Statistica Sinica **10**(2000), 1199-1215

ESTIMATION OF KENDALL'S TAU UNDER CENSORING

Weijing Wang and Martin T. Wells

National Chiao-Tung University and Cornell University

Abstract: We study nonparametric estimation of Kendall's tau, τ, for bivariate censored data. Previous estimators of τ, proposed by Brown, Hollander and Korwar (1974), Weier and Basu (1980) and Oakes (1982), fail to be consistent when marginals are dependent. Here we express τ as an integral functional of the bivariate survival function and construct a natural estimator via the von Mises functional approach. This does not necessarily yield a consistent estimator since tail region information on the survival curve may not be identifiable due to right censoring. To assess the magnitude of the inconsistency we propose some estimable bounds on τ. It is shown that estimates of the bounds shrink to provide consistency if the largest observations on both marginal coordinates are uncensored and satisfy certain regularity conditions. The bounds depend on the sample size, on censoring rates and, in particular, on the estimated probability of the unknown tail region. We also discuss using the bootstrap method for variance estimation and bias correction. Two illustrative data examples are analyzed, as well as some simulation results.

Key words and phrases: Bivariate censored data, bivariate survival function estimation, bootstrap, rank correlation, V-statistic, von Mises functional.

1. Introduction

In many biomedical applications interest focuses on the dependence relationship between two lifetime variables. For example, the analysis of data on lifetimes of twins has been used by geneticists as a tool for assessing genetic effect on mortality (Hougaard, Harvald and Holm (1992)); in $AIDS$ studies, the dependence between the time from HIV infection to $AIDS$ and the time from $AIDS$ to death reveals useful information about the evolution of disease process. Kendall's tau, τ, known as a rank correlation measure, can serve as a simple summary measure of association between two random variables. In contrast to the well-known Pearson correlation coefficient, τ does not require knowledge of the parametric form of the marginal distributions. Its rank invariant property makes it suitable for measuring dependence in non-Gaussian lifetime models. Additionally it has been shown that dependence parameters in the bivariate semi-parametric models proposed by Gumbel (1960), Clayton (1978) and Frank (1979), to name a few, are intimately related to τ. Parameters in these models can then be identified via τ (see Genest (1987), Oakes (1989), Genest and Rivest (1993), Wang and Wells (2000)).

Censoring is a common phenomenon in analysis of lifetime data and it is essential that estimates of τ be available for bivariate censored data. However, few results for this fundamental problem have appeared in the literature. Brown, Hollander and Korwar (1974), Weier and Basu (1980) and Oakes (1982) proposed estimators of τ under censoring, but none of the estimators are consistent when the true value of τ is not equal to zero, that is, when the marginals are dependent. The bias of these estimators increases as the degree of dependence increases. In this article we express τ as an integral of the bivariate survival function. Adopting the ideas of von Mises (1947), a natural way to estimate τ is to plug a suitable bivariate survival estimator into the integral form that defines τ.

In the past decade substantial research effort has been devoted to nonparametric estimation of the bivariate survival function for censored data. A number of nonparametric estimators of the bivariate survival functions, such as those by Campbell (1981), Dabrowska (1988), Prentice and Cai (1992), Lin and Ying (1993), van der Laan (1996) and Wang and Wells (1997), have appeared in the literature. However, just as Kaplan-Meier integrals are biased in the univariate case (Stute (1994)), the von-Mises-type estimator of τ is asymptotically negatively biased. To handle the problem we propose some estimable bounds on the Kendall's tau measure. It will be seen that when the largest observations are uncensored in each marginal coordinate, the bounds shrink to give consistency.

In the next section, notation and previous estimators of τ are introduced. In Section 3 we propose estimators and derive their properties. An application of the bootstrap method for variance estimation and bias correction is also discussed. Illustrative real data examples and simulation results are presented in Section 5. Concluding remarks are given in Section 6.

2. Preliminaries

2.1. Notation

Let (T_1, T_2) be possibly correlated random variables, and let (T_{1i}, T_{2i}) and (T_{1j}, T_{2j}) $(i \neq j)$ be independent realizations from (T_1, T_2). The (i, j)th pair is called *concordant* if $(T_{1i} - T_{1j})(T_{2i} - T_{2j}) > 0$ and *discordant* if $(T_{1i} - T_{1j})(T_{2i} - T_{2j}) < 0$. The population version of Kendall's tau is defined as the difference of concordance and discordance probabilities between the (i, j)th pair. If T_1 and T_2 are continuous, $\tau = \text{pr}\,\{(T_{1i}-T_{1j})(T_{2i}-T_{2j}) > 0\} - \text{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j}) < 0\}$. It is easy to see that $-1 \leq \tau \leq 1$ and if (T_1, T_2) are independent, $\tau = 0$. In the absence of censoring one observes i.i.d replications of (T_1, T_2). Then τ can be easily estimated by taking the difference of sample concordance and discordance proportions. This is equivalent to applying the formula,

$$\hat{\tau} = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} a_{ij} b_{ij}, \tag{2.1}$$

where $a_{ij} = 1$ if $T_{1i} < T_{1j}$, $a_{ij} = -1$ if $T_{2i} > T_{2j}$ and b_{ij} is similarly defined. Notice that the "score", $a_{ij}b_{ij}$, is 1 if the (i,j) pair is concordant and is -1 if discordant. In the complete data setting, it has been shown that $\hat{\tau}$ in (2.1) is a U-statistic, is an unbiased estimate of τ, and $n^{1/2}(\hat{\tau} - \tau)$ is asymptotically normal. See Hoeffding (1948) for further details on the U-statistic representation.

To account for this, Kendall (1962, p.34) proposed two formulas for computing estimates of τ. In both cases, the score is set to zero if a pair has ties, that is $a_{ij} = 0$ if $T_{1i} = T_{1j}$ and $b_{ij} = 0$ if $T_{2i} = T_{2j}$. The first formula, called the "unconditional tau" by Davis and Quade (1968), uses (2.1) with modified scores for ties. The second formula, which excludes tied pairs in computing the total number of combinations, is given by

$$\Gamma = \frac{\sum_{i,j=1}^{n} a_{ij} b_{ij}}{(\sum_{i,j} a_{ij}^2 \sum_{i,j} b_{ij}^2)^{1/2}}. \tag{2.2}$$

It is easy to see that $\Gamma \geq \hat{\tau}$ in all cases and equality holds if there are no ties.

In the case of right censoring, the observable variables are $X = T_1 \wedge C_1$, $Y = T_2 \wedge C_2$, $\delta_j = I\!I(T_j \wedge C_j = T_j)$ $(j = 1,2)$, where (C_1, C_2) are a pair of nuisance censoring variables, "$\wedge$" denotes minimum, and $I\!I(A)$ is the indicator of the event A. Let $F(x,y) = \mathrm{pr}(T_1 > x, T_2 > y)$, $F_j(\cdot)$ $(j = 1,2)$, $H(x,y) = \mathrm{pr}(X > x, Y > y)$ and $H_j(\cdot)$ $(j = 1,2)$ be the joint survival function of (T_1, T_2), the marginal survival function of T_j $(j = 1,2)$, the joint survival function of (X,Y), and the marginal survival functions of X and Y, respectively. Denote the supports of F and H by $\mathcal{S}_F = \{(x,y) : F(x,y) > 0\}$ and $\mathcal{S}_H = \{(x,y) : H(x,y) > 0\}$. The bivariate censored sample is denoted by $\{(X_i, Y_i, \delta_{1i}, \delta_{2i}), i = 1, \ldots, n\}$.

