Ann. Inst. Statist. Math.
Vol. 47, No. 3, 525–549 (1995)

GENERALIZED CRAMÉR-VON MISES TESTS OF GOODNESS OF FIT FOR DOUBLY CENSORED DATA*

JIAN-JIAN REN

Division of Statistics, Department of Mathematics and Statistics, University of Nebraska-Lincoln, 810 Oldfather Hall, P.O. Box 880323, Lincoln, NE 86588-0323, U.S.A.

(Received February 16, 1994; revised December 12, 1994)

Abstract. We generalize Cramér-von Mises statistics to test the goodness of fit of a lifetime distribution when the data are doubly censored. We derive the limiting distributions of our test statistics under the null hypothesis and the alternative hypothesis, respectively. We also give a strong consistent estimator for the asymptotic covariance of the self-consistent estimator for the survival function with doubly censored data. Thereby, a method, called the Fredholm Integral Equation method, is proposed to estimate the null distribution of test statistics. In this work, the perturbation theory for linear operators plays an important role, and some numerical examples are included.

Key words and phrases: Cramér-von Mises statistic, doubly censored data, test of goodness of fit, limiting distribution, self-consistent estimator, survival functions.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be independent observations on a random variable (r.v.) X with a distribution function (d.f.) F. If it is wished to test the null hypothesis $H_0 : F = F_0$, where F_0 is a specified d.f., then the Cramér-von Mises test statistics are given by

$$\Psi(F_n) = \int_{-\infty}^{\infty} [F_n(x) - F_0(x)]^2 dF_0(x),$$

where F_n is the empirical d.f. based on $X_1, X_2, \ldots, X_n$.

However, in practical situations, $X_1, X_2, \ldots, X_n$ are not always available. Lifetime data are often censored. Examples of the lifetime samples' being censored either from above or below, called *doubly censored samples*, have been given by Gehan (1965), Mantel (1967), Peto (1973), Turnbull (1974), and others. Turnbull

* The author's research was supported by a Faculty Fellowship of University of Nebraska-Lincoln.

(1974) constructed a self-consistent estimator $S_X^{(n)}$ (nonparametric MLE) for the survival function $S_X(t) = P\{X > t\}(= [1 - F(x)])$ using a doubly censored sample. Chang and Yang (1987), Chang (1990), Gu and Zhang (1993) have studied the strong consistency and the weak convergence of $S_X^{(n)}$. More discussions on the doubly censored samples can be found in Tsai and Crowley (1985), and Gill (1989). In this paper, we generalize the Cramér-von Mises test statistics through the self-consistent estimator $S_X^{(n)}$ so that we can test the goodness of fit of F when the data are doubly censored.

Let X be a nonnegative random variable denoting the lifetime under investigation. In our study, one observes not $\{X_i\}$ but a doubly censored sample:

$$W_i = \max\{\min\{X_i, Y_i\}, Z_i\}$$

and

$$\delta_i = \begin{cases} 1 & \text{if } Z_i \leq X_i \leq Y_i \\ 2 & \text{if } X_i > Y_i \\ 3 & \text{if } X_i < Z_i \end{cases}$$

where (X_i, Y_i, Z_i), $i = 1, 2, \ldots, n$, are independent observations on (X, Y, Z) for nonnegative variables X, Y and Z, and the r.v.'s Y_i and Z_i are called *right* and *left censoring variables*, respectively. This means that X_i is observable whenever X_i lies in the interval $[Z_i, Y_i]$, and we know whether $X_i < Z_i$ or $X_i > Y_i$ and observe the value of Z_i or Y_i accordingly. The problem considered here is to test the goodness of fit of F based on (W_i, δ_i).

Specifically, in this paper we consider the following hypothesis test:

$$H_0 : F(t) = F_0(t), \qquad t \in [0, T] \tag{1.1}$$

where $T > 0$ is any large number. Gehan (1965) specified T as the upper limit of the observations in his generalized two-sample Wilcoxon test for doubly censored data. This means that we are interested in the goodness of fit of F within the range of our observations.

Let

$$\hat{F}_n = 1 - S_X^{(n)}.$$

Since the estimator $S_X^{(n)}$ of S_X in obtained based on (W_i, δ_i), then a natural extension of $\Psi(F_n)$ for the test (1.1) based on (W_i, δ_i) is given by

$$\psi(\hat{F}_n) = \int_0^T [\hat{F}_n(x) - F_0(x)]^2 dF_0(x). \tag{1.2}$$

Note that the functional $\psi(F) = \int_0^T [F(x) - F_0(x)]^2 dF_0(x)$ induces a functional on the space $D[0, \beta]$ (of right continuous function having left-hand limit) as below:

$$\begin{aligned} \tau(G) = \psi(G \circ F) &= \int_0^T [G(F(x)) - F_0(x)]^2 dF_0(x) \\ &= \int_0^\beta [G(t) - U_0(t)]^2 dU_0(t), \end{aligned} \tag{1.3}$$

where $\beta = F(T) \in (0,1)$, $G \in D[0,\beta]$ and $U_0 = F_0 \circ F^{-1}$. Hence, the test statistic $\psi(\hat{F}_n)$ given by (1.2) is equal to

$$\tau(\hat{U}_n) = \int_0^\beta [\hat{U}_n(t) - U_0(t)]^2 dU_0(t), \tag{1.4}$$

where $\hat{U}_n = \hat{F}_n \circ F^{-1}$.

In Section 2, we present our results on the limiting distributions of $\tau(\hat{U}_n)$ under the null hypothesis and the alternative hypothesis. Along with some discussions about our main results and the simulation results, we also give a strong consistent estimator for the asymptotic covariance of $\sqrt{n}[S_X^{(n)} - S_X]$ in Section 2. In Sections 3 and 4, we give the proofs of our theorems.

2. Main results

Denote $S_X(t) = P\{X > t\}$, $S_Y(t) = P\{Y > t\}$ and $S_Z(t) = P\{Z > t\}$. We present the self-consistent estimators of S_X, S_Y and S_Z as follows. Let (W_i, δ_i) be distributed as (W, δ), and let

$$Q_j(t) = P\{W \le t, \delta = j\}, \qquad j = 1, 2, 3, \tag{2.1}$$

$$Q_j^{(n)}(t) = \frac{1}{n}\sum_{i=1}^n I\{W_i \le t, \delta_i = j\}, \qquad j = 1, 2, 3 \tag{2.2}$$

then the estimators $S_X^{(n)}$, $S_Y^{(n)}$, $S_Z^{(n)}$ of S_X, S_Y, S_Z (Turnbull (1974), Chang and Yang (1987), Chang (1990)) are given by the solutions of the following equations:

$$Q_1^{(n)}(t) = -\int_0^t [S_Y^{(n)} - S_Z^{(n)}] dS_X^{(n)},$$

$$Q_2^{(n)}(t) = -\int_0^t S_X^{(n)} dS_Y^{(n)},$$

$$Q_3^{(n)}(t) = -\int_0^t [1 - S_X^{(n)}] dS_Z^{(n)}.$$

The solutions of $S_X^{(n)}$, $S_Y^{(n)}$, $S_Z^{(n)}$ can be calculated numerically by using the EM algorithm (Turnbull (1974), Tsai and Crowley (1985)) or by the Newton-Raphson method to find the maximum point of the log-likelihood function.

Although more general results on the weak convergence of $S_X^{(n)}$ have been established by Gu and Zhang (1993), the method of calculating the asymptotic variance of $\sqrt{n}[S_X^{(n)} - S_X]$ as processes on $[0,T]$ developed by Chang (1990) are of essential use in our study here. The following are the conditions required by Chang (1990) for the weak convergence and the calculation of the asymptotic variance of $S_X^{(n)}$.

ASSUMPTION A. (A1) The random variable X_i and the vector (Y_i, Z_i) are independent for each i and the vectors (X_i, Y_i, Z_i), $i = 1, \ldots, n$, are independently and identically distributed.

(A2) $P\{Z \leq Y\} = 1$.

(A3) $S_Y(t) - S_Z(t) > 0$ on $(0, \infty)$.

(A4) S_X, S_Y and S_Z are continuous functions of t, on $t \geq 0$, and $0 < S_X(t) < 1$ for $t > 0$.

(A5) $S_X(0) = S_Y(0) = 1$, $S_X(\infty) = S_Y(\infty) = S_Z(\infty) = 0$.

(A6) There exist δ and T, $0 < \delta < T < \infty$, such that $S_Z(t) =$ constant < 1 on $[0, \delta]$ and $S_Z(T) = 0$, i.e., $P\{Z = 0\} > 0$, $P\{Z \in (0, \delta)\} = 0$ and $P\{Z \leq T\} = 1$.

We impose the following condition on F throughout this paper.

Assumption B. F is strictly increasing on $(0, \infty)$.

Remarks. (a) (A6) implies that T in (1.1) can be chosen arbitrarily large.

(b) (A3), (A4) and (A6) imply that $S_Y(t) - S_Z(t)$ is continuous on $[0, T]$ with positive lower bound for any $0 < T < \infty$.

