A bivariate distribution function estimator and its variance under left truncation and right censoring*

Ülkü Gürler
Bilkent University, Turkey
and
Irène Gijbels
Université Catholique de Louvain, Belgium

December, 1996

Abstract

Left truncated and right censored data can naturally arise in survival analysis besides other fields such as insurance and economics. In this paper we consider the situation of bivariate observations where one of the components is subject to left truncation and right censoring. An estimator for the bivariate distribution function is proposed, and its asymptotic properties are established via a strong i.i.d. representation. An explicit expression for the asymptotic covariance function is provided. The asymptotic variance can either be estimated directly via a plug-in procedure or by relying on bootstrap methods. The asymptotic covariance function is quite complicated and in order to investigate the performance of the direct and the bootstrap estimator we focus on the censoring only case.

AMS 1990 subject classifications: 62G05, 62G20, 62G30.

Key words and phrases: Bivariate distribution, bootstrap estimator, censoring, covariance function, truncation.

* This research was supported by NATO Collaborative Research Grant CRG 950271. Part of this research was conducted while the first author was a visiting fellow at the Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve. The second author was supported by 'Projet d'Actions de Recherche Concertées' (No. 93/98 - 164) and by an FNRS-grant (No. 1.5.001.95F).

1 Introduction

In survival or reliability studies, the observed data are often censored and/or truncated. Left truncation and right censoring (LTRC) together occur naturally in cohort follow-up studies. Consider for example the survival of the patients after a cancer surgery. If the follow-up study starts at time t_0, then those individuals who have gone through a surgery and have died before t_0 can not be included in the study, they are left truncated. Furthermore, those who are eligible (alive at t_0) are subject to censoring during the follow-up due to various reasons such as the termination of the study or drop-out. This incomplete nature of the sampling introduces biases and conventional methods do not perform satisfactory, and should be adjusted. The effect of truncation is particularly important if the phenomenon under study is in its early stages and sufficient historical data have not yet been accumulated. Let us first give a more precise description of the observed data. Suppose we are interested in the variable Y. However due to reasons explained above, we can only observe samples of the random vector (T, Z, δ), if $T \leq Z$, where $Z = \min(Y, C)$ and $\delta = I(Y \leq C)$. Here T is the truncating variable and C is the censoring variable which are both assumed to be independent of Y. Note that the LTRC model includes the left truncation ($C \equiv \infty$) and the right censoring ($T \equiv 0$) models as special cases. Here we assume without loss of generality that all random variables are nonnegative. Applications of the LTRC model arise in for example AIDS research, where the survival time is measured from the onset of the disease to the time of death. Then, those patients who died before t_0 (time when the study begins) are clearly not observed. Suppose S denotes the time of the diagnosis, then $T = t_0 - S$. If at time D, a patient drops out from the study, then for this patient $Z = D - S$ and $\delta = 0$. Other examples of LTRC can be found in Andersen *et al.* (1993).

The analysis of bivariate survival data subject to truncation/censoring gained considerable attention recently. For the censored data case Dabrowska (1988) and Prentice and Cai (1992) can be cited among others. Gürler (1996 a, b) suggested nonparametric estimators for the bivariate hazard and the distribution function when a single component is subject to truncation. Recently, van der Laan (1996) introduced efficient estimators when both components are subject to truncation. The proposed estimators are solutions of a set of score equations corresponding to one dimensional submodels. He suggests an iterative algorithm for estimating the limiting variance.

In this paper we suggest a nonparametric estimator for the bivariate distribution function when a component is subject to LTRC. An explicit expression for the covariance function of the limiting distribution is provided, and is further elaborated for special cases. The results are organized as follows. In Section 2, a nonparametric estimator for the bivariate distribution function $F(y, x)$ of the random pair (Y, X), when Y is subject to left truncation and right censoring is derived and its asymptotic properties are studied via an almost sure i.i.d. representation. Some tedious calculations (not presented here) lead to an explicit expression for the asymptotic covariance function, which can be used in inference procedures. The expression for the asymptotic variance-covariance function is quite complicated. For further investigation we focus on the censoring only case. For this special case we studied two estimators for the unknown variance function. The first estimator is a plug-in type nonparametric estimator based on the asymptotic variance. The second estimator uses bootstrap ideas to access the variance of the estimator. The estimation procedures are introduced in Section 3. The performance of the nonparametric estimator for the bivariate distribution function and the two estimators for the variance of the estimator is investigated via extensive simulations in Section 4. Section 5 contains concluding remarks and some discussion.

