\hat{\pi}(x \vee x^*, y \wedge y^*-)} \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\hat{\pi}(x,y) d\hat{\pi}(x^*,y^*) \\ &\equiv -n\Psi(\hat{\pi}). \end{aligned}$$

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 $g(\cdot)$, it follows that

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

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

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

It can be shown that the derivative map of $\Psi(\cdot)$ at $\pi \in D\{[0,\infty)^2\}$ with direction

$h \in D\{[0,\infty)^2\}$ can be written as

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

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

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

where the sequences,

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

are mean 0 i.i.d. random variables. 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.4 Consistency of the Jackknife Estimator

We have shown that $n^{-1/2}L_w$ is asymptotically normal with a finite variance. Now we show consistency of the jackknife variance estimator of L_w. According to Theorem 3.1 of Shao (1993), we need to show the continuous Gateaux differentiability of $\Phi(\pi)$ at $\pi \in D\{[0,\infty)^2\}$. Note that the Hadamard differentiability is stronger than the Gateaux differentiability and hence the Gateaux derivative map is given by $\Phi'_\pi(h)$, available from Section A.1. We only need to show the continuous requirement of the derivative map. For sequence $\pi_k \in D\{[0,\infty)^2\}$ satisfying $\|\pi_k - \pi\|_\infty \to 0$ and $t_k \to 0$ for $k = 1,2,\ldots$, we need to show

$$A_k \equiv \Phi\{\pi_k + t_k(\delta_{u,v} - \pi_k)\} - \Phi(\pi_k) - t_k \Phi'_\pi(\delta_{u,v} - \pi_k) = o(|t_k|),$$

where $\delta_{u,v}(x,y) = I(x \le u, y > v)$ and $o(|t_k|)$ stands for $o(|t_k|)/|t_k| \to 0$ uniformly in (u,v). The following arguments for proving the continuous Gateaux differentiability are similar to those in Example 2.6 of Shao (1993). The continuous differentiability of $w(\cdot)$ and the assumption $\|\pi_k - \pi\|_\infty \to 0$ ensure the following expansion

$$\begin{aligned}
&\frac{w[\pi_k(x\vee x^*, y\wedge y^*-) + t_k\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\}]}{2[\pi_k(x\vee x^*, y\wedge y^*-) + t_k\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\}]} - \frac{w\{\pi_k(x\vee x^*, y\wedge y^*-)\}}{2\pi_k(x\vee x^*, y\wedge y^*-)} \\
&= \frac{w'\{\pi(x\vee x^*, y\wedge y^*-)\}}{2\pi(x\vee x^*, y\wedge y^*-)} t_k\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\} \\
&- \frac{w\{\pi(x\vee x^*, y\wedge y^*-)\}}{2\pi(x\vee x^*, y\wedge y^*-)^2} t_k\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\} + O(|t_k|^2),
\end{aligned}$$

uniformly in (u,v). Hence straightforward but tedious calculations show that

$$A_k = |t_k| B_k + |t_k| C_k + |t_k| D_k + O(|t_k|^2),$$

where

$$\begin{aligned}
B_k \equiv &\iint\iint_{x\vee x^* \le y\wedge y^*} \frac{w[\pi_k(x\vee x^*, y\wedge y^*-) + t_k\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\}]}{\pi_k(x\vee x^*, y\wedge y^*-) + t_k\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\}} \\
&\times \operatorname{sgn}\{(x-x^*)(y-y^*)\} d\{\delta_{u,v}(x,y) - \pi_k(x,y)\} d\pi_k(x^*, y^*) \\
&- \iint\iint_{x\vee x^* \le y\wedge y^*} \frac{w\{\pi(x\vee x^*, y\wedge y^*-)\}}{\pi(x\vee x^*, y\wedge y^*-)} \\
&\times \operatorname{sgn}\{(x-x^*)(y-y^*)\} d\{\delta_{u,v}(x,y) - \pi_k(x,y)\} d\pi(x^*, y^*),
\end{aligned}$$