(c) If there is no left censoring, i.e., if $S_Z \equiv 0$, (A6) is satisfied and $S_X^{(n)}$ is the product-limit estimator of S_X in the right censored case (see Chang and Yang (1987)). Hence, all of our results in this paper hold in the right censored case. In other words, our test statistics include the right censored sample problem as a special case.

(d) Assumption B is to ensure the equivalence of (1.2) and (1.4).

For the sake of convenience, we state some results established by Chang (1990) in the following proposition.

Proposition 2.1. *Under Assumption* A, *we have that for* $t \in [0, T]$

$$\sqrt{n}[\hat{F}_n(t) - F(t)] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i(t) + o_p^{(n)}(1), \tag{2.3}$$

where $T > 0$ *is any large enough real number,* $o_p^{(n)}(1)$ *almost surely converges to* 0 *uniformly on* $[0, T]$ *as* $n \to \infty$,

$$\xi_i(t) = -\sum_{j=1}^{3} \int_0^T IC_j(t, s) d[I\{W_i \leq s, \delta_i = j\} - Q_j(s)], \tag{2.4}$$

and $IC_j(t, s)$ *is the solution of the integral equation*

$$IC_j(t, s) = F_j(t, s) + \int_0^T g(t, u, du) IC_j(u, s), \qquad j = 1, 2, 3 \tag{2.5}$$

for

$$F_1(t, s) = -\frac{I\{0 \leq s \leq t\}}{S_Y(s) - S_Z(s)}, \tag{2.6}$$

$$F_2(t, s) = \frac{I\{0 \leq s \leq t\}}{S_X(s)} \int_s^t \frac{dS_X}{S_Y - S_Z}, \tag{2.7}$$

$$F_3(t, s) = \frac{1}{1 - S_X(s)} \int_0^{s \wedge t} \frac{dS_X}{S_Y - S_Z}, \tag{2.8}$$

$$g(t, u, du) = F_2(t, u) dS_Y(u) - F_3(t, u) dS_Z(u). \tag{2.9}$$

Furthermore, if we additionally assume Assumption B, *we have that for* $t \in [0, \beta]$,

$$\sqrt{n}[\hat{U}_n(t) - t] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i(t) + o_p^{(n)}(1), \tag{2.10}$$

where $\beta = F(T)$, $\eta_i(t) = \xi_i(F^{-1}(t))$ *and* $o_p^{(n)}(1)$ *almost surely converges to* 0 *uniformly on* $[0, \beta]$ *as* $n \to \infty$.

We denote the asymptotic covariance of $\sqrt{n}[\hat{U}_n - U]$ as below:

$$\gamma(s,t) = E\{\eta_i(s)\eta_i(t)\}, \qquad (s,t) \in [0,\beta]^2$$

where U is the uniform d.f. on $[0, 1]$. We also denote λ_j as the eigenvalues for the following eigenvalue problem:

$$\int_0^\beta \gamma(s,t)\phi(t)dt = \lambda\phi(s), \qquad \phi \in L^2[0,\beta]. \tag{2.11}$$

Then we have the following properties of $\gamma(s,t)$ and λ_j's.

LEMMA 2.1. *Under Assumptions* A *and* B, *we have*:
(i) $\gamma(s,t)$ *is bivariate continuous on* $[0,\beta]^2$;
(ii) λ_j*'s are countable, and are positive real numbers with finite multiplicities.*

Our first result gives the limiting distributions of the test statistics given by (1.4) under the null hypothesis and the alternative hypothesis, respectively.

THEOREM 2.1. *Under Assumptions* A *and* B, *we have that*
(i) *under* H_0,

$$n\tau(\hat{U}_n) \xrightarrow{\mathcal{D}} \int_0^\beta T^2(t)dt = \sum_{j=1}^{\infty} \lambda_j Z_j^2, \qquad n \to \infty \tag{2.12}$$

where $T(t)$ *is a Gaussian process for* $t \in [0,\beta]$ *with mean* 0 *and covariate* $\gamma(s,t)$, *and* Z_j *are independent normal random variables with mean* 0 *and variance* 1;
(ii) *when* H_0 *does not hold*,

$$\sqrt{n}[\tau(\hat{U}_n) - \tau(U)] \xrightarrow{\mathcal{D}} N(0, \sigma_\tau^2), \qquad n \to \infty \tag{2.13}$$

where

$$\sigma_\tau^2 = \sum_{j=1}^{3} \int_0^T \left\{ \int_0^T 2[F(x) - F_0(x)] IC_j(x,y) dF_0(x) \right\}^2 dQ_j(y)$$
$$- \left\{ \sum_{j=1}^{3} \int_0^T \int_0^T 2[F(x) - F_0(x)] IC_j(x,y) dQ_j(y) dF_0(x) \right\}^2.$$

The proofs of Lemma 2.1 and Theorem 2.1 are given in Section 3.

Note that the covariance function $\gamma(s,t)$ is given by

$$\begin{aligned}\gamma(s,t) = &\sum_{j=1}^{3}\int_0^T IC_j(F^{-1}(s),x)IC_j(F^{-1}(t),x)dQ_j(x)\\ &-\left\{\sum_{j=1}^{3}\int_0^T IC_j(F^{-1}(s),x)dQ_j(x)\right\}\\ &\times\left\{\sum_{j=1}^{3}\int_0^T IC_j(F^{-1}(t),y)dQ_j(y)\right\}.\end{aligned}\tag{2.14}$$

Hence, we can easily see that even under H_0, the eigenvalues λ_j of (2.11) depend on unknown survival functions S_Y and S_Z. In order to approximate the limiting distribution of $\tau(\hat{U}_n)$ (under H_0) given by (2.12), we observe that in (2.5), we have that for $j=1,2,3$,

$$\begin{aligned}IC_j(t,s) &= F_j(t,s) + \int_0^T IC_j(u,s)F_2(t,u)dS_Y(u)\\ &\quad -\int_0^T IC_j(u,s)F_3(t,u)dS_Z(u)\\ &= F_j(t,s) - \int_0^T IC_j(u,s)F_2(t,u)g(u)du\\ &\quad + IC_j(0,s)F_3(t,0)[1-S_Z(0)]\\ &\quad + \int_0^T IC_j(u,s)F_3(t,u)h(u)du,\end{aligned}\tag{2.15}$$

where g is the density function of $1-S_Y$, $1-S_Z$ is absolutely continuous on $[\delta,\infty)$ with derivative h and we assign $h(x)=0$ for $x\in[0,\delta]$ (δ is given in (A6)). The kernel estimators of g and h based on (W_i,δ_i) are given by

$$g_n(x) = \frac{1}{a_n}\int K_1\left(\frac{x-y}{a_n}\right)d[1-S_Y^{(n)}(y)],\qquad x\in[0,\infty),\tag{2.16}$$

$$h_n(x) = \frac{1}{b_n}\int K_2\left(\frac{x-y}{b_n}\right)d[1-S_Z^{(n)}(y)],\qquad x\in[0,\infty),\tag{2.17}$$

respectively, where K_1 and K_2 are density functions, and a_n, b_n are sequences of positive numbers such that $a_n\to 0$, $b_n\to 0$, as $n\to\infty$. These kernel estimators are similar to those for right censored data (Marron and Padgett (1987), Mielniczuk (1986), among others). We obtain $F_j^{(n)}(t,s)$, $j=1,2,3$, from replacing S_X, S_Y, S_Z by their estimators $S_X^{(n)}$, $S_Y^{(n)}$, $S_Z^{(n)}$ in (2.6)–(2.8), respectively, and obtain

$IC_j^{(n)}$, $j = 1, 2, 3$, from the equations:

$$IC_j^{(n)}(t,s) = F_j^{(n)}(t,s) - \int_0^T IC_j^{(n)}(u,s)F_2^{(n)}(t,u)g_n(u)du + IC_j^{(n)}(0,s)F_3^{(n)}(t,0)[1 - S_Z^{(n)}(0)] + \int_0^T IC_j^{(n)}(u,s)F_3^{(n)}(t,u)h_n(u)du, \tag{2.18}$$

respectively. Let $\|\cdot\|$ be the supremum norm throughout the paper. Furthermore, we obtain

$$\hat{\gamma}_n(s,t) = \sum_{j=1}^{3}\int_0^T IC_j^{(n)}(F^{-1}(s),x)IC_j^{(n)}(F^{-1}(t),x)d\hat{Q}_j^{(n)}(x) - \left\{\sum_{j=1}^{3}\int_0^T IC_j^{(n)}(F^{-1}(s),x)d\hat{Q}_j^{(n)}(x)\right\} \times \left\{\sum_{j=1}^{3}\int_0^T IC_j^{(n)}(F^{-1}(t),y)d\hat{Q}_j^{(n)}(y)\right\}, \tag{2.19}$$

where $(s,t) \in [0,\beta]^2$ and $\hat{Q}_j^{(n)}$ is the continuous version of $Q_j^{(n)}$ satisfying:

$$\|\hat{Q}_j^{(n)} - Q_j^{(n)}\| \le 1/n, \tag{2.20}$$

with probability 1. We should notice that under H_0, we have $F = F_0$ in (2.19). Hence, under H_0, the eigenvalues $\lambda_j^{(n)}$ for the following eigenvalue problem:

$$\int_0^\beta \hat{\gamma}_n(s,t)\phi(t)dt = \lambda\phi(s), \qquad \phi \in L^2[0,\beta] \tag{2.21}$$

is determined through the doubly censored sample (W_i, δ_i), $i = 1, \ldots, n$. Under the following conditions on the kernel estimators, we have the theorem on the limiting distribution of $\tau(\hat{U}_n)$ under H_0.