2 Estimation of the bivariate distribution function

2.1 The estimator: definition and properties

In this section an estimator for the bivariate distribution function $F(y,x)$ of (Y,X) under the LTRC model is introduced. Suppose that the component Y is subject to LTRC and we observe $(Z_i, X_i, T_i, \delta_i)$, $i = 1, \ldots, n$, for which $T_i \leq Z_i$. Let F_Y denote the marginal distribution function of Y, the variable of primary interest, and assume that T and C and the vector (Y,X) are mutually independent. The distribution functions of T and C are denoted by G and H respectively. For any distribution function L we denote $\bar{L}(t) = 1 - L(t)$. Let V_Z denote the distribution function of the random variable $Z = \min(Y,C)$. Then $\bar{V}_Z = (1 - F_Y)(1 - H)$. For any distribution function L, let $a_L = \inf\{t : L(t) > 0\}$ and $b_L = \inf\{t : L(t) = 1\}$ be the left and right endpoints of its support. Then, as in the univariate truncation case (see Woodroofe (1985)), $F(y,x)$ is identifiable only if $a_G \leq a_{V_Z}$ and $b_G \leq b_{V_Z}$, which is assumed to be true in this study. Some further notations are needed. The distribution functions of the *observed* random variables are denoted by W with subscript(s) indicating the particular variable(s) that are involved. For example, let $W^1_{Z,X}(z,x)$ be the bivariate sub-distribution function of the observed uncensored variables, i.e.

$$W^1_{Z,X}(z,x) = P\,(\,Z \leq z, X \leq x, \delta = 1 \mid T \leq Z) = \alpha^{-1} \int_0^\infty \int_0^x \int_0^{z \wedge c} G(u)\; F(du,dv)\; H(dc),$$

where $0 < \alpha = P(T \leq Z)$, $t \wedge u = \min(t,u)$, and $t \vee u = \max(t,u)$. The following marginals will be used in what follows:

$$W^1_{Z,T}(z,t) = P\,(\,Z \leq z, T \leq t, \delta = 1 \mid T \leq Z) = \alpha^{-1} \int_0^z G(t \wedge u)\bar{H}(u-)F_Y(du),$$

$$\begin{aligned} W^1_Z(z) &= \alpha^{-1} \int_0^z G(u)\bar{H}(u-)F_Y(du), \\ W^1_{Z,X}(dz,dx) &= \alpha^{-1}\; G(z)\; \bar{H}(z-)\; F(dz,dx), \\ W^1_Z(dz) &= \alpha^{-1}\; G(z)\; \bar{H}(z-)\; F_Y(dz). \end{aligned} \tag{2.1}$$

The function

$$C(u) = P(T \leq u \leq Z | T \leq Z)$$

plays a crucial role in the truncation model and its estimation. Note that

$$C(u) = \alpha^{-1} G(u) \bar{F}_Y(u-) \bar{H}(u-) = W_T(u) - W_Z(u-), \tag{2.2}$$

where $W_T(t) = P(T \le t | T \le Z)$ and $W_Z(z) = P(Z \le z | T \le Z)$ is the distribution function of the observed r.v. T, respectively Z. From (2.1) and (2.2) it is easily seen that

$$F(dy, dx) = \frac{\bar{F}_Y(y-)}{C(y)} W^1_{Z,X}(dy, dx) \equiv A(y) W^1_{Z,X}(dy, dx), \tag{2.3}$$

with

$$A(u) = \frac{\bar{F}_Y(u-)}{C(u)}.$$

The empirical counterpart of $A(u)$ is defined as

$$A_n(u) = \frac{\bar{F}_{Y,n}(u-)}{C_n(u)},$$

where

$$\bar{F}_{Y,n}(y) = \prod_{i: Z_i \le y} \left[1 - \frac{s(Y_i)}{n C_n(Y_i)} \right]^{\delta_i}$$

with

$$n C_n(u) = \#\{ i : T_i \le u \le Z_i \},$$

and for $u > 0$, $s(u) = \#\{ i : Z_i = u \}$. Expression (2.3) motivates the following estimator for $F(y, x)$:

$$F_n(y, x) = \frac{1}{n} \sum_{i=1}^{n} A_n(Z_i) I(Z_i \le y, X_i \le x, \delta_i = 1). \tag{2.4}$$

It is easy to verify that $F_n(y, x)$ is a bivariate distribution function, and when $x \to \infty$ it reduces to a univariate product-limit type of estimator for $F_Y(y)$ based on left truncated and right censored data.

The following theorem provides a strong i.i.d. representation for the estimator given in (2.4). From this representation, consistency and asymptotic normality are obtained in Corollary 1. The proof of Theorem 1 is deferred to the Appendix.

Define

$$\begin{aligned} L_i(z) &= \frac{I(Z_i \le z, \delta_i = 1)}{C(Z_i)} - \int_0^z \frac{I(T_i \le u \le Z_i)}{C^2(u)} W^1_Z(du), \\ \bar{L}_n(z) &= \frac{1}{n} \sum_{i=1}^{n} L_i(z), \end{aligned} \tag{2.5}$$

and denote by

$$W^1_{Z,X,n}(z, x) = \frac{1}{n} \sum_{i=1}^{n} I(Z_i \le z, X_i \le x, \delta_i = 1),$$

the empirical counterpart of $W^1_{Z,X}(z, x)$.