When censoring is present, the relative concordance/discordance relationship is not clear for some pairs. Figure 1 lists possible pair relationships under censoring. In Figure 1 a point is used to denote an observation which is completely observed. A singly censored observation with $\delta_1 = 0$ and $\delta_2 = 1$ is denoted by a right arrow indicating that the possible failure time is located there, similarly if $\delta_1 = 1$ and $\delta_2 = 0$. When $\delta_1 = 0$ and $\delta_2 = 0$, the true failure time is located in the upper right quadrant relative to the observed point. In Figure 1 the only certain pair relationships are (i)-(iii). For example in (i), any point on the right arrow would yield a concordant relationship for the pair.

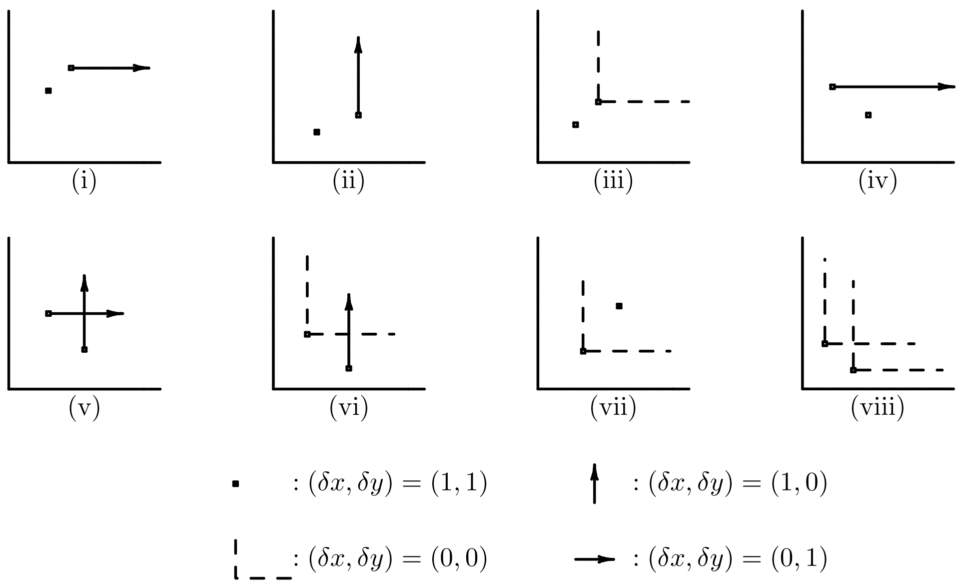

Figure 1. Possible pair relationship between bivariate censored data.

2.2. Previous estimators of τ

Several estimators of τ have been proposed that modify the scores for those pairs whose concordance/discordance relationships are not clear. Brown *et al.* (1974) proposed an estimator of τ which utilized the marginal Kaplan-Meier estimates. Except for ties, they assigned $a_{ij} = 2\ \mathrm{pr}\{T_{1i} > T_{1j} \mid (X_i, X_j, \delta_{1i}, \hat{\mathbb{F}}_1)\} - 1$ and $b_{ij} = 2\ \mathrm{pr}\{T_{2i} > T_{2j} \mid (Y_i, Y_j, \delta_{2i}, \hat{\mathbb{F}}_2)\} - 1$, where $\hat{\mathbb{F}}_i(\cdot)$ $(i = 1, 2)$ are the marginal Kaplan-Meier estimators of $F_i(\cdot)$ $(i = 1, 2)$, respectively. Table 1 lists the values of a_{ij} given in Brown *et al.* (1974). The values of b_{ij} are similarly defined.

Table 1. Values of a_{ij} of Brown *et al.*'s estimator.

$(\delta_{1i}, \delta_{1j})$	$X_i > X_j$	$X_i = X_j$	$X_i < X_j$
(1, 1)	1	0	-1
(0, 1)	1	1	$2\{\hat{\mathbb{F}}_1(x_j)/\hat{\mathbb{F}}_1(x_i)\} - 1$
(1, 0)	$1 - 2\{\hat{\mathbb{F}}_1(x_j)/\hat{\mathbb{F}}_1(x_i)\}$	-1	-1
(0, 0)	$1 - \{\hat{\mathbb{F}}_1(x_j)/\hat{\mathbb{F}}_1(x_i)\}$	$1 - \{\hat{\mathbb{F}}_1(x_j)/\hat{\mathbb{F}}_1(x_i)\}$	$\{\hat{\mathbb{F}}_1(x_j)/\hat{\mathbb{F}}_1(x_i)\} - 1$

To normalize the measure to lie between $[-1, 1]$, Brown *et al.* (1974) adopted Γ as their estimate of τ, denoted as $\hat{\tau}_B$. Note that this method takes partial information provided by the Kaplan-Meier estimates into account. For singly censored observations, as illustrated in Figure 1 (iv) and (v), this approach seems quite intuitive for determining the unknown relationship. However for pairs with doubly censored observations, as in Figure 1 (vi)-(viii), the modifications may not be sufficient because joint information is ignored.

Weier and Basu (1980) discussed other ways of modifying the scores. One alternative they proposed was to impute censored observations by their expected values under Kaplan-Meier estimates. All methods, as discussed in Weier and Basu (1980), suffer from the same drawback as $\hat{\tau}_B$, since only marginal adjustments are made. It should be mentioned that the joint relationship becomes more important when the association becomes higher. Consequently these estimators are inconsistent under $\tau \neq 0$ and have bias increasing in τ.

Oakes (1982) proposed an estimator of τ for testing the independence hypothesis. It has the form (2.1) with $a_{ij} = 1$ if $\delta_{1i} = 1$ and $X_i < X_j$; $a_{ij} = -1$ if $\delta_{1j} = 1$ and $X_j < X_i$; $a_{ij} = 0$ otherwise (b_{ij} defined similarly). Notice that $a_{ij}b_{ij} \neq 0$ only when the relative concordance/discordance relationship is certain (such as (i) - (iii) in Figure 1). Oakes' estimator ignores partial information, provided by censored data, that can be unreliable when $\tau \neq 0$ and some data are censored.

3. New Estimators of τ

3.1. The V-statistic approach

When (T_1, T_2) are both continuous positive random variables,

$$\tau = 4\text{pr}(T_{1i} > T_{1j}, T_{2i} > T_{2j}) - 1 = 4\int_0^\infty \int_0^\infty F(x,y)F(dx,dy) - 1. \qquad (3.1)$$

Therefore one can write $\tau = \mathcal{T}(F)$, where $\mathcal{T} : \mathcal{D}[\mathcal{S}_F] \to I\!R$ and $\mathcal{D}[\mathcal{S}_F]$ is the space of cadlag functions on $\mathcal{S}_F$. When there is no censoring τ can be estimated by $\mathcal{T}(\bar{I\!F})$, where $\bar{I\!F}$ is the empirical estimator of F. $\mathcal{T}(\bar{I\!F})$ has the form of a so-called V-statistic, which differs from the U-statistic estimator in (2.1) only at order $O(n^{-3})$ (Serfling (1980)). Asymptotic properties of $\mathcal{T}(\bar{I\!F})$ can be developed by applying a Taylor series expansion on $\mathcal{T}(\cdot)$ (Fernholz (1983)). The technique was first introduced by von Mises (1947) and then further extended to the so-called "functional δ method" by Gill (1989). Generally speaking, if the estimator of F is a reasonable estimator and $\mathcal{T}(\cdot)$ is a "differentiable" functional (specifically, compactly or Hadamard differentiable), $\mathcal{T}(\bar{I\!F})$ will inherit nice properties of $\bar{I\!F}$. In our case, it is easy to show that $\mathcal{T}(F) = 4\int F\, dF - 1$ is compactly differentiable with F being a *cadlag* function of bounded variation. Since $\bar{I\!F}$ is strongly consistent and asymptotically normal (at a rate $n^{-1/2}$), nice properties of $\mathcal{T}(\bar{I\!F})$ follow.