Assumption C. Let $1 - S_Y$ have a density function g, and for $\delta > 0$ in (A6), let $1 - S_Z$ be absolutely continuous on $[\delta, \infty)$ with derivative h. For appropriate K_1, K_2, a_n and b_n, we have

$$\int_0^T |g_n(x) - g(x)|dx \to 0, \qquad \text{as} \quad n \to \infty,$$
$$\int_0^T |h_n(x) - h(x)|dx \to 0, \qquad \text{as} \quad n \to \infty,$$

with probability 1.

THEOREM 2.2. *Under Assumptions* A, B *and* C, *we have that*

(i) *for sufficiently large* n, $\hat{\gamma}_n(s,t)$ *is bivariate continuous on* $[0,\beta]^2$ *and is nonnegative*;

(ii) *for sufficiently large* n, $\lambda_j^{(n)}$ *'s are countable, and are positive real numbers with finite multiplicities*;

(iii) *with probability* 1,

$$\|\hat{\gamma}_n - \gamma\| \to 0, \qquad as \quad n \to \infty; \tag{2.22}$$

(iv) *with probability* 1,

$$\sum_{j=1}^{\infty} (\lambda_j^{(n)} - \lambda_j)^2 \to 0, \qquad as \quad n \to \infty \tag{2.23}$$

where we arrange λ_j *and* $\lambda_j^{(n)}$ *as* $\lambda_1 \geq \lambda_2 \geq \cdots$, *and* $\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \cdots$;

(v) *for those* Z_i*'s in* (2.12), *which are independent from* (W_i, δ_i), $i = 1, 2, \ldots, n$,

$$\sum_{j=1}^{\infty} \lambda_j^{(n)} Z_j^2 \xrightarrow{\mathcal{D}} \sum_{j=1}^{\infty} \lambda_j Z_j^2, \qquad as \quad n \to \infty, \tag{2.24}$$

and moreover, under H_0:

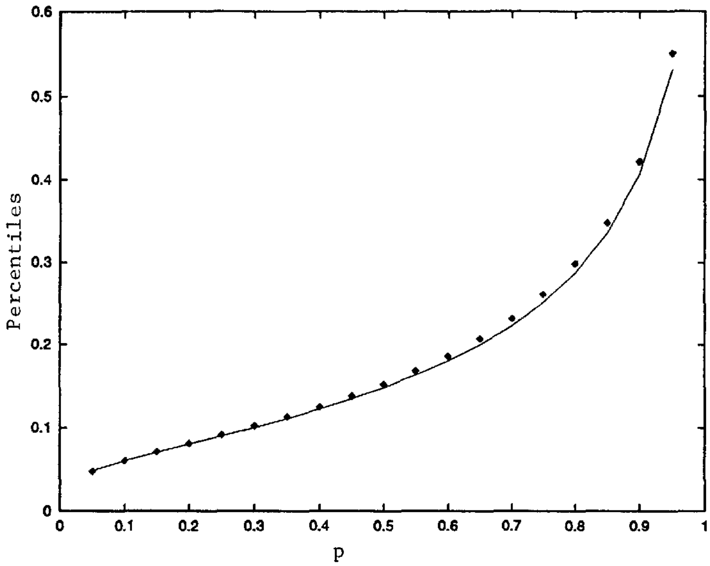
$$n\tau(\hat{U}_n) \overset{\mathcal{D}}{\approx} \sum_{j=1}^{\infty} \lambda_j^{(n)} Z_j^2, \qquad as \quad n \to \infty. \tag{2.25}$$

The proof of Theorem 2.2 is given in Section 4, where the perturbation theory for linear operators plays an important role.

Remarks. (e) The choices of K_1, K_2, a_n and b_n in Assumption C are not our concern in this study. For the special case of $S_Z \equiv 0$, Assumption C has been well established by Mielniczuk (1986), among others. For the general doubly censored case, Assumption C was established under regularity conditions by Ren (1994).

(f) There are two reasons that we use density estimation in equations (2.18). One is that for any fixed s, (2.18) are linear Fredholm integral equations of the second kind. The numerical methods for finding the solutions of this type of integral equations have been well studied. Typically, Nystrom method and Galerkin method (Delves and Walsh (1974), Delves and Mohamed (1985)) can be used to numerically calculate $IC_j^{(n)}(\cdot, s)$, $s \in [0,T]$, in (2.18). Another reason is that the use of density estimation in (2.18) avoids some difficulties encountered in our proof of Theorem 2.2 (iii). It is not clear whether the results are valid without using density estimation in (2.18).

(g) Theorem 2.2 (iii) shows that our $\hat{\gamma}_n$ given by (2.19) is a strong consistent estimator for the asymptotic covariance function γ of $\sqrt{n}[\hat{U}_n - U]$. This result is of great importance in many other statistical inference problems for doubly censored



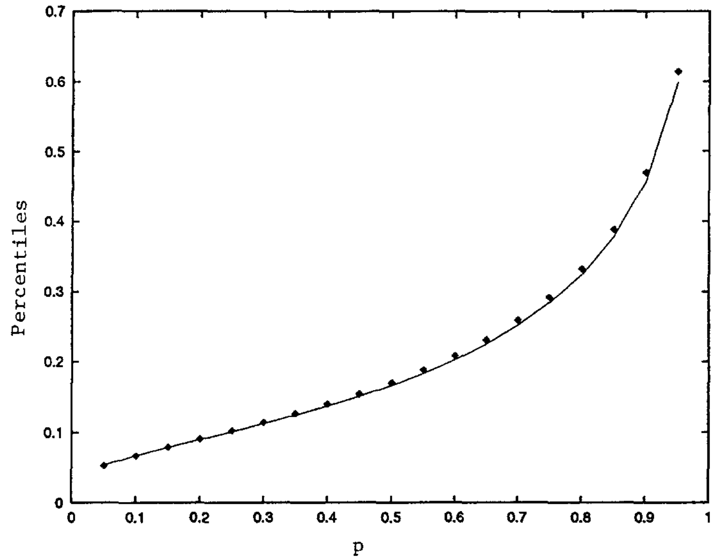
$X \sim \mathrm{Exp}(1)$, $Y \sim \mathrm{Exp}(3)$
Monte Carlo: ——— ; FIE: ······

Fig. 1.

data. For instance, the generalized L-, M- and R-estimators for doubly censored data are shown to be asymptotically normal (Ren and Zhou (1993, 1994)). The asymptotic variances of these estimators are all given through $\gamma(s,t)$. Hence, our result here provides a strong consistent estimator for the asymptotic variances of all these statistics.

(h) Based on our Theorem 2.2, in practice the estimation of the limiting distribution of $\tau(\hat{U}_n)$ may be done as follows. (1) For each $s_i \in [0,T]$, $i = 1,\ldots,M$, we can numerically calculate the solution $IC_j^{(n)}(\cdot, s_i)$ of the integral equation (2.18) using the methods mentioned above; (2) Obtain $IC_j^{(n)}(\cdot,\cdot)$ numerically through $IC_j^{(n)}(\cdot, s_i)$ using Cubic Spline Interpolation method (Burden and Faires (1989), the CSI method is available in *Mathematica*); (3) Calculate the covariance estimator $\hat{\gamma}_n$ based on formula (2.19); (4) Find finite eigenvalues $\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_N^{(n)}$ of the eigenvalue problem (2.12) numerically using Nystrom method or expansion method (Delves and Walsh (1974)), where N can be chosen arbitrarily large; (5) Generate observations $W_i^{(n)} = \sum_{j=1}^N \lambda_j^{(n)} Z_{ji}^2$, $i = 1,\ldots,m$, through computer, where Z_{ji} are i.i.d. standard normal r.v.'s, and obtain the empirical d.f. $G_m^{(n)}$ based on $W_1^{(n)},\ldots,W_m^{(n)}$; (6) Use the sample quantile estimator ${G_m^{(n)}}^{-1}(p)$, $0 < p < 1$, as the desired critical point for the test (1.1). We call our procedure described here the *Fredholm Integral Equation* (FIE) method.