Theorem 1: Assume $F(y,x)$ is continuous in both components, $b < b_{V_Z}$ and let $T_b = \{(y,x) : 0 < y < b; 0 < x < \infty\}$. Then $F_n(y,x)$ admits the following representation:

$$
\begin{aligned}
F_n(y,x) - F(y,x) &= \int_0^y A(u)[W^1_{Z,X,n}(du,x) - W^1_{Z,X}(du,x)] \\
&\quad - \int_0^y \left\{ \frac{A(u)}{C(u)}[C_n(u) - C(u)] - \bar{L}_n(u)A(u) \right\} W^1_{Z,X}(du,x) + R_n(y,x) \\
&\equiv \bar{\xi}_n(y,x) + R_n(y,x) \qquad (2.6)
\end{aligned}
$$

and

(i) If $a_G < a_{V_Z}$, then

$$\sup_{(y,x)\,\epsilon\, T_b} |R_n(y,x)| = O\left(n^{-1}\log^2 n\right) \qquad \text{a.s.}$$

(ii) If $a_G = a_{V_Z}$, and $\int G^{-3}(u)V_Z(du) < \infty$, then

$$\sup_{(y,x)\,\epsilon\, T_b} |R_n(y,x)| = O\left(n^{-1}\log^3 n\right) = o\left(n^{-1/2}\right) \qquad \text{a.s.}$$

Strong consistency and the law of the iterated logarithm now follow via standard arguments.

Corollary 1: Under the assumptions of Theorem 1,
(a) For $(y,x)\,\epsilon\, T_b$

$$F_n(y,x) \to F(y,x) \qquad \text{a.s.}$$

(b) $\sup_{(y,x)\,\epsilon\, T_b} |F_n(y,x) - F(y,x)| = O\left(n^{-1/2}(\log n)^{1/2}\right) \qquad$ a.s.

2.2 Asymptotic variance-covariance function of the estimator

The covariance structure of the limiting process is quite complicated, particularly due to the effect of truncation. Deriving the covariance function is a tedious task and this derivation can be found in a separate paper by Gijbels and Gürler (1996). The explicit formula for the covariance function of the scaled estimator $\sqrt{n}\,(F_n(y,x) - F(y,x))$ is given in the next theorem. The following notations are used:

$$
\begin{aligned}
a_1(t,x) &= \int_0^t \frac{W^1_{Z,X}(dv,x)}{G(v)} \\
a_2(t,x) &= \int_0^t \frac{G(v)}{C(v)} F(dv,x) \\
b_1(t,x) &= \int_0^t \frac{W^1_{Z,X}(dv,x)}{C(v)} \\
b(t) &= \int_0^t \frac{W^1_Z(dz)}{C^2(z)} \\
h(t) &= \int_0^t \frac{G(v)}{C^2(v)} W^1_Z(dv)
\end{aligned}
$$

$$h(t,x) = \int_0^t \frac{G(v)}{C^2(v)} W^1_{Z,X}(dv,x)$$
$$d(u,v,x) = \int_0^{u\wedge v} [a_1(u,x) - a_1(s,x)]h(ds).$$

Theorem 2: Suppose $\int G^{-1}(u)V_Z(du) < \infty$. Then, the asymptotic covariance function of the scaled version of estimator (2.4) is given by

$$\begin{aligned}\mathrm{Cov}(y_1,y_2,x_1,x_2) &= \int_0^{y_1\wedge y_2} A^2(u)W^1_{Z,X}(du, x_1\wedge x_2)\\ &+\mathcal{T}(y_1,y_2,x_1,x_2) + \mathcal{T}(y_2,y_1,x_2,x_1) + \mathcal{X}(y_1,y_2,x_1,x_2),\end{aligned}$$

with

$$\begin{aligned}\mathcal{T}(y_1,y_2,x_1,x_2) &= -a_2(y_2,x_2)\int_{y_1\wedge y_2}^{y_1} A(u)a_1(du,x_1) - \int_0^{y_1\wedge y_2} A(u)a_2(u,x_2)a_1(du,x_1)\\ &-\int_0^{y_1\wedge y_2}[F(y_2,x_2) - F(u,x_2)]A(u)b_1(du,x_1)\\ &+\int_0^{y_1}\int_0^{y_2}[d(y_1,v,x_1) - d(u,v,x_1)]A(du)F(dv,x_2),\end{aligned}$$

and

$$\mathcal{X}(y_1,y_2,x_1,x_2) = \int_0^{y_1}\int_0^{y_2} K(u,v)F(du,x_1)F(dv,x_2),$$

where

$$K(u,v) = \frac{G(u\wedge v)}{G(u\vee v)C(u\wedge v)} - h(u\wedge v)\frac{G(u)+G(v)}{G(u)G(v)} + b(u\wedge v).$$

The variance function of the scaled version of the bivariate distribution function estimator (2.4) is given by

$$\sigma^2(y,x) \equiv \mathrm{Var}(y,x) = \mathrm{Cov}(y,y,x,x) = \int_0^y A^2(u)W^1_{Z,X}(du,x) + 2\mathcal{T}(y,x) + \mathcal{X}(y,x), \qquad (2.7)$$

where $\mathcal{X}(y,x) = \mathcal{X}(y,y,x,x)$ and $\mathcal{T}(y,x) = \mathcal{T}(y,y,x,x)$.