When either T_1 or T_2 has a discrete component the data may have ties and (3.1) has to be modified. Notice that by the law of total probability, $\mathrm{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j})>0\}+\mathrm{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j})<0\}+\mathrm{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j})=0\}=1$. Based on the spirit of the unconditional tau, a modified population version of tau for ties is

$$\tilde{\tau}=4\mathrm{pr}\{T_{1i}>T_{1j},T_{2i}>T_{2j}\}-1+\mathrm{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j})=0\}. \tag{3.2}$$

Another alternative, based on the spirit of Γ in (2.2), is

$$\gamma=\frac{\tilde{\tau}}{1-\mathrm{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j})=0\}}. \tag{3.3}$$

Note $\gamma \geq \tilde{\tau}$. Since T_{ki} and T_{kj} $(k=1,2)$ are independent and have identical distributions, it follows that $\mathrm{pr}(T_{ki}=T_{kj})=\sum_{s\in\Omega_k}\mathrm{pr}(T_k=s)^2$, and $\mathrm{pr}(T_{1i}=T_{1j},\ T_{2i}=T_{2j})=\sum_{s\in\Omega_1}\sum_{t\in\Omega_2}\mathrm{pr}(T_1=s,T_2=t)$, where Ω_k $(k=1,2)$ are the sets in which T_k take on their discrete values. With complete data, $\mathrm{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j})=0\}$ can be easily estimated by empirical estimates. For example $\mathrm{pr}(T_k=s)$ can be estimated by $\sum_{i=1}^n I\!I(T_{ki}=s)/n$ for $s\in O_k$ where O_k is the set of observed tied points in T_k $(k=1,2)$.

As mentioned earlier, a variety of nonparametric estimators of $F(x,y)$ have been proposed. These bivariate survival estimators can serve as candidates of the plug-in estimator, denoted $\hat{I\!F}$. Let $x_{(0)}=y_{(0)}=0$, $x_{(1)}<\cdots<x_{(n)}$ and $y_{(1)}<\cdots<y_{(n)}$ be ordered observations of X and Y and $\delta_{1:(j)}$ and $\delta_{2:(k)}$ be the corresponding censoring indicators of $X_{(j)}$ and $Y_{(k)}$, respectively. Without ties, one can write $\mathcal{T}(\hat{I\!F})=4\sum_{i=1}^n\sum_{j=1}^n\hat{I\!F}(x_{(i)},y_{(j)})\hat{I\!F}(\triangle x_{(i)},\triangle y_{(j)})-1$, where $\hat{I\!F}$ is a bivariate survival estimator and $\hat{I\!F}(\triangle x_{(i)},\triangle y_{(j)})=\hat{I\!F}(x_{(i)},y_{(j)})-\hat{I\!F}(x_{(i)},y_{(j-1)})-\hat{I\!F}(x_{(i-1)},y_{(j)})+\hat{I\!F}(x_{(i-1)},y_{(j-1)})$ is the estimated mass on the rectangle $[x_{(i-1)},x_{(i)}]\times[y_{(j-1)},y_{(j)}]$. Thus $\hat{I\!F}(\triangle x_{(i)},\triangle y_{(j)})$ is the estimated mass at $(x_{(i)},y_{(j)})$ when $\hat{I\!F}$ is discrete at data points. Note that $\hat{I\!F}(x_{(i)},0)=\hat{I\!F}_1(x_{(i)})$ and $\hat{I\!F}(0,y_{(j)})=\hat{I\!F}_2(y_{(j)})$ and that, for most existing estimators of the survival function, $\hat{I\!F}_k(\cdot)$ reduces to the Kaplan-Meier estimator of $F_k(\cdot)$ $(k=1,2)$. When ties are present, the probability of $\mathrm{pr}\{(T_{1i}-T_{1j})(T_{2i}-T_{2j})=0\}$ can also be estimated under censoring. Specifically $\mathrm{pr}(T_k=s)$ can be estimated by $\hat{I\!F}_k(s-)-\hat{I\!F}_k(s)$ $(k=1,2)$ and $\mathrm{pr}(T_1=s,T_2=t)$ can be estimated by $\hat{I\!F}(\Delta s,\Delta t)$.

To simplify the analysis, we assume T_1 and T_2 are both continuous. As in the univariate case where Kaplan-Meier integrals are biased (Stute (1994)), $\mathcal{T}(\hat{I\!F})$ is an asymptotically biased estimator of τ. Specifically, it may happen that $\hat{I\!F}$ does not go to zero as x or y goes to ∞ if $\delta_{1:(n)}=0$ or $\delta_{2:(n)}=0$ and, as a result, $\sum_{i=1}^n\sum_{j=1}^n\hat{I\!F}(\triangle x_{(i)},\triangle y_{(j)})<1$. Recall $\mathcal{S}_H=\{(x,y):H(x,y)>0\}$

and $\mathcal{S}_F = \{(x,y) : F(x,y) > 0\}$. Under right censoring, $\mathcal{S}_H \subseteq \mathcal{S}_F$. It should be mentioned that asymptotic properties of the aforementioned bivariate survival estimators are valid only in $\mathcal{S}_H$.

Define $\tau_0 = 4\int\int_{\mathcal{S}_H} F(x,y)F(dx,dy) - 1$. Note that $\beta = \tau_0 - \tau$ measures the bias of τ_0 from τ. It is easy to see that $\beta \leq 0$ with equality when $\mathcal{S}_H = \mathcal{S}_F$. Since no data can be obtained outside the range of $\mathcal{S}_H$, β is not identifiable. It follows that

$$\mathcal{T}(\hat{I\!\!F}) = 4\int_0^\infty\int_0^\infty \hat{I\!\!F}(x,y)\hat{I\!\!F}(dx,dy) - 1 = 4\int\int_{\mathcal{S}_H} \hat{I\!\!F}(x,y)\hat{I\!\!F}(dx,dy) - 1, \quad (3.4)$$

which implies that $\mathcal{T}(\hat{I\!\!F})$ actually estimates τ_0 instead of τ. From now on we let $\hat{\tau}_0 = \mathcal{T}(\hat{I\!\!F})$. The asymptotic properties of $\hat{\tau}_0$ are stated in the following theorem.