In Figs. 1 and 2, we compare the distribution of $n\tau(\hat{U}_n)$ by our FIE method with that obtained by Monte Carlo method, where the special case of $S_Z \equiv 0$ and the sample size $n = 200$ of the observations are considered. In our study, the



$X \sim N(0,1)$, $Y \sim N(1,4)$
Monte Carlo: ——— ; FIE: ······

Fig. 2.

simulation results are based on 300 runs and for each run, the percentiles of $n\tau(\hat{U}_n)$ are obtained from 400 replicates of the procedure. From the figures, we can easily see that our FIE method provides very accurate estimation for the distribution of the test statistics. The routines of our FIE method are available in Fortran and may be obtained from the author. Our experiences show that our routines compute the solutions rapidly even for quite large samples, and in step (4), the number of the eigenvalues N does not have to be overly large, as in general the eigenvalues decrease rapidly. We used $N = 100$ in Figs. 1 and 2. For more details on the implementation of FIE, see Ren and Ledder (1995).

(i) An alternative method to our FIE method for the problem considered here is the nonparametric bootstrap. We should point out that the usual n out of n bootstrap fails in our problem here and that the $m = o(n)$ out of n bootstrap should be used (see Bickel and Ren (1995)). Simulation studies show that for the tests considered here, the power curve by FIE is better than that by bootstrap (Bickel and Ren (1995), Ren and Ledder (1995)). Moreover, FIE method provides a strong consistent covariance estimator $\hat{\gamma}_n$ of the self-consistent estimators $S_X^{(n)}$. Nonetheless, the bootstrap method is still quite appealing for the testing problem, since it is very easy to program.

3. Proof of Lemma 2.1 and Theorem 2.1

PROOF OF LEMMA 2.1. (i) By Chang (1990), IC_j are measurable and bounded. Hence, the second term of (2.5) is continuous in t for any fixed s, because both F_2 and F_3 are bivariate continuous. Therefore, for $j = 1, 2, 3$ and any fixed s, $IC_j(t, s)$ has at most one discontinuity point at $t = s$. Since Q_j and F are continuous, by the Dominated Convergence Theorem and the Assumption B, we know that $\gamma(s, t)$ is bivariate continuous, because for any $t_0 \in [0, T]$,

$$\begin{aligned}\lim_{t\to t_0} \int_0^T IC_j(t, x)dQ_j(x) &= \int_0^T \lim_{t\to t_0} IC_j(t, x)dQ_j(x) \\ &= \int_0^T IC_j^*(t_0, x)dQ_j(x) = \int_0^T IC_j(t_0, x)dQ_j(x)\end{aligned}$$

where $IC_j^*(t_0, x) = IC_j(t_0, x)$ for $x \neq t_0$.

(ii) Consider an operator $\mathcal{L}$: $L^2[0, \beta] \to L^2[0, \beta]$ given by

$$\mathcal{L}(G) = \int_0^\beta \gamma(s, t)G(t)dt, \qquad G \in L^2[0, \beta], \tag{3.1}$$

and denote $\Sigma(\mathcal{L})$ as the set of all eigenvalues of $\mathcal{L}$. Note that $\gamma(s, t)$ is symmetric. Hence, by Kato ((1980), pp. 157, 257 and 269), $\mathcal{L}$ is a compact and selfadjoint linear operator. From Theorem 2.26 of Kato ((1980), p. 185), $\Sigma(\mathcal{L})$ is a countable set with no accumulation point different from zero (i.e., all eigenvalues are isolated), all eigenvalues in $\Sigma(\mathcal{L})$ are nonzero and have finite multiplicities.

Since $\gamma(s, t)$ is a real covariance function, it is nonnegative. Therefore, all λ_j are positive real numbers. □

PROOF OF THEOREM 2.1. (i) Under H_0, we have

$$n\tau(\hat{U}_n) = n \int_0^\beta [\hat{U}_n(t) - t]^2 dt.$$

Let

$$\begin{aligned} T_n(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(t) \\ &= -\frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j=1}^3 \int_0^T IC_j(F^{-1}(t), s)d[I\{W_i \le s, \delta_i = j\} - Q_j(s)]. \end{aligned} \tag{3.2}$$

Since

$$\begin{aligned} &\int_0^\beta \eta_i(t)dt \\ &= -\sum_{j=1}^3 \int_0^\beta \left\{ \int_0^T IC_j(F^{-1}(t), s)d[I\{W_i \le s, \delta_i = j\} - Q_j(s)] \right\} dt \end{aligned} \tag{3.3}$$

are i.i.d.r.v.'s, by the Central Limit Theorem, we have

$$\int_0^\beta T_n(t)dt = \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\beta \eta_i(t)dt = O_p^{(n)}(1), \tag{3.4}$$

where $O_p^{(n)}(1)$ is bounded in probability as $n \to \infty$. Hence, by (2.10), we have

$$n\tau(\hat{U}_n) = \int_0^\beta T_n^2(t)dt + o_p^{(n)}(1). \tag{3.5}$$

Note that $E\{T_n(t)\} = 0$ and the covariance of $T_n(t)$ is given by

$$\begin{aligned} E\{T_n(s)T_n(t)\} &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n E\{\eta_i(s)\eta_j(t)\} \\ &= \frac{1}{n}\sum_{i=1}^n E\{\eta_i(s)\eta_i(t)\} = E\{\eta_i(s)\eta_i(t)\} = \gamma(s,t). \end{aligned} \tag{3.6}$$

By Mercer's Theorem and 'Proper Orthogonal Decomposition Theorem' (Loéve (1963), p. 478), we know that

$$\gamma(s,t) = \sum_{j=1}^\infty \lambda_j \phi_j(s)\phi_j(t), \tag{3.7}$$

$$T_n(t) = \sum_{j=1}^\infty \sqrt{\lambda_j}\phi_j(t)\xi_{nj} \tag{3.8}$$

where ϕ_j are the orthonormal eigenfunctions (i.e., $\int_0^\beta \phi_i(t)\phi_j(t)dt = \delta_{ij}$, for $\delta_{ij} = 1,\ i = j;\ 0,\ i \neq j$) corresponding to the eigenvalues λ_j, and ξ_{nj} are given by

$$\xi_{nj} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{1}{\sqrt{\lambda_j}}\int_0^\beta \eta_i(t)\phi_j(t)dt. \tag{3.9}$$

Note that we have $E\{\xi_{nj}\} = 0$ and $E\{\xi_{ni}\xi_{nj}\} = \delta_{ij}$. By a similar proof of Theorem 1 of Rosenblatt (1952), we have that

$$\int_0^\beta T_n^2(t)dt \xrightarrow{\mathcal{D}} \int_0^\beta T^2(t)dt = \sum_{j=1}^\infty \lambda_j Z_j^2, \qquad \text{as} \quad n \to \infty, \tag{3.10}$$

where $T(t)$ is a Gaussian process for $t \in [0,\beta]$ with mean 0 and covariance $\gamma(s,t)$, and Z_j are independent normal random variables with mean 0 and variance 1. Therefore, by (3.5) and (3.10), we have

$$n\tau(\hat{U}_n) \xrightarrow{\mathcal{D}} \int_0^\beta T^2(t)dt = \sum_{j=1}^\infty \lambda_j Z_j^2, \qquad \text{as} \quad n \to \infty.$$

(ii) Assume that H_0 does not hold. We will derive the limiting distribution of $\tau(\hat{U}_n)$ through the Hadamard differentiability approach. References on this approach can be found in Ren and Sen (1991), Gill (1989) or Fernholz (1983).

We observe that the functional τ given by (1.3) can be expressed as a composition of the following Hadamard differentiable transformations:

$\tau_1 : D[0,\beta] \to D[0,\beta]$, defined by $\tau_1(G) = (G - U_0)^2$, is Hadamard differentiable at U with derivative $\tau'_{1_U}(G) = 2(U - U_0)G$, for $G \in D[0,\beta]$. The proof follows easily from the definition of Hadamard differentiability (Fernholz (1983), Ren and Sen (1991)).

$\tau_2 : D[0,\beta] \to R$, defined by $\tau_2(G) = \int_0^\beta G(t)dU_0(t)$, is Fréchet differentiable at any $S \in D[0,\beta]$ with derivative $\tau'_{2_S}(G) = \int_0^\beta G(t)dU_0(t)$, because it is a linear and continuous functional.