The result of Theorem 2 leads to the following corollary on the weak convergence of $Q_n(y,x) \equiv \sqrt{n}(F_n(y,x) - F(y,x))$, which follows from the weak convergence of empirical processes in multidimensional time studied by for example Neuhaus (1971). Let (D_2, d) be the complete separable metric space defined on $[0,1]\times[0,1]$ as described in Neuhaus (1971). Then

Corollary 2: Suppose the conditions of part (ii) of Theorem 1 hold. Then for $(y,x) \in T_b$, $Q_n(y,x)$ converges weakly to a mean zero, two dimensional time Gaussian process on (D_2, d) with the covariance structure as given in Theorem 2.

3 A Special Case: Censoring Only

3.1 The asymptotic variance function and a plug-in type estimator

To gain some further insight in the covariance structure of the limiting process of the bivariate distribution function estimator we consider in this section the special and important case of censoring only. The censoring case is obtained by letting $T \equiv 0$ so that $G(u) = 1$ for all u, and $\alpha = 1$. Note also that in this special case

$$\begin{aligned} C(u) &= \bar{F}_Y(u-)\bar{H}(u-) \\ A(u) &= \frac{1}{\bar{H}(u-)} \\ b(u) &= \int_0^u \frac{1}{\bar{F}_Y^2(z-)\bar{H}(z-)} F_Y(dz), \end{aligned}$$

and

$$\int_0^y A^2(u) W^1_{Z,X}(du, x) = \int_0^y \frac{1}{\bar{H}(u-)} F(du, x).$$

It can be shown (see Gijbels and Gürler (1996)) that the variance function given in (2.7) reduces in the special case of censoring only to

$$\sigma^2(y, x) \equiv \mathrm{Var}(y, x) = \int_0^y A^2(u) W^1_{Z,X}(du, x) - 2\int_0^y [F(y, x) - F(v, x)] \left[\frac{1}{C(v)} - b(v)\right] F(dv, x).$$

This variance formula can be used to make some standard inferences, such as hypothesis testing and construction of confidence intervals. For such applications an estimator of the variance is required. We first rewrite the expression of the variance function. Since $A(z)W^1_{Z,X}(dz, x) = F(dz, x)$ in general and $A(z)W^1_{Z,X}(dz, x) = [\bar{H}(z-)]^{-1} W^1_{Z,X}(dz, x) = F(dz, x)$ the above expression can be rewritten as

$$\sigma^2(y, x) = \int_0^y A^2(u) W^1_{Z,X}(du, x) - 2\int_0^y [F(y, x) - F(v, x)] \left[\frac{1}{C(v)} - b(v)\right] A(v) W^1_{Z,X}(dv, x). \tag{3.1}$$

From expression (3.1) it now becomes transparent how to estimate the asymptotic variance function. Let

$$\begin{aligned} A_n(y, x) &= \frac{1}{n}\sum_{i=1}^n A_n^2(Z_i) I(Z_i \le y, X_i \le x, \delta_i = 1), \\ B_n(y, x) &= 2\frac{1}{n}\sum_{i=1}^n A_n(Z_i)\{F_n(y, x) - F_n(Z_i, x)\}\left\{\frac{1}{C_n(Z_i)} - b_n(Z_i)\right\} I(Z_i \le y, X_i \le x, \delta_i = 1), \end{aligned}$$

where

$$b_n(u) = \frac{1}{n}\sum_{i=1}^n \frac{I(Z_i \le u, \delta_i = 1)}{C_n^2(Z_i)}.$$

Then, an estimator of the asymptotic variance is given by:

$$\sigma_n^2(y, x) = A_n(y, x) - B_n(y, x). \tag{3.2}$$

3.2 A bootstrap estimator for the variance of the estimator

In the previous section we discussed estimation of the asymptotic variance of the bivariate distribution function estimator (2.4). An alternative to this asymptotic approach is to use bootstrap methods to access the unknown variance of the estimator. Bootstrap methods for truncated and censored data are discussed recently in Gross and Lai (1996). In that paper the authors show that the "simple" and the "obvious" methods suggested by Efron (1981) for censored data are no longer equivalent as was the case in censoring. They provide an asymptotic theory for the "simple" bootstrap method, which we also adopted in this paper. This bootstrap procedure reads as follows. Given the sample $(Z_i, X_i, T_i, \delta_i)$, $i = 1, \ldots, n$ we generate B bootstrap samples $(Z_i^{*b}, X_i^{*b}, T_i^{*b}, \delta_i^{*b})$, $i = 1, \ldots, n$, $b = 1, \cdots, B$. For each of the bootstrap samples we obtain

$$F_n^{*b}(y,x) = \frac{1}{n}\sum_{i=1}^{n} A_n(Z_i^{*b}) I(Z_i^{*b} \le y, X_i^{*b} \le x, \delta_i^{*b} = 1),$$

the bootstrap version of the estimator. An estimator of the variance of the estimator $F_n(y,x)$ is given by the sample variance along the bootstrap samples, i.e.