$$\begin{aligned}
C_k = &\int\iiint_{x\vee x^* \le y\wedge y^*} \frac{w'\{\pi_k(x\vee x^*, y\wedge y^*-)\}}{2\pi_k(x\vee x^*, y\wedge y^*-)} \left\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\right\} \\
&\times \operatorname{sgn}\left\{(x-x^*)(y-y^*)\right\} d\pi_k(x,y) d\pi_k(x^*, y^*) \\
&- \int\iiint_{x\vee x^* \le y\wedge y^*} \frac{w'\{\pi(x\vee x^*, y\wedge y^*-)\}}{2\pi(x\vee x^*, y\wedge y^*-)} \left\{\delta_{u,v}(x\vee x^*, y\wedge y^*-) - \pi_k(x\vee x^*, y\wedge y^*-)\right\} \\
&\times \operatorname{sgn}\left\{(x-x^*)(y-y^*)\right\} d\pi(x,y) d\pi(x^*, y^*)
\end{aligned}$$

and

$$D_k \equiv \iint\iint_{x\vee x^*\le y\wedge y^*} \frac{w\{\pi_k(x\vee x^*, y\wedge y^*-)\}}{2\pi_l(x\vee x^*, y\wedge y^*-)^2}\{\delta_{u,v}(x\vee x^*, y\wedge y^*-)-\pi_k(x\vee x^*, y\wedge y^*-)\}$$
$$\times \operatorname{sgn}\{(x-x^*)(y-y^*)\}d\pi_k(x,y)d\pi_k(x^*,y^*)$$
$$-\iint\iint_{x\vee x^*\le y\wedge y^*} \frac{w\{\pi(x\vee x^*, y\wedge y^*-)\}}{2\pi(x\vee x^*, y\wedge y^*-)^2}\{\delta_{u,v}(x\vee x^*, y\wedge y^*-)-\pi_k(x\vee x^*, y\wedge y^*-)\}$$
$$\times \operatorname{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*).$$

Under the assumption that $\|\pi_k-\pi\|_\infty \to 0$, it can be seen that B_k C_k and D_k have order $o(1)$. This proves the theorem 3.

To show the consistency of the jackknife variance estimator for L_w^* stated in theorem 4, we only need to check whether the continuity of the Gateaux differential map of $\Psi(\pi)$ which is available in Section A.3. The continuity requirement for the Gateaux differentiable map $\Psi'_\pi(h)$ in Section A.3 can be verified after tedious algebraic operations similar to the above arguments.

A.5 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

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

By some algebraic work, the event B_{ij} can be written as $I\{B_{ij}\}=I(\breve{X}_{ij}\le\tilde{Y}_{ij}<\tilde{C}_{ij})$. Then, we obtain the following functional expression:

$$L_w=\iint_{x\le y} w\{\hat{v}(x,y-)\}\left\{N_{11}(dx,dy)-\frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)}\right\}=\sum_{i<j} I\{B_{ij}\}\frac{2w\{\varphi(\hat{H};\breve{X}_{ij},\tilde{Z}_{ij})\}}{R(\breve{X}_{ij},\tilde{Z}_{ij})}\left(\Delta_{ij}-\frac{1}{2}\right)$$
$$=-\frac{n}{2}\times\frac{1}{n^2}\sum_{i,j} I\{\breve{X}_{ij}\le\tilde{Y}_{ij}<\tilde{C}_{ij}\}\frac{w\{\varphi(\hat{H};\breve{X}_{ij},\tilde{Y}_{ij}\wedge\tilde{C}_{ij})\}}{\hat{H}(\breve{X}_{ij},\tilde{Y}_{ij}\wedge\tilde{C}_{ij}-,\tilde{Y}_{ij}\wedge\tilde{C}_{ij}-)}\operatorname{sgn}\{(X_i-X_j)(Y_i-Y_j)\}$$

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

$$\times\,\mathrm{sgn}\{(x-x^*)(y-y^*)\}d\hat{H}(x,y,c)d\hat{H}(x^*,y^*,c^*)$$

Here, the last equation follows from the property

$$d\hat{H}(x,y,c) = \begin{cases} 1/n & X_j = x, Y_j = y, C_j = c \quad \text{for some } j \\ 0 & \text{otherwise} \end{cases}.$$

Based on similar arguments with Section A3, we can express the estimator $\hat{c}^*$ as a function of $\hat{H}$ such that $\hat{c}^* = g^*(\hat{H})$. Similar algebraic operations can be applied to obtain the functional expression of L_w^*.

REFERENCES

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

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.

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.

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.

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

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.