We have $\tau(G) = \tau_2(\tau_1(G))$, and by Chain Rule (Fernholz (1983)), τ is Hadamard differentiable at U with derivative

$$\tau'_U(G) = 2\int_0^\beta [t - U_0(t)]G(t)dU_0(t), \qquad G \in D[0,\beta]. \tag{3.11}$$

From Gu and Zhang (1993), we know that $\sqrt{n}[\hat{U}_n - U]$ weakly converges in the space $(D[0,1], \|\cdot\|)$ to a Gaussian process with continuous covariate $\gamma(s,t)$ given by (2.14). Hence, by Gill (1989), we have

$$\sqrt{n}[\tau(\hat{U}_n) - \tau(U)] = \sqrt{n}\tau'_U(\hat{U}_n - U) + o_p^{(n)}(1).$$

Therefore, by (2.10) of Proposition 2.1, we have

$$\sqrt{n}[\tau(\hat{U}_n) - \tau(U)] = \frac{1}{\sqrt{n}}\sum_{i=1}^n \tau'_U(\eta_i) + o_p^{(n)}(1). \tag{3.12}$$

Since,

$$\begin{aligned}
\tau'_U(\eta_i) &= 2\int_0^\beta [t - U_0(t)]\eta_i(t)dU_0(t) \\
&= 2\int_0^\beta [t - U_0(t)]\xi_i(F^{-1}(t))dU_0(t) \\
&= -2\int_0^\beta [t - U_0(t)] \\
&\qquad \times \left\{\sum_{j=1}^3 \int_0^T IC_j(F^{-1}(t),s)d[I\{W_i \le s, \delta_i = j\} - Q_j(s)]\right\} dU_0(t) \\
&= -2\int_0^\beta [t - U_0(t)]\left\{\sum_{j=1}^3\left(IC_j(F^{-1}(t),W_i)I\{W_i \le T, \delta_i = j\}\right.\right. \\
&\qquad\qquad \left.\left. - \int_0^T IC_j(F^{-1}(t),s)dQ_j(s)\right)\right\}dU_0(t) \\
&= -2\int_0^T [F(x) - F_0(x)]\left\{\sum_{j=1}^3\left(IC_j(x,W_i)I\{W_i \le T, \delta_i = j\}\right.\right. \\
&\qquad\qquad \left.\left. - \int_0^T IC_j(x,s)dQ_j(s)\right)\right\}dF_0(x),
\end{aligned}$$

it is easy to see that $E\{\tau'_U(\eta_i)\} = 0$ and we can compute

$$\sigma_\tau^2 = \mathrm{Var}\{\tau'_U(\eta_i)\} = \sum_{j=1}^{3} \int_0^T \left\{ 2\int_0^T [F(x) - F_0(x)] \times IC_j(x,y)dF_0(x) \right\}^2 dQ_j(y) - \left\{ 2\sum_{j=1}^{3} \int_0^T \int_0^T [F(x) - F_0(x)] \times IC_j(x,y)dQ_j(y)dF_0(x) \right\}^2.$$

Therefore, by the Central Limit Theorem and Slutsky's Lemma, we have that

$$\sqrt{n}[\tau(\hat{U}_n) - \tau(U)] \xrightarrow{\mathcal{D}} N(0, \sigma_\tau^2), \qquad \text{as} \quad n \to \infty$$

follows from (3.12). □

Remark. Note that in (3.11), we have $\tau'_U(G) \equiv 0$ under H_0. This is why under H_0, the limiting distribution of $\tau(\hat{U}_n)$ degenerates to the one given by (2.12).

4. Proof of Theorem 2.2

Before proving Theorem 2.2, we first need to establish some preliminary results. Consider the operators $\mathcal{T}$ and $\mathcal{T}_n : D[0,T] \to D[0,T]$, defined by

$$\mathcal{T}(G) = G(0)F_3(t,0)[1 - S_Z(0)] + \int_0^T G(u)F_3(t,u)h(u)du - \int_0^T G(u)F_2(t,u)g(u)du, \tag{4.1}$$

and

$$\mathcal{T}_n(G) = G(0)F_3^{(n)}(t,0)[1 - S_Z^{(n)}(0)] + \int_0^T G(u)F_3^{(n)}(t,u)h_n(u)du - \int_0^T G(u)F_2^{(n)}(t,u)g_n(u)du, \tag{4.2}$$

where $G \in D[0,T]$, and use $\|\cdot\|$ to denote the norm of the operators when the meaning is clear in the context. Clearly, we have that IC_j and $IC_j^{(n)}$ are the solutions of the following equations:

$$IC_j(\cdot, s) = F_j(\cdot, s) + \mathcal{T}(IC_j(\cdot, s)), \tag{4.3}$$

$$IC_j^{(n)}(\cdot, s) = F_j^{(n)}(\cdot, s) + \mathcal{T}_n(IC_j^{(n)}(\cdot, s)), \tag{4.4}$$

where $s \in [0,T]$, respectively. Moreover, we have the following lemma.

LEMMA 4.1. *Under Assumptions* A, B *and* C, *we have*:

(i) $\mathcal{T}$ *and* $\mathcal{T}_n$ *are bounded and compact operators for* $n \geq 1$;

(ii) $(I-\mathcal{T})^{-1}$ *exists on* $D[0,T]$, *and* $IC_j(\cdot,s) = (I-\mathcal{T})^{-1}(F_j(\cdot,s))$, $j=1,2,3$, *for* $s \in [0,T]$;

(iii) *with probability* 1, *for* $j=1,2,3$,

$$\|F_j^{(n)} - F_j\| \to 0, \qquad as \quad n \to \infty,$$

and

$$\|\mathcal{T}_n - \mathcal{T}\| \to 0, \qquad as \quad n \to \infty;$$

(iv) *with probability* 1, $(I-\mathcal{T}_n)^{-1}$ *exists on* $D[0,T]$ *for sufficiently large* n *and*

$$IC_j^{(n)}(\cdot,s) = (I-\mathcal{T}_n)^{-1}(F_j^{(n)}(\cdot,s)), \qquad j=1,2,3, \qquad for \quad s \in [0,T];$$

(v) *with probability* 1,

$$\|IC_j^{(n)} - IC_j\| \to 0, \qquad as \quad n \to \infty,$$

where $j=1,2,3$.

PROOF. (i) Clearly, $\mathcal{T}$ and $\mathcal{T}_n$ are bounded on $D[0,T]$. Note that F_2 and F_3 are bivariate continuous, and that for any t_1 and t_2,

$$\sup_{s\in[0,T]} |F_j^{(n)}(t_1,s) - F_j^{(n)}(t_2,s)|, \qquad j=2,3 \tag{4.5}$$

can be made arbitrarily small by taking $|t_1-t_2|$ small. Hence, $\mathcal{T}(G)$ and $\mathcal{T}_n(G)$ are the elements of the space of $(C[0,T], \|\cdot\|)$, where $G \in D[0,T]$ and $C[0,T]$ is the space of all continuous real valued functions on $[0,T]$. Moreover, by Kato ((1980), p. 157), we have that $\mathcal{T}$ and $\mathcal{T}_n$ are compact.

(ii) From (13) of Chang (1990), we have that for

$$\begin{aligned} k_{12}(t,s) &= \frac{I\{0<s<t\}}{S_Y(s)-S_Z(s)}, \\ k_{21}(t,s) &= \frac{I\{0<s<t\}}{S_X(s)}, \\ k_{31}(t,s) &= \frac{I\{t<s<T\}}{1-S_X(s)}, \end{aligned}$$

and some $u_1,u_2,u_3 \in D[0,T]$, the following equations:

$$\begin{aligned} \int_0^T k_{12}(t,s)[u_3(s)-u_2(s)]dS_X(s) &= u_1(t), \\ -\int_0^T k_{21}(t,s)u_1(s)dS_Y(s) &= u_2(t), \\ -\int_0^T k_{31}(t,s)u_1(s)dS_Z(s) &= u_3(t), \end{aligned} \tag{4.6}$$

imply that $u_1 = u_2 = u_3 = 0$. Since we can express $\mathcal{T}$ as the following:

$$\mathcal{T}(G) = \int_0^T k_{12}(\cdot, u)\left\{ \int_0^T k_{21}(u,v)G(v)dS_Y(v) - \int_0^T k_{31}(u,v)G(v)dS_Z(v)\right\}dS_X(u), \tag{4.7}$$

then $\mathcal{T}(G) = G$ implies (4.6) with

$$u_1 = G, \qquad u_2 = -\int_0^T k_{21}(\cdot, v)G(v)dS_Y(v),$$
$$u_3 = -\int_0^T k_{31}(\cdot, v)G(v)dS_Z(v).$$

Hence, we have $G = 0$. Therefore, we know that 1 is not an eigenvalue of $\mathcal{T}$. Since $\mathcal{T}$ is compact, by Theorem 6.26 of Kato ((1980), p. 185), we know that $(I - \mathcal{T})^{-1}$ exists and is defined on $D[0,T]$. Hence, we have $IC_j(\cdot, s) = (I - \mathcal{T})^{-1}(F_j(\cdot, s))$, because $F_j(\cdot, s)$ is an element of $D[0,T]$ for any $s \in [0,T]$.

(iii) Note that

$$\int_0^t \frac{dS_X^{(n)}}{S_Y^{(n)} - S_Z^{(n)}} - \int_0^t \frac{dS_X}{S_Y - S_Z} = \int_0^t \left\{ \frac{1}{S_Y^{(n)} - S_Z^{(n)}} - \frac{1}{S_Y - S_Z}\right\} dS_X^{(n)} + \int_0^t \frac{1}{S_Y - S_Z} d[S_X^{(n)} - S_X\}.$$

Since $[S_Y - S_Z]^{-1}$ is continuous and bounded on $[0,T]$, by Theorem 4.2 of Chang and Yang (1987), we know that with probability 1,

$$\left| \int_0^t \frac{dS_X^{(n)}}{S_Y^{(n)} - S_Z^{(n)}} - \int_0^t \frac{dS_X}{S_Y - S_Z}\right| \to 0, \qquad \text{as} \quad n \to \infty,$$

uniformly for $t \in [0,T]$. Hence, we have that with probability 1,

$$\|F_j^{(n)} - F_j\| \to 0, \qquad \text{as} \quad n \to \infty, \tag{4.8}$$

where $j = 1, 2, 3$.