Theorem 1. *Suppose (X,Y) are jointly continuous. If $\hat{I\!\!F}(x,y)$ is a strongly consistent estimator of $F(x,y)$ on $\mathcal{S}_H = \{(x,y) : H(x,y) > 0\}$, and $H_j^{-1}(t)$ $(j = 1,2)$ are continuous at $t = 0$ then $\hat{\tau}_0$ converges to τ_0 $(\leq \tau)$ in probability and, if $n^{1/2}\{\hat{I\!\!F}(x,y) - F(x,y)\}$ converges weakly to a mean-zero Gaussian process for $(x,y) \in \mathcal{S}_H$, $n^{1/2}(\hat{\tau}_0 - \tau_0)$ converges to a mean-zero normal random variable.*

3.2. Estimable bounds on the bias

We study bounds on the bias that can be easily estimated with reasonable precision. Sharper bounds on τ, such as those on Kaplan-Meier integrals (for a review, see Stute (1994)), may be derived in a more rigorous fashion. First note that

$$|\beta| \leq 4\ \mathrm{pr}(\mathcal{S}_F \backslash \mathcal{S}_H) \sup_{(x,y):\mathcal{S}_F\backslash\mathcal{S}_H} |F(x,y)|\,, \quad (3.5)$$

where $A\backslash B = A \cap B^c$ and B^c is the complement of B. Without loss of generality, assume $\mathcal{S}_F = [0,\infty)^2$. Precise knowledge on $\mathcal{S}_H$ is crucial in constructing a bound on $F(x,y)$ in $\mathcal{S}_F\backslash\mathcal{S}_H$. Generally $\mathcal{S}_H$ is unknown in the nonparametric setting. But, given the data, one can obtain some information about it.

Let $\xi_j = \sup\{t : H_j(t) > 0,\ 0 \leq t < \infty\}$ $(j = 1,2)$ and $cl(\mathcal{S}_H)$ be the closure of $\mathcal{S}_H$. Partition $\mathcal{S}_F\backslash cl(\mathcal{S}_H)$ into four disjoint sets (see Figure 2), $\mathcal{C}_1 \equiv (\xi_1,\infty) \times [0,\xi_2]$, $\mathcal{C}_2 \equiv [0,\xi_1] \times (\xi_2,\infty)$ $\mathcal{C}_3 \equiv (\xi_1,\infty) \times (\xi_2,\infty)$ and $I\!\!R_H \equiv [0,\infty)^2\backslash\{cl(\mathcal{S}_H) \cup \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3\}$.

From the decomposition of $\mathcal{S}_F\backslash cl(\mathcal{S}_H)$ as $\mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \cup I\!\!R_H$, it follows that

$$F(x,y) \leq \begin{array}{ll} F_1(\xi_1) & \text{if } (x,y) \in \mathcal{C}_1 \\ F_2(\xi_2) & \text{if } (x,y) \in \mathcal{C}_2 \\ F(\xi_1,\xi_2) & \text{if } (x,y) \in \mathcal{C}_3 \\ \max_{(x,y)\in\tilde{\mathcal{S}}_H} F(x,y) & \text{if } (x,y) \in I\!\!R_H, \end{array} \quad (3.6)$$

where $\tilde{\mathcal{S}}_H$ is the boundary of $\mathcal{S}_H$. From (3.4) and (3.5),

$$|\beta| \le 4\,\{F_1(\xi_1)\mathrm{pr}(\mathcal{C}_1) + F_2(\xi_2)\mathrm{pr}(\mathcal{C}_2) + F(\xi_1,\xi_2)\mathrm{pr}(\mathcal{C}_3) + \max_{(x,y)\in\tilde{\mathcal{S}}_H} F(x,y)\mathrm{pr}(I\!\!R_H)\}. \tag{3.7}$$

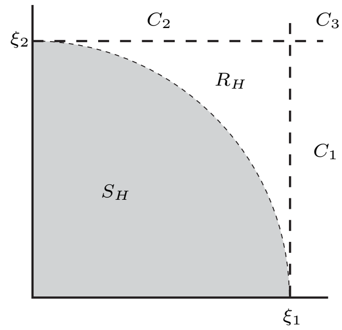

Figure 2. Illustration of the sets in construction of the bounds.

The bound of β in (3.7) can be estimated. Note that the endpoints ξ_1 and ξ_2 can be consistently estimated by the maximum order statistics $x_{(n)}$ and $y_{(n)}$, respectively. The Glivenko-Cantelli Theorem implies the probabilities of the sets, $cl(\mathcal{S}_H)$, $\mathcal{C}_k$ $(k=1,2,3)$ and $I\!\!R_H$ can be consistently estimated by $\hat{\mathrm{pr}}(cl(\mathcal{S}_H)) = \sum_{i=1}^n\sum_{j=1}^n \hat{I\!\!F}(\triangle x_{(i)},\triangle y_{(j)})$, $\hat{\mathrm{pr}}(\mathcal{C}_1) = \hat{I\!\!F}_1(x_{(n)}) - \hat{I\!\!F}(x_{(n)},y_{(n)})$, $\hat{\mathrm{pr}}(\mathcal{C}_2) = \hat{I\!\!F}_2(y_{(n)})$ $-\hat{I\!\!F}(x_{(n)},y_{(n)})$, $\hat{\mathrm{pr}}(\mathcal{C}_3) = \hat{I\!\!F}(x_{(n)},y_{(n)})$, and

$$\hat{\mathrm{pr}}(I\!\!R_H) = 1 - \{\sum_{i=1}^n\sum_{j=1}^n \hat{I\!\!F}(\triangle x_{(i)},\triangle y_{(j)}) + \hat{\mathrm{pr}}(\mathcal{C}_1) + \hat{\mathrm{pr}}(\mathcal{C}_2) + \hat{\mathrm{pr}}(\mathcal{C}_3)\},$$

respectively. Note that $\max_{(x,y)\in\tilde{\mathcal{S}}_H} F(x,y)$ can be estimated by $\max_{(i,j)\in\check{\mathcal{S}}_H}$ $\hat{I\!\!F}(x_i,y_j)$, where $\check{\mathcal{S}}_H = \{(x_i,y_j) : \hat{H}(x_i,y_j) = 0,\ (i,j=1,\ldots,n)\}$. Therefore, based on (3.6), a bound for τ can be estimated by $[\hat{\tau}_0,\hat{\tau}_1]$, where

$$\begin{aligned}\hat{\tau}_1 = \hat{\tau}_0 + 4\Big\{&\hat{I\!\!F}_1(x_{(n)})\hat{\mathrm{pr}}(\mathcal{C}_1) + \hat{I\!\!F}_1(y_{(n)})\hat{\mathrm{pr}}(\mathcal{C}_2) + \hat{I\!\!F}(x_{(n)},y_{(n)})\hat{\mathrm{pr}}(\mathcal{C}_3)\\ &+ \max_{(i,j)\in\check{S}_H} \hat{I\!\!F}(x_i,y_j)\hat{\mathrm{pr}}(I\!\!R_H)\Big\}.\end{aligned} \tag{3.8}$$

When $\delta_{1:(n)} = \delta_{2:(n)} = 1$ it follows that $\hat{\tau}_0 = \hat{\tau}_1$ since, for most existing estimators of $F(x,y)$, $\hat{I\!\!F}_1(x_{(n)}) = \hat{I\!\!F}_2(y_{(n)}) = 0$, $\hat{I\!\!F}(x,y_{(n)}) = \hat{I\!\!F}(x_{(n)},y) = 0$ and hence $\hat{\mathrm{pr}}(I\!\!R_H) = 0$. In such cases the bounds shrink to a point estimate of τ. In general the length of the bound depends on the sample size, the censoring proportion and in particular the magnitude of $\hat{I\!\!F}_1(x_{(n)})$, $\hat{I\!\!F}_2(y_{(n)})$ and $\hat{I\!\!F}(x_{(n)},y_{(n)})$. Chen and Lo (1997) derived some results on the consistency of the Kaplan-Meier estimates $\hat{I\!\!F}_i(t)$ $(i=1,2)$ for t near the endpoint. Specifically they showed that $sup_{t\le\xi_i}|\hat{I\!\!F}_i(t)-F_i(t)|=o(n^{-p})$ if for $0<p<1/2$,

$$\int_0^{\xi_i}(1-G_i)^{-\frac{p}{1-p}}dF<\infty \quad (i=1,2), \tag{3.9}$$

where G_i is the survival function of the censoring variable C_i. The size of p in (3.9) essentially reflects the heaviness of the censoring in the tail region. The smaller the p the less the uncensored observations are near the endpoint. This in turn is reflected in the convergence rate of $\hat{I\!\!F}_i$. The consistency of $\hat{\tau}_1$ is summarized in the following proposition.