$$\hat{\sigma}_B^2(y,x) = \frac{1}{B-1}\sum_{b=1}^{B}\left\{F_n^{*b}(y,x) - \overline{F_n^{*b}}(y,x)\right\}^2, \tag{3.3}$$

where

$$\overline{F_n^{*b}}(y,x) = \frac{1}{B}\sum_{b=1}^{B} F_n^{*b}(y,x).$$

4 Simulation study

4.1 The simulation models

We consider two simulation models. In the first model we take

$$\begin{aligned} X &\sim \text{Exp}(1) \qquad Y \sim \text{Exp}(1), \qquad X \text{ and } Y \text{ independent;} \\ C &\sim \text{Exp}(\lambda), \end{aligned} \tag{4.1}$$

where the parameter λ determines the censoring proportion $P(Y > C) = \lambda/(\lambda+1)$. This model will be referred to as IND. EXP. in the graphs and the tables. It is easy to calculate that under simulation model (4.1),

$$\begin{aligned} C(u) &= \exp\{-(1+\lambda)u\} \\ A(u) &= \exp\{\lambda u\} \\ b(u) &= \frac{1}{1+\lambda}\left[\exp\{(1+\lambda)u\} - 1\right] \\ W_{Z,X}^1(du,x) &= [1-\exp\{-x\}]\exp\{-(\lambda+1)u\}du, \end{aligned}$$

From (3.1), we find the asymptotic variance:

$$\begin{aligned}\sigma^2(y,x) &= [1-\exp\{-x\}]\frac{1}{\lambda-1}[\exp\{(\lambda-1)y]\}-1] \\ &\quad -[1-\exp\{-x\}]^2\frac{2}{\lambda+1}\left[\frac{1}{\lambda-1}\exp\{(\lambda-1)y\}-\frac{\lambda}{\lambda-1}+\frac{1}{2}\exp\{-2y\}+\frac{1}{2}\right].\end{aligned}$$

The distribution function $F(y,x)$ is displayed in Figure 1(a). The asymptotic variance function given above is shown in Figure 1(d) for $\lambda = 3/7$ which corresponds to a theoretical proportion of 70 percent uncensored observations. For reference purposes we also displayed the variance surface of the empirical distribution function for uncensored observations in Figure 1(c). Comparison of the figures in parts (c) and (d) illustrate the loss in terms of the increased variance due to censoring.

In the second simulation model we consider a bivariate exponential distribution for the vector (Y,X), which will be referred to as BIV. EXP. model. In this model,

$$\begin{array}{ll} Y=\min(U_1,U_3) & U_i \sim \text{Exp}(1),\ i=1,2,3 \\ X=\min(U_2,U_3), & U_1,U_2 \text{ and } U_3 \text{ mutually independent}, \end{array} \tag{4.2}$$

which yields the following distribution function for the random pair (Y,X)

$$F(y,x)=1-\exp\{-2x\}-\exp\{-2y\}+\exp\{2(y\vee x)-(y\wedge x)\}.$$

For a detailed discussion on multivariate exponential distributions see Marshall and Olkin (1967). Of interest is to mention that the bivariate exponential distribution function has both an absolutely continuous and a singular part. It is easy to see that in the bivariate exponential model

$$P(X=Y)=P\{U_3\le(U_1\wedge U_2)\}=\frac{1}{3},$$

and hence the distribution function has a singularity along the line $x=y$. Following the arguments in Marshall and Olkin (1967) one can show that

$$F(y,x)=\frac{2}{3}F_a(y,x)+\frac{1}{3}F_s(y,x),$$

where the absolutely continuous part is given by

$$F_a(y,x) = \frac{3}{2}[1-\exp\{-2x\}-\exp\{-2y\}+\exp\{-2(y\vee x)-(y\wedge x)\}]-\frac{1}{2}[1-\exp\{-3(y\wedge x)\}],$$

and the singular part is

$$F_s(y,x)=1-\exp\{-3(y\wedge x)\}.$$

It is easy to derive the following expressions:

$$\begin{aligned} C(u) &= \exp\{-(2+\lambda)u\} \\ A(u) &= \exp\{\lambda u\} \\ b(u) &= \frac{2}{2+\lambda}[\exp\{(2+\lambda)u\}-1]. \end{aligned}$$

Calculating the expression for the asymptotic variance is not a straightforward task, and relies on the following expressions related to the bivariate distribution function:

For $u < v$:

$$F_a(u,v) = \frac{3}{2}\left[1-\exp\{-2u\}-\exp\{-2v\}+\exp\{-2v-u\}\right]-\frac{1}{2}\left[1-\exp\{-3u\}\right]$$
$$F_a(du,dv) = 3\exp\{-2v-u\};$$