Table 1. Comparison of two (jackknife vs. analytic) variance estimators of $n^{-1/2}L_{\rho=0}$ in absence of censoring when 500 replications for pairs X and Y are generated from independent exponential distributions

$c = \Pr(X \le Y)$	n	Variance of $n^{-1/2}L_{\rho=0}$	Average (MSE) of the variance estimators		Type I error (5%)	
			Jackknife	Analytic	Jackknife	Analytic
0.75	50	0.819	0.870 (0.146)	0.653 (0.077)	0.048	0.092
	100	0.872	0.955 (0.121)	0.786 (0.054)	0.062	0.078
	200	0.930	0.997 (0.081)	0.875 (0.039)	0.066	0.078
0.67	50	0.735	0.885 (0.188)	0.652 (0.059)	0.058	0.084
	100	0.859	0.941 (0.110)	0.774 (0.051)	0.062	0.074
	200	0.923	0.957 (0.065)	0.846 (0.039)	0.064	0.072
0.50	50	0.688	0.872 (0.194)	0.630 (0.056)	0.048	0.068
	100	0.820	0.943 (0.146)	0.762 (0.053)	0.056	0.076
	200	0.814	0.981 (0.118)	0.851 (0.043)	0.048	0.056
0.33	50	0.723	0.897 (0.231)	0.630 (0.071)	0.044	0.078
	100	0.733	0.931 (0.172)	0.738 (0.052)	0.050	0.064
	200	0.837	0.941 (0.092)	0.815 (0.039)	0.050	0.064
0.25	50	0.696	0.852 (0.163)	0.604 (0.053)	0.046	0.062
	100	0.802	0.903 (0.123)	0.720 (0.053)	0.064	0.082
	200	0.896	0.960 (0.100)	0.825 (0.050)	0.052	0.066

Table 2. Comparison of two (jackknife vs. analytic) variance estimators of $n^{-1/2}L_{\rho=1}$ in absence of censoring when 500 replications for pairs of X and Y are generated from independent exponential distributions

$c = \Pr(X \le Y)$	n	Variance of $n^{-1/2}L_{\rho=1}$ ($\times 100$)	Average (MSE) of the variance estimators ($\times 100$)		Type I error (5%)	
			Jackknife	Analytic	Jackknife	Analytic
0.75	50	7.493	7.468 (0.0263)	7.604 (0.0264)	0.058	0.036
	100	7.037	7.072 (0.0110)	7.159 (0.0118)	0.048	0.040
	200	6.794	6.800 (0.0055)	6.850 (0.0057)	0.062	0.056
0.67	50	5.438	6.553 (0.0366)	6.614 (0.0377)	0.036	0.026
	100	5.935	6.194 (0.0115)	6.263 (0.0119)	0.056	0.042
	200	5.616	5.889 (0.0062)	5.928 (0.0065)	0.056	0.040
0.50	50	4.097	5.321 (0.0330)	5.352 (0.0340)	0.028	0.014
	100	4.487	4.917 (0.0094)	4.964 (0.0101)	0.050	0.040
	200	4.086	4.665 (0.0074)	4.689 (0.0077)	0.044	0.038
0.33	50	3.901	4.369 (0.0157)	4.426 (0.0168)	0.038	0.024
	100	3.454	3.981 (0.0098)	4.015 (0.0103)	0.038	0.032
	200	3.481	3.756 (0.0042)	3.779 (0.0044)	0.034	0.030
0.25	50	3.414	4.085 (0.0190)	4.124 (0.0203)	0.036	0.024
	100	3.336	3.624 (0.0068)	3.660 (0.0073)	0.040	0.034
	200	3.334	3.376 (0.0027)	3.400 (0.0028)	0.046	0.042

Table 3. Performance of $n^{-1/2}L_W$ *standardized by the jackknife variance estimator in absence of censoring when 500 replications for pairs* X *,* Y *and* C *are generated from independent exponential distributions*

Test statistics	n	Variance of $n^{-1/2}L_W$	Average (MSE) of the jackknife variance estimator	Type I error (5%)
$n^{-1/2}L_{\rho=0}$	50	0.4678	0.5858 (0.06398)	0.050
	100	0.5205	0.6142 (0.04585)	0.046
	200	0.5828	0.6461 (0.02702)	0.050
$n^{-1/2}L_{\rho=1}$	50	0.0745	0.0777 (0.00103)	0.050
	100	0.0720	0.0710 (0.00036)	0.052
	200	0.0613	0.0658 (0.00020)	0.054
$n^{-1/2}L_{inv\log}$	50	0.2075	0.2560 (0.01580)	0.038
	100	0.2382	0.2342 (0.00402)	0.052
	200	0.2010	0.2190 (0.00253)	0.050