Proposition 1. *Suppose that the assumption in Theorem* 1 *and* (3.9) *are satisfied. Define $h_x(y)\equiv H(x,y)$ for a fixed x and $h_y(x)\equiv H(x,y)$ for a fixed y. Assume that $h_y^{-1}(t)$ and $h_x^{-1}(t)$ exist for any $(x,y)\in\mathcal{S}_H$ and are continuous at $t=0$. Then $\hat{\tau}_1$, defined in* (3.8)*, converges to $\tau_1=\tau_0+4\{F_1(\xi_1)pr(\mathcal{C}_1)+F_2(\xi_2)pr(\mathcal{C}_2)$ $+F(\xi_1,\xi_2)pr(\mathcal{C}_3)+\max_{(x,y)\in\tilde{\mathcal{S}}_H}F(x,y)pr(I\!\!R_H)\}\ge\tau$ in probability.*

Alternatively for $(x,y)\in\mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3$, it is easy to see that $F(x,y)\le F_1(\xi_1)\vee F_2(\xi_2)$. If

$$\int\int_{(x,y)\in\boldsymbol{R}_H}F(x,y)F(dx,dy)\le \mathrm{pr}(I\!\!R_H)\{F_1(\xi_1)\vee F_2(\xi_2)\}, \tag{3.10}$$

then a simple crude bound on β is given by

$$|\beta|\le 4\ \{F_1(\xi_1)\vee F_2(\xi_2)\}\{1-\mathrm{Pr}(\mathcal{S}_H)\}, \tag{3.11}$$

and the second bound for τ can be estimated by $[\hat{\tau}_0,\hat{\tau}_2]$, where

$$\hat{\tau}_2=\hat{\tau}_0+4\left\{\hat{I\!\!F}_1(x_{(n)})\vee\hat{I\!\!F}_2(x_{(n)})\right\}\{1-\hat{\mathrm{pr}}(\mathcal{S}_H)\}. \tag{3.12}$$

Again, the length of this bound depends on the sample size and the censoring proportion. This bound also shrinks to a point estimate of τ when the largest observations are uncensored. Under (3.9) and (3.10), $\hat{\tau}_2$ converges to τ_2 in probability, where $\tau_2=\tau_0+4\{F_1(\xi_1)\vee F_2(\xi_2)\}\{1-\mathrm{Pr}(\mathcal{S}_H)\}\ge\tau$. Note that the condition stated in (3.10) is weaker than the assumption that $\max_{(x,y)\in\tilde{\mathcal{S}}_H}F(x,y)\le F_1(\xi_1)\vee F_2(\xi_2)$. Although the marginal endpoints $(\xi_1,0)$ and $(0,\xi_2)$ both lie in $\tilde{\mathcal{S}}_H$, the latter assumption may not be true since $\tilde{\mathcal{S}}_H$ is not a contour curve of F. Furthermore both assumptions can not be verified nonparametrically.

In general $\mathcal{S}_H\subseteq[0,\xi_1)\times[0,\xi_2)$. When $\mathcal{S}_H\equiv[0,\xi_1)\times[0,\xi_2)$, that is $I\!\!R_H=\emptyset$, it is easy to see that $\hat{\tau}_1=\hat{\tau}_3\le\hat{\tau}_2$ where

$$\hat{\tau}_3=\hat{\tau}_0+4\left\{\hat{I\!\!F}_1(x_{(n)})\hat{\mathrm{pr}}(\mathcal{C}_1)+\hat{I\!\!F}_1(y_{(n)})\hat{\mathrm{pr}}(\mathcal{C}_2)+\hat{I\!\!F}(x_{(n)},y_{(n)})\hat{\mathrm{pr}}(\mathcal{C}_3)\right\}. \tag{3.13}$$

Even in the case when $\mathcal{S}_H$ is strictly contained in $[0,\xi_1)\times[0,\xi_2)$, the set $I\!R_H$ is usually much smaller than $[0,\infty)^2\backslash[0,\xi_1)\times[0,\xi_2)$. Therefore the loose bound on $F(x,y)$ for points in $[0,\infty)^2\backslash[0,\xi_1)\times[0,\xi_2)$ leaves some space for $(x,y)\in I\!R_H$. This implies that $[\hat\tau_0,\hat\tau_3]$, in most cases, still provides a bound on τ and is sharper than $[\hat\tau_0,\hat\tau_j]$ $(j=1,2)$.

4. Variance Estimation and Bias Correction Using Bootstrap

The limiting variances of $\hat\tau_j$ $(j=0,1,2,3)$ depend on the asymptotic variance of $\hat{I\!F}$, which in turn depends on the unknown F. However most existing survival function estimators are too complex to have closed forms for their asymptotic variances. This leads to the use of the bootstrap. Let $\{(x_j^*,y_j^*,\delta_{1j}^*,\delta_{2j}^*),\ (j=1,\ldots,m)\}$ be a sample of size m from the original censored data with replacement. Based on the bootstrap sample, one can compute $\hat{I\!F}^*$ and then $\hat\tau_j^*$ $(j=0,1,2,3)$. It has been shown that as $n\wedge m\to\infty$, $m^{1/2}\{\hat{I\!F}^*(x,y)-\hat{I\!F}(x,y)\}$ will mimic the asymptotic distribution of $n^{1/2}\{\hat{I\!F}(x,y)-F(x,y)\}$ on $\mathcal{S}_H$ (Dabrowska (1989)). The behavior of $m^{1/2}(\hat\tau_j^*-\hat\tau_j)$ will also mimic that of $n^{1/2}(\hat\tau_j-\tau_j)$ for $j=0,1,2,3$. The bootstrap resampling procedure can be repeated B times. The sample variance of $\hat\tau_{jb}^*$ $(b=1,\ldots,B)$ can be used to estimate the variance of $\hat\tau_j$ for $j=0,1,2,3$. There are two ways of constructing confidence intervals on τ_j. The first method is to apply the normality result and construct "t-type" confidence intervals using bootstrap variance estimates. Another alternative, which does not need the limiting normality result, is to use bootstrap percentiles. Refer to Efron and Tibshirani (1993) for detailed descriptions of both of these methods.