Furthermore, since

$$\|\mathcal{T}_n - \mathcal{T}\| = \sup_{G \in D[0,T], \|G\|=1} \|\mathcal{T}_n(G) - \mathcal{T}(G)\|$$
$$\leq \|F_3^{(n)}(\cdot, 0)[1 - S_Z^{(n)}(0)] - F_3(\cdot, 0)[1 - S_Z(0)]\|$$
$$+ \sup_{G \in D[0,T], \|G\|=1} \left\| \int_0^T G(u)F_2^{(n)}(\cdot, u)g_n(u)du \right.$$

$$
\begin{aligned}
&\qquad\qquad - \int_0^T G(u)F_2(\cdot,u)g(u)du \Bigg\| \\
&+ \sup_{G\in D[0,T],\|G\|=1} \Bigg\| \int_0^T G(u)F_3^{(n)}(\cdot,u)h_n(u)du \\
&\qquad\qquad - \int_0^T G(u)F_3(\cdot,u)h(u)du \Bigg\| \\
\le\ & \|F_3^{(n)}(\cdot,0)[1-S_Z^{(n)}(0)] - F_3(\cdot,0)[1-S_Z(0)]\| \\
&+ \int_0^T \|F_2^{(n)}(\cdot,u)g_n(u) - F_2(\cdot,u)g(u)\|du \\
&+ \int_0^T \|F_3^{(n)}(\cdot,u)h_n(u) - F_3(\cdot,u)h(u)\|du,
\end{aligned}
$$

by Assumption C, (4.8) and the strong consistency of $S_Z^{(n)}$ (Chang and Yang (1987)), we have that with probability 1,

$$\|\mathcal{T}_n - \mathcal{T}\| \to 0, \quad \text{as} \quad n\to\infty. \tag{4.9}$$

(iv) Since $\mathcal{T}$ and $\mathcal{T}_n$ are bounded, by (4.9) and Theorem 2.23 of Kato ((1980), p. 206), we have that with probability 1, $\mathcal{T}_n \to \mathcal{T}$ in generalized sense. Since $(I-\mathcal{T})^{-1}$ exists on $D[0,T]$, from Theorem 2.23 of Kato ((1980), p. 206), we know that with probability 1, $(I-\mathcal{T}_n)^{-1}$ exists on $D[0,T]$ for sufficiently large n, and

$$\|(I-\mathcal{T}_n)^{-1} - (I-\mathcal{T})^{-1}\| \to 0, \quad \text{as} \quad n\to\infty, \tag{4.10}$$

with probability 1. Hence, we have $IC_j^{(n)}(\cdot,s) = (I-\mathcal{T}_n)^{-1}(F_j^{(n)}(\cdot,s))$ for sufficiently large n, because $F_j^{(n)}(\cdot,s)$ is an element of $D[0,T]$ for any $s\in[0,T]$.

(v) Since

$$
\begin{aligned}
IC_j(\cdot,s) &= (I-\mathcal{T})^{-1}(F_j(\cdot,s)),\\
IC_j^{(n)}(\cdot,s) &= (I-\mathcal{T}_n)^{-1}(F_j^{(n)}(\cdot,s))
\end{aligned}
$$

and

$$
\begin{aligned}
\|IC_j^{(n)}(\cdot,s) - IC_j(\cdot,s)\| &\le \|(I-\mathcal{T}_n)^{-1} - (I-\mathcal{T})^{-1}\|\|F_j^{(n)}(\cdot,s)\| \\
&\quad + \|(I-\mathcal{T})^{-1}\|\|F_j^{(n)}(\cdot,s) - F_j(\cdot,s)\|,
\end{aligned}
$$

by (4.8) and (4.10), we have that

$$\sup_{s\in[0,T]} \|IC_j^{(n)}(\cdot,s) - IC_j(\cdot,s)\| \to 0, \quad \text{as} \quad n\to\infty, \tag{4.11}$$

with probability 1. □

PROOF OF THEOREM 2.2. (i) Since by (4.5), we know that $F_2^{(n)}$ and $F_3^{(n)}$ are continuous in t for any fixed s, from (v) of Lemma 4.1 and the Dominated Convergence Theorem, we know that $\mathcal{T}_n(IC_j^{(n)}(\cdot, s))$ is continuous for any fixed s and $j = 1, 2, 3$. Since $F_1^{(n)}(t, s)$ is continuous in t for any fixed s, except at $t = s$, by (4.4), we have that $IC_j^{(n)}(t, s)$ are continuous in t for any fixed s, except at most at point $t = s$. Since $\hat{Q}_j^{(n)}$ are continuous, using a similar argument to the one in the proof of (i) of Lemma 2.1, we know that by the Dominated Convergence Theorem, $\hat{\gamma}_n$ is bivariate continuous.

Let Δ be a r.v. with p.d.f.: $P\{\Delta = j\} = \frac{1}{n}\sum_{i=1}^n I\{\delta_i = j\}$, $j = 1, 2, 3$. Since $\hat{Q}_j^{(n)}/P\{\Delta = j\}$ are d.f.'s, there exist r.v.'s ζ_j such that

$$P\{\zeta_j \leq x \mid \Delta = j\} = \hat{Q}_j^{(n)}(x)/P\{\Delta = j\}.$$

Therefore, we have $P\{\zeta_j \leq x, \Delta = j\} = \hat{Q}_j^{(n)}(x)$. Let

$$\zeta = \zeta_1 I\{\Delta = 1\} + \zeta_2 I\{\Delta = 2\} + \zeta_3 I\{\Delta = 3\}$$

and

$$X(t) = \sum_{j=1}^3 IC_j^{(n)}(t, \zeta) I\{\Delta = j\},$$

then

$$E\{X(t)\} = \sum_{j=1}^3 \int_0^T IC_j^{(n)}(t, x) d\hat{Q}_j^{(n)}(x),$$

and

$$\begin{aligned} E\{X(t)X(s)\} &= E\left\{\sum_{i=1}^3 \sum_{j=1}^3 IC_i^{(n)}(t, \zeta) I\{\Delta = i\} IC_j^{(n)}(s, \zeta) I\{\Delta = j\}\right\} \\ &= \sum_{j=1}^3 E\{IC_j^{(n)}(t, \zeta) IC_j^{(n)}(s, \zeta) I\{\Delta = j\}\} \\ &= \sum_{j=1}^3 \int_0^T IC_j^{(n)}(t, x) IC_j^{(n)}(s, x) d\hat{Q}_j^{(n)}(x). \end{aligned}$$

Hence, by (2.19), we have that $\hat{\gamma}_n(s, t) = \operatorname{cov}\{X(F^{-1}(s)), X(F^{-1}(t))\}$ is a covariance function. Therefore, $\hat{r}_n$ is nonnegative.

(ii) The proof is similar to the one of (ii) of Lemma 2.1.

(iii) It can be shown that with probability 1,

$$[Q_j^{(n)} - Q_j] \to 0, \qquad \text{as} \quad n \to \infty \tag{4.12}$$

uniformly for $t \in [0, T]$. Hence, from (2.20), we have that with probability 1,

$$[\hat{Q}_j^{(n)} - Q_j] \to 0, \qquad \text{as} \quad n \to \infty \tag{4.13}$$