For $u > v$:

$$F_a(u,v) = \frac{3}{2}\left[1-\exp\{-2u\}-\exp\{-2v\}+\exp\{-2u-v\}\right]-\frac{1}{2}\left[1-\exp\{-3v\}\right]$$
$$F_a(du,dv) = 3\exp\{-2u-v\};$$

For $u = v$:

$$F_a(ds) = 3\exp\{-3s\}.$$

With all these ingredients we can calculate that

$$\sigma^2(y,x) = T_1(y,x) - 2T_2(y,x),$$

where

$$T_1(y,x) = \begin{cases} \dfrac{2}{2-\lambda}\left[1-\exp\{-(2-\lambda)y\}\right]-\dfrac{\exp\{-2x\}}{1-\lambda}\left[1-\exp\{-(1-\lambda)y\}\right] & \text{for } y \le x, \\[2ex] \dfrac{2}{2-\lambda}\left[1-\exp\{-(3-\lambda)x\}\right]-\dfrac{\exp\{-2x\}}{1-\lambda}\left[1-\exp\{-(1-\lambda)x\}\right] & \\ \qquad -\dfrac{2}{2-\lambda}\exp\{-(2-\lambda)y\}\left[1-\exp\{-x\}\right] & \text{for } y > x, \end{cases}$$

and

$$T_2(y,x) = \begin{cases} \int_0^y K(y,x,u)h(u)\left[2\exp\{-2u\}-\exp\{-2x-u\}\right]du & \text{for } y \le x, \\[2ex] \int_0^y K(y,x,u)h(u)\left[-\exp\{-u-2x\}+2\exp\{-2u-x\}\right]du & \\ \qquad +2\left[1-\exp\{-x\}\right]\int_0^y K(y,x,u)h(u)\exp\{-2u\}du & \text{for } y > x, \end{cases}$$

with

$$h(u) = \frac{1}{2+\lambda}\left[\lambda\exp\{(2+\lambda)u\}+2\right],$$

and

$$K(y,x,u) = \exp\{-2(x\vee y)-(x\wedge y)\}-\exp\{-2y\}+\exp\{-2u\}-\exp\{-2x-u\}.$$

Details about these nontrivial calculations can be obtained from the authors. The distribution function $F(y,x)$ under this bivariate exponential model is depicted in Figure 2 (a), and the asymptotic variance $\sigma^2(y,x)$ is shown in Figure 2 (d) for $\lambda = 2/3$, which results (theoretically) in 70 percent uncensored observations. Again, in Figure (c) we display the variance of the empirical distribution function for i.i.d. observations. For this model we observe more increase in the variance due to censoring.

4.2 Results of the simulations

The first aim of the simulation study is to investigate the performance of the proposed estimator (2.4) for the bivariate distribution function. For illustration purposes, a typical simulated realization of the suggested estimators for a sample of size $n = 400$ are displayed in Figure 1(b) and Figure 2(b) for the two models. In both samples the theoretical proportion of uncensored observations was 0.7. For these two simulated samples, the performance of the estimator seems quite satisfactory. However to observe their stable behavior, we took the average of these estimators over $N = 500$ simulation samples. We considered samples of size $n = 100$, 400 and 800. Based on the N samples we calculate $F_n^\ell(y, x)$, $\ell = 1, \cdots, N$ and evaluate

$$\frac{1}{N}\sum_{\ell=1}^{N}\left\{F_n^\ell(y,x) - F(y,x)\right\}^2,$$

for three (y, x) points. These points were (0.2, 2), (1.22, 2) and (2.5, 2) for the IND.EXP. model; and (0.2, 2), (0.4, 2) and (1.2, 2) for the BIV.EXP. model. At the same points we also evaluate the performance of the plug-in and the bootstrap estimators for the variance. The choice of the above points were motivated by the curvature structure of the variance function, as will be explained below. Table 1 summarizes the results of the simulations for the distribution function estimator. It is seen that as the sample size increases, the MSE reduces much faster in the IND.EXP. model than the BIV. EXP. model. We also observe that the MSE's in both models increase with the value of y, which may be attributed to the censoring in the right. Considering the variance functions displayed in Figures 1 (d) and 2 (d), we wanted to measure the quality of the variance estimators at three points with distinct features. For a fixed x value, three y values are chosen which correspond to an 'increasing' variance, 'maximum' variance and a relatively 'smaller' variance. The results for the plug-in estimator of the variance are presented in Table 2. We see again here that the estimation error is higher for the bivariate exponential model, the highest being again for the largest 'y' value. Observe that the impact of the increase in the sample size is highest when it increases to 400 from 100.

In order to evaluate the bootstrap procedure for accessing the unknown variance of the bivariate distribution estimate, we proceed as follows. For each simulation $(\ell = 1, \cdots, N)$ we obtain $\hat{\sigma}_\ell^2(y, x)$ as explained in (3.3). In our simulations we took B, the number of bootstrap replicates, equal to 500. To measure the performance of the bootstrap estimator we calculate

$$\frac{1}{N}\sum_{\ell=1}^{N}\hat{\sigma}_\ell^2(y,x).$$

The results for the bootstrap estimator are presented in Table 3. It was seen that the cpu time for bootstrapping is very high and the results are not very satisfactory. We therefore performed this experiment only for $n = 100$ and 400. It is seen that the bootstrap method significantly underestimates the variances at all three points for both models. Based on this comparison, we therefore can not recommend the bootstrap procedure for estimating the variance function.