The bootstrap can also be used in bias estimation (Efron and Tibshirani (1993, §10). Let $\hat{I\!F}_b^*$ $(b=1,\ldots,\ B)$ be bootstrap estimates of F from B bootstrap samples. It turns out that $\hat\tau_0^*-\hat\tau_0$ can be used to estimate the bias $\hat\tau_0-\tau_0$, where $\hat\tau_0^*=\sum_{b=1}^B\mathcal{T}(\hat{I\!F}_b^*)/B$. The improvement by reducing the bias of $\hat\tau_0$ to τ_0 may still be useful, especially for small sample sizes, where $\hat\tau_0$ is a biased estimate of τ_0. We investigate the effect of bias correction via examples in the next section.

5. Examples

5.1. Simulation results

A series of Monte Carlo simulations was carried out to examine finite sample performance of $\hat\tau_j$ $(j=0,1,2,3,B)$ (defined by (3.4), (3.8), (3.12), (3.13), and the modification of (2.2) by Brown *et al.* (1974)), for cases with $\tau=0.1,0.5,0.8$ and $n=100$, 200. The replicates of the vector (T_1,T_2) were generated from Clayton's family (1978), whose survival functions are of the form

$$F(x,y)=\Big\{[\frac{1}{F_1(x)}]^{\alpha-1}+[\frac{1}{F_2(y)}]^{\alpha-1}-1\Big\}^{-1/(\alpha-1)}\ (\alpha>0), \tag{5.1}$$

where $\tau = (\alpha - 1)/(\alpha + 1)$ and $F_i(t) = \exp(-t)$ $(i = 1, 2)$. Censoring vectors (C_1, C_2) were generated from independent uniform distributions. The marginal censoring rate was about 30% in both dimensions, and double censoring varied from 10% to 15% as τ increased from 0.1 to 0.8. The estimator of τ proposed by Oakes (1982) was also studied but these results are not included: Oakes' estimator performed well when τ was near zero but its bias became substantial when τ got larger.

Table 2. Performance of estimators of τ with $n = 200$. In each cell the top number $(\times 10^{-2})$ is the average of $\hat{\tau}_j - \tau$ and the number in the parenthesis $(\times 10^{-2})$ is the standard deviation of $\hat{\tau}_j$ based on $1,000$ replications.

	$\tau = 0.8$	$\tau = 0.5$	$\tau = 0.1$
$\hat{\tau}_0$	-1.2 (5.5)	-1.2 (5.1)	-1.2 (5.4)
$\hat{\tau}_1$	0.6 (5.6)	0.7 (5.5)	1.2 (5.7)
$\hat{\tau}_2$	0.9 (5.4)	1.0 (5.3)	1.5 (5.8)
$\hat{\tau}_B$	-9.7 (2.2)	-5.7 (4.1)	-1.3 (5.3)

Note that in all simulations, $\hat{\tau}_1$ is very close to $\hat{\tau}_3$ and only the result on $\hat{\tau}_1$ is reported. Tables 2 and 3 list the results for $n = 100$ and $n = 200$, respectively, based on $1,000$ simulation runs. In each cell, the first number $(\times 10^{-2})$ is the average of $\hat{\tau}_j - \tau$ $(j = 0, 1, 2, B)$ and the number in parenthesis $(\times 10^{-2})$ is the standard deviation of $\hat{\tau}_j$. As Theorem 1 indicates, $\hat{\tau}_0 - \tau \leq 0$ and $\hat{\tau}_j - \tau \geq 0$ $(j = 1, 2)$. The length of the two estimated bounds can be obtained by calculating $\hat{\tau}_j - \hat{\tau}_0$ $(j = 1, 2)$. It is easy to see that the length of the bounds increases as the value of τ decreases and the sample size increases. On the average $[\hat{\tau}_0, \hat{\tau}_1]$ is shorter than $[\hat{\tau}_0, \hat{\tau}_2]$. Notice that $\hat{\tau}_B$ performs well at $\tau = 0.1$ but it becomes more biased as the value of τ increases. The proposed bounds provide much more accurate information on τ for moderate and high correlation in both sample sizes.

We also studied the effect of using the bootstrap method to correct the bias of $\hat{\tau}_0$. Let

$$\hat{\tau}_{bc} = \hat{\tau}_0 - \{\sum_{b=1}^{B} [\mathcal{T}(\hat{\mathbb{F}}^*_b)/B - \hat{\tau}_0]\}$$

be the bias-corrected estimate of τ_0. We hope that $\hat{\tau}_{bc}$ is closer to τ than $\hat{\tau}_0$ is. We studied three cases $\tau = 0.1, 0.5, 0.8$ with $n = 100$, $m = 100$ and $B = 200$. Because of the extensive computing time, we ran only 200 replications. The average of $\hat{\tau}_{bc} - \tau$ $(\times 10^{-2})$ is -1.0, -0.2 and -0.8 for $\tau = 0.8, 0.5, 0.1$, respectively. The standard deviation of $\hat{\tau}_{bc}$ ($\times 10^{-2}$) is 7.2, 7.9 and 8.8 for $\tau = 0.8, 0.5, 0.1$, respectively. Compared with the results in Table 1 and 2 we see that $\hat{\tau}_{bc}$, calculated from an original sample with $n = 100$, is even less biased than $\hat{\tau}_0$ with $n = 200$ and various degrees of association. However $\hat{\tau}_{bc}$ tends to have larger variation. It seems worthwhile to use the bootstrap to correct bias for $\tau = 0.5$ and 0.8.

5.2. Illustrative examples

Two data sets were analyzed, with Dabrowska's estimator used as the plug-in estimator of F. The first data set was from a study on the length of exercise time required to induce angina pectoris in 21 heart disease patients (Danahy, Burwell, Aranow and Prakash (1977)). Here T_1 is the exercise time to angina pectoris at time 0 and T_2 is the exercise time to angina pectoris 3 hours after taking oral isosorbide dinitrate. Only 4 observations of T_2 were censored due to patient fatigue. Since the largest observations of T_1 and T_2 are both observed, the proposed approach also yields a point estimate. Two observations of T_1 are tied with $O_1 = \{250\}$ and the estimated probability of $\mathrm{pr}(T_1 = 250)$ is $(2/21)^2$. After adjusting for ties, the proposed estimate of $\tilde{\tau}$ is $0.384 + 0.009 = 0.393$ and the estimate of γ, in (3.3), is 0.397. The proposed bootstrap bias-corrected estimate of $\tilde{\tau}$ is 0.399 and that of γ is 0.40. However $\hat{\tau}_B = 0.48$, which is very different from the proposed estimates. Without knowing the true value of τ, it seems impossible to compare the two estimates. Nevertheless, the model selection procedure proposed by Genest and Rivest (1993) allows one to informally evaluate relative accuracy of the estimates. Specifically, suppose that (T_1, T_2) belongs to the Archimedean copula class whose survival function is of the form

$$F(x, y) = \phi_\alpha^{-1}\{\phi_\alpha(F_1(x)) + \phi_\alpha(F_2(y))\}, \tag{5.2}$$

for some convex decreasing function $\phi_\alpha(\cdot)$ satisfying $\phi_\alpha(1) = 0$, α an association parameter. They showed that

$$\tau = 4\int_0^1 \lambda_\alpha(v)dv + 1, \tag{5.3}$$

where $\lambda_\alpha(v) = v - \mathrm{pr}(F_\alpha(T_1, T_2) \le v)$ and $\phi_\alpha(v) = \exp\{\int_{v_0}^{v} 1/\lambda_\alpha(t)dt\}$, where $0 < v_0 < 1$ is an arbitrary constant. Wang and Wells (1997) proposed a nonparametric estimator of $\lambda(v)$ for bivariate censored data. We found that theoretical curves of $\lambda_{\hat{\alpha}}(v)$ for some selected models with $\hat{\alpha}$ inverted using $\hat{\tau} = 0.393$ are much closer to the nonparametric estimate of $\lambda(v)$ than those with $\hat{\alpha}$ inverted using $\hat{\tau} = 0.484$. This suggests that the proposed estimate of τ is more accurate than $\hat{\tau}_B$.