iformly for $t \in [0, T]$. Note that for $(s,t) \in [0,\beta]^2$,

$$
\begin{aligned}
|\hat{\gamma}_n(s,t) - \gamma(s,t)| \le \sum_{j=1}^{3} \bigg| & \int_0^T IC_j^{(n)}(F^{-1}(s),x) IC_j^{(n)}(F^{-1}(t),x) d\hat{Q}_j^{(n)}(x) \\
& - \int_0^T IC_j(F^{-1}(s),x) IC_j(F^{-1}(t),x) dQ_j(x) \bigg| \\
& + \Bigg| \left\{ \sum_{j=1}^{3} \int_0^T IC_j^{(n)}(F^{-1}(s),x) d\hat{Q}_j^{(n)}(x) \right\} \\
& \times \left\{ \sum_{j=1}^{3} \int_0^T IC_j^{(n)}(F^{-1}(t),y) d\hat{Q}_j^{(n)}(y) \right\} \\
& - \left\{ \sum_{j=1}^{3} \int_0^T IC_j(F^{-1}(s),x) dQ_j(x) \right\} \\
& \times \left\{ \sum_{j=1}^{3} \int_0^T IC_j(F^{-1}(t),y) dQ_j(y) \right\} \Bigg| \\
\le & \sum_{j=1}^{3} \bigg| \int_0^T \{ IC_j^{(n)}(F^{-1}(s),x) IC_j^{(n)}(F^{-1}(t),x) \\
& - IC_j(F^{-1}(s),x) IC_j(F^{-1}(t),x) \} d\hat{Q}_j^{(n)}(x) \bigg| \\
& + \sum_{j=1}^{3} \bigg| \int_0^T IC_j(F^{-1}(s),x) \\
& \times IC_j(F^{-1}(t),x) d[\hat{Q}_j^{(n)}(x) - Q_j(x)] \bigg| \\
& + \left\{ \sum_{j=1}^{3} \bigg| \int_0^T IC_j^{(n)}(F^{-1}(s),x) d\hat{Q}_j^{(n)}(x) \right. \\
& \left. - \int_0^T IC_j(F^{-1}(s),x) dQ_j(x) \bigg| \right\} \\
& \times \left| \sum_{j=1}^{3} \int_0^T IC_j^{(n)}(F^{-1}(t),y) d\hat{Q}_j^{(n)}(y) \right| \\
& + \left\{ \sum_{j=1}^{3} \bigg| \int_0^T IC_j^{(n)}(F^{-1}(t),y) d\hat{Q}_j^{(n)}(y) \right.
\end{aligned}
$$

$$- \int_0^T IC_j(F^{-1}(t), y) dQ_j(y) \Bigg| \Bigg\}$$

$$\times \left| \sum_{j=1}^{3} \int_0^T IC_j(F^{-1}(s), x) dQ_j(x) \right|.$$

From Chang (1990), we know that $IC_j(t, s)$ is continuous in s for any fixed t and $j = 1, 2, 3$, except at most at $s = t$. Since F_2 and F_3 are bivariate continuous, we have

$$\int_0^T IC_j(u, s) F_2(t, u) dS_Y(u) - \int_0^T IC_j(u, s) F_3(t, u) dS_Z(u)$$

is bivariate continuous on $[0, T]^2$. Since $F_j(t, s)$ is monotone in t for $j = 1, 2, 3$ and any fixed s, by (2.15), (4.11) and (4.13), we have that

$$\|\hat{\gamma}_n - \gamma\| \to 0, \qquad \text{as} \quad n \to \infty,$$

with probability 1.

(iv) Consider operators $\mathcal{L}_n : L^2[0, \beta] \to L^2[0, \beta]$ given by

$$\mathcal{L}_n(G) = \int_0^\beta \hat{\gamma}_n(s, t) G(t) dt, \qquad G \in L^2[0, \beta],$$

and denote $\Sigma(\mathcal{L}_n)$ as the set of all eigenvalues of $\mathcal{L}_n$. Let

$$\lambda_1 \geq \lambda_2 \geq \cdots, \tag{4.14}$$

$$\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \cdots. \tag{4.15}$$

By (3.7), we observe that

$$\int_0^\beta \gamma(t, t) dt = \sum_{j=1}^{\infty} \lambda_j, \tag{4.16}$$

$$\int_0^\beta \int_0^\beta \gamma^2(s, t) ds dt = \sum_{j=1}^{\infty} \lambda_j^2, \tag{4.17}$$

and that similarly for $\hat{\gamma}_n$, we have

$$\int_0^\beta \hat{\gamma}_n(t, t) dt = \sum_{j=1}^{\infty} \lambda_j^{(n)}, \tag{4.18}$$

$$\int_0^\beta \int_0^\beta \hat{\gamma}_n^2(s, t) ds dt = \sum_{j=1}^{\infty} (\lambda_j^{(n)})^2. \tag{4.19}$$

Since by (2.22), we have that with probability 1,

$$\int_0^\beta \hat{\gamma}_n(t, t) dt = \sum_{j=1}^{\infty} \lambda_j^{(n)} \to \int_0^\beta \gamma(t, t) dt = \sum_{j=1}^{\infty} \lambda_j, \qquad \text{as} \quad n \to \infty, \tag{4.20}$$

we have that with probability 1, there exists $M > 0$ such that

$$0 < \lambda_j^{(n)} < M, \quad \text{for} \quad j \geq 1 \text{ and } n \geq 1. \tag{4.21}$$

Let $\lambda_j^{(n_k)}$ be any convergent subsequence of $\lambda_j^{(n)}$ such that for $j = 1, 2, \ldots$

$$\lambda_j^{(n_k)} \to \gamma_j \in [0, M], \quad \text{as} \quad k \to \infty, \tag{4.22}$$

and let $\Sigma = \{\gamma_j; \gamma_j > 0, j = 1, 2, 3, \ldots\}$, then by (4.15), we have

$$\gamma_1 \geq \gamma_2 \geq \cdots. \tag{4.23}$$

For any $\gamma_j \in \Sigma$, let $\phi_j^{(n)}$ be the orthonormal eigenfunctions, i.e., $\int_0^\beta \phi_i^{(n)}(t)\phi_j^{(n)}(t)dt = \delta_{ij}$, corresponding to the eigenvalues $\lambda_j^{(n)}$ for the eigenvalue problem (2.21). Since $\mathcal{L}$ is compact and $\|\phi_j^{(n)}\| = 1$ and since

$$\begin{aligned}\int_0^\beta \hat{\gamma}_n(s,t)\phi_j^{(n)}(t)dt &= \int_0^\beta [\hat{\gamma}_n(s,t) - \gamma(s,t)]\phi_j^{(n)}(t)dt + \int_0^\beta \gamma(s,t)\phi_j^{(n)}(t)dt \\ &= \lambda_j^{(n)}\phi_j^{(n)}(s),\end{aligned}$$

by (iii) of Theorem 2.2, there exists convergent subsequence, also denoted by $\phi_j^{(n_k)}$, such that

$$\phi_j^{(n_k)} \to \psi_j \neq 0, \quad \text{as} \quad k \to \infty. \tag{4.24}$$

Hence, we have

$$\int_0^\beta \gamma(s,t)\psi_j(t)dt = \gamma_j\psi_j(s),$$

which implies that γ_j is an eigenvalue of $\mathcal{L}$. Therefore, we have $\Sigma \subset \Sigma(\mathcal{L})$.

Now, we show that $\lambda_1 \in \Sigma$ and $\lambda_1 = \gamma_1 > 0$. First, note that

$$\begin{aligned}&\|\mathcal{L}_n(G) - \mathcal{L}(G)\| \leq \|\mathcal{L}_n - \mathcal{L}\|\|G\|, \\ &\|\mathcal{L}_n - \mathcal{L}\| \leq \beta\|\hat{\gamma}_n - \gamma\|,\end{aligned}$$

and that T is selfadjoint and only has isolated eigenvalues λ with finite multiplicity m_λ and isolation distance $d_\lambda > 0$. From the discussion by Kato ((1980), p. 290), we know that when $\|\mathcal{L}_n - \mathcal{L}\| < \frac{1}{2}d_\lambda$, in the circle of $\Gamma_\lambda = \{u; |u - \lambda| \leq \frac{1}{2}d_\lambda\}$, there are exactly m_λ repeated eigenvalues of $\mathcal{L}_n$ and no other points of $\Sigma(\mathcal{L}_n)$. Hence, by (iii) of Theorem 2.2, when n is sufficiently large, in Γ_λ there are exactly $m_1 = m_{\lambda_1}$ repeated eigenvalues of $\mathcal{L}_n$ and no other points of $\Sigma(\mathcal{L}_n)$. Since $\mathcal{L}$ and $\mathcal{L}_n$ are selfadjoint, by Theorem 4.10 of Kato ((1980), p. 291) and (iii) of Theorem 2.2, we know that for any $j \geq 1$,

$$\operatorname{dist}\{\lambda_j^{(n)}, \Sigma(\mathcal{L})\} \leq \|\mathcal{L}_n - \mathcal{L}\| \leq \beta\|\hat{\gamma}_n - \gamma\| \to 0, \quad \text{as} \quad n \to \infty, \tag{4.25}$$

with probability 1. If for any large n, there exists $\kappa_n \in \Sigma(\mathcal{L}_n)$ such that $\kappa_n \geq \lambda_1 + \frac{1}{2}d_{\lambda_1}$, then we have

$$\text{dist}\{\kappa_n, \Sigma(\mathcal{L})\} \geq \frac{1}{2}d_{\lambda_1} > 0,$$

which contradicts (4.25). Hence, by (4.15), we have that for sufficiently large n, those m_1 repeated eigenvalues of $\mathcal{L}_n$ in Γ_{λ_1} are $\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_{m_1}^{(n)}$. Therefore, by (4.22), we have $\lambda_1 = \gamma_1 \in \Sigma$.

Similarly, we can show the same for $\lambda_2, \lambda_3, \ldots$. Hence, we have $\Sigma = \Sigma(\mathcal{L})$ and $\gamma_j = \lambda_j > 0$ for $j = 1, 2, \ldots$. Therefore, we know that every convergent subsequence of $\lambda_j^{(n)}$ converges to λ_j for every $j \geq 1$. This gives that with probability 1,

$$\lambda_j^{(n)} \to \lambda_j, \qquad \text{and} \quad n \to \infty, \tag{4.26}$$

for $j = 1, 2, \ldots$.