	Model		IND.EXP				BIV.EXP.	
n	points:	(y,x)	(y,x)	(y,x)		(y,x)	(y,x)	(y,x)
	real	(0.2,2)	(1.22,2)	(2.5,2)		(0.2,2)	(0.4,2)	(1.2,2)
	$F(y,x)$	.1278	.4967	.6469		.2482	.4025	.5208
100	AV.	.1241	.4931	.6426		.2325	.3680	.4666
	MSE	.0012	.0039	.0039		.0023	.0039	.0061
400	AV.	.1260	.4967	.6458		.2344	.3700	.4696
	MSE	.0003	.0008	.0009		.0006	.0017	.0033
800	AV.	.1277	.4979	.6475		.2353	.3702	.4696
	MSE.	.0001	.0004	.0004		.0004	.0014	.0030

Table 1: Average estimated $F(y,x)$ and the MSE, 500 replications.

	Model		IND.EXP				BIV.EXP.	
n	points:	(y,x)	(y,x)	(y,x)		(y,x)	(y,x)	(y,x)
	Real	(0.2,2)	(1.22,2)	(2.5,2)		(0.2,2)	(0.4,2)	(1.2,2)
	Var.	.1381	.3178	.2770		.2360	.2890	.1580
100	AV.	.1355	.3338	.3500		.2161	.2652	.2522
	MSE	.0007	.0011	.0130		.0008	.0006	.0103
400	AV.	.1371	.3219	.2948		.2160	.2609	.2309
	MSE	.0002	.0002	.0012		.0005	.0008	.0056
800	AV.	.1380	.3194	.2863		.2156	.2601	.2272
	MSE.	.0001	.0001	.0005		.0005	.0008	.0050

Table 2: Average estimated variance and the MSE, 500 replications, Direct method.

	Model		IND.EXP				BIV.EXP	
n	points:	(y,x)	(y,x)	(y,x)		(y,x)	(y,x)	(y,x)
	Real	(0.2,2)	(1.22,2)	(2.5,2)		(0.2,2)	(0.4,2)	(1.2,2)
	Var.	.1381	.3178	.2770		.2360	.2890	.1580
100	AV.	.0338	.1246	.1845		.0338	.0603	.1240
	MSE	.0109	.0375	.0102		.0409	.0520	.0014
400	AV.	.0343	.1236	.1713		.0343	.0609	.1226
	MSE	.0108	.0378	.0115		.0407	.0521	.0014

Table 3: Average estimated variance and the MSE, 500 replications, Bootstrap method.

(a) (b)

(c) (d)

Figure 1: (Y, X) Independent Exponential, (a): Real distribution function, (b): Estimated distribution function for $n = 400$, censoring proportion $= 0.3$, (c): Variance of the empirical distribution function with i.i.d. observations, (d): Variance of the empirical distribution function with censored observations.

(a) (b)

(c) (d)

Figure 2 : (Y, X) Bivariate Exponential, (a): Real distribution function, (b): Estimated distribution function with $n = 400$, censoring proportion = 0.3, (c): Variance of the empirical distribution function for i.i.d. observations, (d): Variance of the empirical distribution function for censored observations.

5 Discussion and conclusion

In this paper we considered a nonparametric estimator for the bivariate distribution function when the observations for a component are subject to left truncation and right censoring. The large sample properties of the estimator are obtained via an almost sure representation by the mean of i.i.d. random variables. An expression for the covariance function of the limiting bivariate process is provided. This expression is further elaborated for the special case where the observations are subject to censoring only. An estimator for the variance for this case is proposed and its performance is investigated via simulation. In the simulations, two important models are studied for the (Y, X) pair, namely the case when they are independent with exponential distributions and when they are dependent with a joint bivariate exponential distribution. The latter one is an important common model for bivariate distributions and the nontrivial expression for the variance of the distribution function estimator for these models are provided. A bootstrap procedure for estimating the variances is also considered for comparison purposes. The results however indicated that the bootstrap alternative behaves very poorly. We therefore can not recommend the use of bootstrap for the particular cases that are studied. These results also emphasize the significance of introducing the plug-in estimator for the variance function.

References

Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993). *Statistical models based on counting processes*. Springer-Verlag, New York.

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

Efron, B. (1981). Censored data and the bootstrap. *Jour. Amer. Statist. Assoc.*, **89**, 452 – 462.

Gijbels, I. and Gürler, Ü. (1996). Covariance function of a bivariate distribution function estimator for left truncated and right censored data. *Submitted.*

Gijbels I., and Wang, J.L. (1993). Strong representations of the Survival function estimator for truncated and censored data with applications. *J. Multivariate Anal.*, **47**, 210 – 229.