The second data set (McGilchrist and Aisbett (1991)) is from a study of the recurrence time of infection in kidney patients who are using a portable dialysis machine. Two successive recurrence times, measured from insertion until the next infection, were recorded. The catheter must be removed if an infection occurs. After infection cleared up, the catheter was then reinserted. Censoring may be due to removal for other reasons or the end-of-study effect (for the second infection). Let T_1 be the time to the first infection and T_2 be the time to the second. There are 38 observations, 6 observations of T_1 were censored, 12 observations of T_2 were censored, and 3 observations were doubly censored. Since the largest observations of X and Y are both observed, the proposed method also produces a point estimate. Observations of X with $\delta_1 = 1$ and Y with $\delta_2 = 1$ both have ties, specifically $O_1 = \{7, 15, 152\}$ and $O_2 = \{30\}$. Using the proposed method, the estimated $\mathrm{pr}(T_{1i} = T_{1j} \text{ or } T_{2i} = T_{2j})$ is 0.022, the estimate of $\tilde{\tau}$ is 0.213 and that for γ, in (3.3), is $0.213/0.978 = 0.218$. The bootstrap bias-corrected estimates of $\tilde{\tau}$ and γ are almost identical to the uncorrected version. The estimated standard deviation of the proposed estimate is 0.072. The estimate by Brown *et al.* is $\hat{\tau}_B = 0.209$, quite close to the proposed estimate of γ. Notice that in this data set 3 observations (out of 38) are doubly censored but τ is low. That is why the proposed estimator does not show much improvement.

6. Discussion

Although Kendall's tau is an important quantity of interest in many applications, until now there has not been a practical estimator of τ under censored data. Previous estimators fail to account for joint information and are not reliable if the degree of association is above a moderate level. The proposed method imposes estimable bounds on τ which can produce a consistent estimator if the largest observations in both dimensions are uncensored. The lengths of the bounds depend heavily on the estimated tail probabilities of F and F_j $(j = 1, 2)$. Then if some proportion of large observations are censored, the bounds can be very wide and contain very little information about τ. This is the main drawback of the proposed approach. It should be mentioned that we encountered negative mass when using the Dabrowska estimator. The negative mass problem has been pointed out by Pruitt (1991) and is actually common to most bivariate estimators mentioned earlier, except for the so called "repaired nonparametric MLE" proposed by van der Laan (1996). We think that negative mass is not serious in estimating τ. Specifically we observe that $\int_0^x \int_0^y \hat{\mathbb{F}}(du, dv)$ is a good estimate of $F(x, y)$ even when some $\hat{\mathbb{F}}(du, dv)$ are negative. It seems that statistics of the form $\int\int \phi(u, v) \hat{\mathbb{F}}(du, dv)$ can still be good estimates of $\int\int \phi(u, v) F(du, dv)$, since the effect of negative mass can be averaged out. However negative mass may cause problems in computing the probability of tied observations. In such a case improvement is possible if a proper survival function estimate of F is used.

Acknowledgements

The support of NSC 87-2118 and NSF Grant DMS 9625440 is gratefully acknowledged. The authors thank Professor Phillip Hougaard for providing the reference for the first data set, and the referee for helpful comments.

Appendix

Proof of Theorem 1

Proof of Consistency. One has

$$1/4\{\hat{\tau}_0 - \tau_0\} = \int\int_{\mathcal{S}_H} \hat{I\!\!F}(x,y)\hat{I\!\!F}(dx,dy) - \int\int_{\mathcal{S}_H} F(x,y)F(dx,dy)$$
$$= \int\int_{\mathcal{S}_H} \{\hat{I\!\!F}(x,y) - F(x,y)\}\hat{I\!\!F}(dx,dy) + \int\int_{\mathcal{S}_H} F(x,y)\{\hat{I\!\!F}(dx,dy) - F(dx,dy)\}.$$

By strong consistency of $\hat{I\!\!F}$ on $\mathcal{S}_H$, and applying the Bounded Convergence Theorem, the first term in the above equation converges to zero. To prove convergence of the second term (to zero) when $\mathcal{S}_H$ is a rectangle, one can apply integration by parts successively to make functions of $\{\hat{I\!\!F}(x,y) - F(x,y)\}$ appear in the integrand and then use the previous arguments. When $\mathcal{S}_H$ is not a rectangle, one can approximate it by a union of small rectangles so that the integration by parts formula can be applied. The approximation will approach the target region $\mathcal{S}_H$ as the mesh of rectangle grids becomes increasingly finer. Then $\hat{\tau}_0 \stackrel{P}{\rightarrow} \tau_0$ as $n \rightarrow \infty$ can be proved.

Proof of Normality. Let $\hat{W}(x,y) = n^{1/2}\{\hat{I\!\!F}(x,y) - F(x,y)\}$, weakly convergent to a zero mean Gaussian process on $\mathcal{S}_H$. Denote the limiting process of $\hat{W}(x,y)$ by $W(x,y)$. It can be shown that

$$n^{1/2}\{\hat{\tau}_0 - \tau_0\} = \int\int_{\mathcal{S}_H} \hat{W}(x,y)F(dx,dy) + \int\int_{\mathcal{S}_H} F(x,y)\hat{W}(dx,dy)$$
$$+ \int\int_{\mathcal{S}_H} \hat{W}(x,y)\{\hat{I\!\!F}(dx,dy) - F(dx,dy)\}.$$

Weak convergence of $\hat{W}(x,y)$ to $W(x,y)$ on $\mathcal{S}_H$ implies that

$$\int\int_{\mathcal{S}_H} \hat{W}(x,y)F(dx,dy) \Rightarrow \int\int_{\mathcal{S}_H} W(x,y)F(dx,dy)$$

and

$$\int\int_{\mathcal{S}_H} F(x,y)\hat{W}(dx,dy) \Rightarrow \int\int_{\mathcal{S}_H} F(x,y)W(dx,dy).$$

Now the remaining work is to show that $\left|\int\int_{\mathcal{S}_H}\hat{W}(x,y)\{\hat{I\!\!F}(dx,dy)-F(dx,dy)\}\right|$ $\to 0$. Without loss of generality, let $\mathcal{S}_H=[0,\xi_1)\times[0,\xi_2)$. If $\mathcal{S}_H$ is not a rectangle we can apply the argument in the previous proof. Let $V(x,y)=\{\hat{I\!\!F}(x,y)-F(x,y)\}$ and by successive integration by parts, we can obtain the following expression:

$$\int_0^{\xi_1}\int_0^{\xi_2}n^{1/2}V(x,y)V(dx,dy)=\int_0^{\xi_1}n^{1/2}V(x,\xi_2)V(dx,\xi_2)-\int_0^{\xi_1}n^{1/2}V(x,0)V(dx,0)$$
$$+\int_0^{\xi_2}n^{1/2}V(\xi_1,y)V(\xi_1,dy)-\int_0^{\xi_2}n^{1/2}V(0,y)V(0,dy)-\int_0^{\xi_1}\int_0^{\xi_2}n^{1/2}V(x,y)V(dx,dy).$$

Notice that by one more integration by parts in the first term of the above,

$$\int_0^{\xi_1}n^{1/2}V(x,\xi_2)V(dx,\xi_2)=n^{1/2}\{V(\xi_1,\xi_2)^2-V(0,\xi_2)^2\}-\int_0^{\xi_1}n^{1/2}V(x,\xi_2)V(dx,\xi_2).$$

Since $n^{1/2}V(\xi_1,\xi_2)^2\to 0$ and $n^{1/2}V(0,\xi_2)^2\to 0$, it follows that $\int_0^{\xi_1}n^{1/2}V(x,\xi_2)$ $V(dx,\xi_2)\to 0$. Using similar techniques, one can show that the second to the fourth terms in (A.1) converge to zero, and consequently $\int_0^{\xi_1}\int_0^{\xi_2}n^{1/2}V(x,y)$ $V(dx,dy)\to 0$. This completes the proof.