Since

$$\int_0^\beta \gamma(t,t)dt = \sum_{j=1}^\infty \lambda_j < \infty,$$

then for any $\epsilon > 0$, there exists $N_\epsilon > 0$ such that $\sum_{j=N_\epsilon}^\infty \lambda_j \leq \epsilon$. Hence, we have that with probability 1,

$$\left| \sum_{j=1}^\infty \lambda_j(\lambda_j^{(n)} - \lambda_j) \right| \leq \sum_{j=1}^{N_\epsilon} \lambda_j |\lambda_j^{(n)} - \lambda_j| + 2M\epsilon. \tag{4.27}$$

By (4.26), we have that with probability 1,

$$\sum_{j=1}^\infty \lambda_j(\lambda_j^{(n)} - \lambda_j) \to 0, \qquad \text{as} \quad n \to \infty. \tag{4.28}$$

Note that by (2.22), we have

$$\int_0^\beta \int_0^\beta \hat{\gamma}_n^2(s,t)dsdt \to \int_0^\beta \int_0^\beta \gamma^2(s,t)dsdt, \qquad \text{as} \quad n \to \infty \tag{4.29}$$

with probability 1. Hence, (4.17), (4.19) and (4.29) imply

$$\sum_{j=1}^\infty (\lambda_j^{(n)})^2 \to \sum_{j=1}^\infty \lambda_j^2, \qquad \text{as} \quad n \to \infty \tag{4.30}$$

with probability 1. Also, note that

$$\begin{aligned}\sum_{j=1}^\infty (\lambda_j^{(n)} - \lambda_j)^2 &= \sum_{j=1}^\infty (\lambda_j^{(n)})^2 - 2\sum_{j=1}^\infty \lambda_j^{(n)}\lambda_j + \sum_{j=1}^\infty \lambda_j^2 \\ &= \sum_{j=1}^\infty (\lambda_j^{(n)})^2 - 2\sum_{j=1}^\infty \lambda_j(\lambda_j^{(n)} - \lambda_j) - \sum_{j=1}^\infty \lambda_j^2.\end{aligned}$$

Hence, by (4.28) and (4.30), we have that with probability 1,

$$\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j)^2 \to 0, \quad \text{as} \quad n \to \infty.$$

(v) Let $W = \sum_{j=1}^{\infty} \lambda_j Z_j^2$, $W_n = \sum_{j=1}^{\infty} \lambda_j^{(n)} Z_j^2$ and $\boldsymbol{\lambda}^{(n)} = (\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots)$, then by (4.20), we have that as $n \to \infty$,

$$(4.31)\quad c_n = E\{W_n - W\} = E\{E\{W_n - W \mid \boldsymbol{\lambda}^{(n)}\}\} = E\left\{\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j)\right\} \to 0.$$

Note that $\boldsymbol{\lambda}^{(n)}$ and $\boldsymbol{Z} = (Z_1, Z_2, \ldots)$ are independent, and that we have

$$\begin{aligned}
(4.32)\quad & E\{[(W_n - W) - E(W_n - W)]^2 \mid \boldsymbol{\lambda}^{(n)}\} \\
&= E\left\{\left(\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j)Z_j^2 - c_n\right)^2 \middle| \boldsymbol{\lambda}^{(n)}\right\} \\
&= E\left\{\left(\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j)(Z_j^2 - 1) + \sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j) - c_n\right)^2 \middle| \boldsymbol{\lambda}^{(n)}\right\} \\
&= E\left\{\left(\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j)(Z_j^2 - 1)\right)^2 \middle| \boldsymbol{\lambda}^{(n)}\right\} \\
&\quad + \left(\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j) - c_n\right)^2 \\
&= 2\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j)^2 + \left(\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j) - c_n\right)^2.
\end{aligned}$$

Since for any $\epsilon > 0$, we have that by (4.32),

$$\begin{aligned}
& P\{|(W_n - W) - E(W_n - W)| \geq \epsilon\} \\
&\quad = E\{P(|(W_n - W) - E(W_n - W)| \geq \epsilon \mid \boldsymbol{\lambda}^{(n)})\} \\
&\quad \leq \frac{1}{\epsilon^2} E\{E([(W_n - W) - E(W_n - W)]^2 \mid \boldsymbol{\lambda}^{(n)})\} \\
&\quad = \frac{1}{\epsilon^2} E\left\{2\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j)^2 + \left(\sum_{j=1}^{\infty}(\lambda_j^{(n)} - \lambda_j) - c_n\right)^2\right\},
\end{aligned}$$

hence by (iv) of Theorem 2.2, (4.20) and (4.31), we have

$$(4.33)\qquad P\{|(W_n - W) - E(W_n - W)| \geq \epsilon\} \to 0, \quad \text{as} \quad n \to \infty.$$

Therefore, by (4.31), we have $W_n \xrightarrow{\mathcal{D}} W$, as $n \to \infty$. Clearly, (2.25) follows easily from (2.24) and (i) of Theorem 2.1. □

Acknowledgements

I would like to thank Professor Pranab K. Sen and Professor Jie Xiong for helpful discussions and suggestions while this manuscript was prepared, and would like to thank Professor Glenn Ledder and Professor Tom Shore, two applied mathematicians, for useful suggestions on the numerical methods for solving the integral equations and the eigenvalue problems. I would also like to thank the referee for useful comments and suggestions on the manuscript.

References

Bickel, P. J. and Ren, J. (1995). The m out of n bootstrap and goodness of fit tests with doubly censored data (preprint).

Burden, R. L. and Faires, J. D. (1989). *Numerical Analysis*, PWS-KENT Publishing Company, Boston, Massachusetts.

Chang, M. N. (1990). Weak convergence of a self-consistent estimator of the survival function with doubly censored data, *Ann. Statist.*, **18**, 391–404.

Chang, M. N. and Yang, G. L. (1987). Strong consistency of a nonparametric estimator of the survival function with doubly censored data, *Ann. Statist.*, **15**, 1536–1547.

Delves, L. M. and Mohamed, J. L. (1985). *Computational Methods for Integral Equations*, Cambridge University Press, Massachusetts.

Delves, L. M. and Walsh, J. (1974). *Numerical Solution of Integral Equations*, Clarendon Press, Oxford.

Fernholz, L. T. (1983). Von Mises calculus for statistical functional, *Lecture Notes in Statist.*, **19**, Springer, New York.

Gehan, E. A. (1965). A generalized two-sample Wilcoxon test for doubly censored data, *Biometrika*, **52**, 650–653.

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

Gu, M. G. and Zhang, C. H. (1993). Asymptotic properties of self-consistent estimators based on doubly censored data, *Ann. Statist.*, **21**(2), 611–624.

Kato, T. (1980). *Perturbation Theory for Linear Operators*, Springer, New York.

Loève, M. (1963). *Probability Theory*, Van Nostrand, Princeton, New Jersey.

Mantel, N. (1967). Ranking procedures for arbitrarily restricted observations, *Biometrics*, **23**, 65–78.

Marron, J. S. and Padgett, W. J. (1987). Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples, *Ann. Statist.*, **15**(4), 1520–1535.

Mielniczuk, J. (1986). Some asymptotic properties of kernel estimators of a density function in case of censored data, *Ann. Statist.*, **14**(2), 766–773.

Peto, R. (1973). Experimental survival curves for interval-censored data, *Appl. Statist.*, **22**, 86–91.

Ren, J. (1994). On self-consistent estimators and kernel density estimators with doubly censored data, Tech. Report, #22, Division of Statistics, Department of Mathematics & Statistics, University of Nebraska-Lincoln (submitted).

Ren, J. and Ledder, G. (1995). On Cramér-von Mises goodness of fit tests for doubly censored data (submitted).

Ren, J. and Sen P. K. (1991). On Hadamard differentiability of extended statistical functional, *J. Multivariate Anal.*, **39**, 30–43.

Ren, J. and Zhou, M. (1993). Extended L-estimators and M-estimators for doubly censored data, Tech. Report, #20, Division of Statistics, Department of Mathematics & Statistics, University of Nebraska-Lincoln (submitted).

Ren, J. and Zhou, M. (1994). Generalized R-statistics for doubly censored data, Tech. Report, #25, Division of Statistics, Department of Mathematics & Statistics, University of Nebraska-Lincoln.

Rosenblatt, M. (1952). Limit theorems associated with variants of the von Mises statistic, *Ann. Math. Statist.*, **23**, 617–623.

Tsai, W. Y. and Crowley, J. (1985). A large sample study of generalized maximum likelihood estimators from incomplete data via self-consistency, *Ann. Statist.*, **13**, 1317–1334.

Turnbull, B. W. (1974). Nonparametric estimation of a survivorship function with doubly censored data, *J. Amer. Statist. Assoc.*, **69**, 169–173.