Proof of Proposition 1. It is easy to see that $\tau_1\geq\tau$. One can write $x_{(n)}=\hat{H}_1^{-1}(0)$ and $y_{(n)}=\hat{H}_2^{-1}(0)$, where $\hat{H}_j$ $(j=1,2)$ are the empirical survival functions of X and Y, respectively. By the elementary Glivenko-Cantelli theorem and continuity of $H_j^{-1}(t)$ $(j=1,2)$ at $t=0$, it follows that $\lim_{n\to\infty}\hat{H}_j^{-1}(0)=H_j^{-1}(0)=\xi_j$ $(j=1,2)$. Then under (3.9), the strong consistency of the Kaplan-Meier estimator follows from Theorem 1 of Chen and Lo (1997). With the continuity of F_j, $\hat{I\!\!F}_1(x_{(n)})\to F_1(\xi_1)$, $\hat{I\!\!F}_2(y_{(n)})\to F_2(\xi_1)$ and $\hat{I\!\!F}_1(x_{(n)},y_{(n)})\to F(\xi_1,\xi_2)$. By the continuity of $F(\cdot,\cdot)$, $F(x,y)\to F(x^*,y^*)$ if $(x^*,y^*)\to(x,y)$. With the continuity of $h_x^{-1}(t)$ and $h_y^{-1}(t)$ at $t=0$, one can show that consistency of $\hat{I\!\!F}$ on $\mathcal{S}_H$ can be extended to $\tilde{\mathcal{S}}_H$. Since $\hat{H}(\cdot,\cdot)$ converges to $H(\cdot,\cdot)$, it follows that $\max_{(i,j)\in\check{\mathcal{S}}_H}\hat{I\!\!F}(x_i,y_j)$ converges to $\max_{(x,y)\in\tilde{\mathcal{S}}_H}F(x,y)$. Because $\hat{\text{pr}}(\mathcal{C}_k)$, $\hat{\text{pr}}(\mathcal{S}_H)$ and $\hat{\text{pr}}(I\!\!R_H)$ are consistent estimates of $\text{pr}(\mathcal{C}_k)$ for $k=1,2,3$, $\text{pr}(\mathcal{S}_H)$ and $\text{pr}(I\!\!R_H)$, respectively, $\hat{\tau}_1\xrightarrow{P}\tau_1$ can be established under the assumed conditions.

References

Brown, B. W., Hollander, M. and Korwar, R. M. (1974). Nonparametric tests of independence for censored data, with applications to heart transplant studies. *Reliability and Biometry*, 327-354.

Campbell, G. (1981). Nonparametric bivariate estimation with censored data. *Biometrika* **68**, 417-423.

Chen, K. and Lo, S. H. (1997). On the rate of convergence of the product-limit estimator: strong and weak laws. *Ann. Statist.* **25**, 1050-1087.

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

Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane. *Ann. Statist.* **16**, 1475-1489.

Dabrowska, D. M. (1989). Kaplan-Meier estimate on the plane: weak convergence, LIL, and the bootstrap. *J. Multivariate Anal.* **29**, 308-325.

Danahy, D. J., Burwell, D. T., Aranow, W. S. and Prakash, R. (1977). Sustained henodynamic and antianginal effect of high dose oral isosorbide dinitrate. *Circulation* **55**, 381-387.

Davis, C. E. and Quade, D. (1968). On comparing the correlations within two pairs of variables. *Biometrics* **24**, 987-995.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* Chapman and Hall, London.

Fernholz, L. T. (1983). *Von Mises Calculus for Statistical Functions.* Springer-Verlag, New York.

Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Math.* **19**, 194-226.

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

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.* **88**, 1034-1043.

Gill, R. D. (1989). Non- and semi-parametric maximum likelihood estimators and the von Mises method (Part I). *Scand. J. Statist.* **16**, 97-128.

Gumbel, E. J. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc.* **55**, 698-707.

Hanley, J. A. and Parnes, M. N. (1983). Nonparametric estimation of a multivariate distribution in the presence of censoring. *Biometrics* **39**, 129-139.

Hoeffding, W. (1948). A class of statistics with asymptotic normal distributions. *Ann. Statist.* **19**, 293-325.

Hougaard, P., Harvald, B. and Holm, N. V. (1992). Assessment of dependence in the life times of twins. In *Survival Analysis: State of the Art*, 77-97. Kluwer Academic Publishers.

Holt, J. D. and Prentice, R. L. (1974). Survival analysis in twin studies and matched-pair experiments. *Biometrika* **61**, 17-30.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457-481.

Kendall, M. G. (1962). *Rank Correlation Methods.* Griffin, London.

Lin, D. Y. and Ying, Z. (1993). A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika* **80**, 573-581.

McGilchrist, C. A. and Aisbett, C. W. (1991). Regression with frailty in survival analysis. *Biometrics* **47**, 461-466.

Oakes, D. (1982). A concordance test for independence in the presence of censoring. *Biometrics* **38**, 451-455.

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

Prentice, R. L. and Cai, J. (1992). Covariance and survival function estimation using censored multivariate failure time data. *Biometrika* **79**, 495-512.

Pruitt, R. C. (1991). On negative mass assigned by the bivariate Kaplan-Meier estimator. *Ann. Statist.* **19**, 443-453.

Serfling, W. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, New York.

Stute, W. (1994). The bias of Kaplan-Meier integrals. *Scand. J. Statist.* **21**, 475-484.

van der Laan, M. J. (1996). Efficient estimation in the bivariate censoring model and repairing MLE. *Ann. Statist.* **2**, 596-627.

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

Wang, W. and Wells, M. T. (1997). Nonparametric estimators of the bivariate survival function under simplified censoring structures. *Biometrika* **84**, 863-880.

Wang, W. and Wells, M. T. (2000). Model selection and semi-parametric inference for bivariate failure time data. *J. Amer. Statist. Assoc.* **95**, 62-72; 75-76.

Weier, D. R. and Basu, A. P. (1980). An investigation of Kendall's τ modified for censored data with applications. *J. Statist. Plann. Inference* **4**, 381-390.

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

Email: wjwang@stat.nctu.edu.tw

Department of Social Statistics, Cornell University, Ithaca, NY 14853, U.S.A.

Email: mtw1@cornell.edu

(Received February 1999; accepted February 2000)