Gross, S.T. and Lai T.L. (1996). Bootstrap methods for truncated and censored data. *Statistica Sinica*, **6**, 509 – 530.

Gürler, Ü. (1996a). Bivariate estimation with right truncated data. *Jour. Amer. Statist. Assoc.*, **91**, 1152 – 1165.

Gürler, Ü. (1996b). Bivariate distribution and hazard functions when a component is randomly truncated. *J. Multivariate Anal.*, to appear.

Hajek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.*, **39**, 325 – 346.

Marshall, A.W. and Olkin, I. (1967). A multivariate exponential distribution. *Jour. Amer. Statist. Assoc.*, **62**, 30 – 44.

Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameters. *Ann. Math. Statist.*, **42**, 1285 – 1295.

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

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Stute, W. (1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.,* **21**, 146 – 156.

van der Laan, M. (1996). Nonparametric estimation of the bivariate survival function with truncated data. *J. Multivariate Anal.*, **58**, 107 – 131.

Woodroofe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.*, **13**, 163 – 177.

APPENDIX

A.1. Proof of Theorem 1.

Using expressions (2.3) and (2.4) one can write

$$F_n(y,x) - F(y,x) = \int_0^y \frac{\bar{F}_{Y,n}(u-)}{C_n(u)} W^1_{Z,X,n}(du,x) - \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du,x). \tag{A.1}$$

We first decompose the above difference. To facilitate the notation we omit in a first writing the integral signs and the arguments of the functions. Note that

$$\begin{aligned} \frac{\bar{F}_{Y,n}}{C_n} W^1_{Z,X,n} - \frac{\bar{F}_Y}{C} W^1_{Z,X} &= \frac{\bar{F}_Y}{C}\left(W^1_{Z,X,n} - W^1_{Z,X}\right) + \left[\frac{\bar{F}_{Y,n}}{C_n} W^1_{Z,X,n} - \frac{\bar{F}_Y}{C} W^1_{Z,X,n}\right] \\ &\equiv I_1 + A_1. \end{aligned} \tag{A.2}$$

The term A_1 can be further decomposed as follows

$$A_1 = \frac{\bar{F}_{Y,n} - \bar{F}_Y}{C} W^1_{Z,X,n} + \left(\frac{\bar{F}_{Y,n}}{C_n} - \frac{\bar{F}_{Y,n}}{C}\right) W^1_{Z,X,n} \equiv I'_3 + A_2. \tag{A.3}$$

where

$$A_2 = \frac{\bar{F}_Y}{CC_n}(C - C_n) W^1_{Z,X,n} + \frac{\bar{F}_{Y,n} - \bar{F}_Y}{CC_n}(C - C_n) W^1_{Z,X,n} \equiv A_3 + R_{1,n}, \tag{A.4}$$

and the first term in this expression can be rewritten as

$$
\begin{aligned}
A_3 &= \frac{\bar{F}_Y}{CC_n}(C-C_n)W^1_{Z,X} + \frac{\bar{F}_Y}{CC_n}(C-C_n)\left(W^1_{Z,X,n} - W^1_{Z,X}\right) \\
&= \frac{\bar{F}_Y}{C^2}(C-C_n)W^1_{Z,X} + \frac{\bar{F}_Y}{C^2C_n}(C-C_n)^2W^1_{Z,X} + \frac{\bar{F}_Y}{CC_n}(C-C_n)\left(W^1_{Z,X,n} - W^1_{Z,X}\right) \\
&= \frac{\bar{F}_Y}{C^2}(C-C_n)W^1_{Z,X} + \frac{\bar{F}_Y}{C^2C_n}(C-C_n)^2W^1_{Z,X} + \frac{\bar{F}_Y}{C^2}(C-C_n)\left(W^1_{Z,X,n} - W^1_{Z,X}\right) \\
&\quad + \frac{\bar{F}_Y}{C^2C_n}(C-C_n)^2\left(W^1_{Z,X,n} - W^1_{Z,X}\right) \\
&= \frac{\bar{F}_Y}{C^2}(C-C_n)W^1_{Z,X} + \frac{\bar{F}_Y}{C^2}(C-C_n)\left(W^1_{Z,X,n} - W^1_{Z,X}\right) + \frac{\bar{F}_Y}{C^2C_n}(C-C_n)^2W^1_{Z,X,n} \\
&\equiv I_2 + R_{2,n} + R_{3,n}. \qquad \text{(A.5)}
\end{aligned}
$$

For the term I_3' in expression (A.3) we get, by using the representation of Gijbels and Wang (1993)

$$
\begin{aligned}
I_3' &= -\frac{1}{C}\bar{F}_Y\left(\bar{L}_n + \varepsilon_n\right)W^1_{Z,X,n} \\
&= -\frac{\bar{F}_Y}{C}\bar{L}_nW^1_{Z,X} - \frac{\bar{F}_Y}{C}\bar{L}_n\left(W^1_{Z,X,n} - W^1_{Z,X}\right) - \frac{\bar{F}_Y}{C}\varepsilon_nW^1_{Z,X,n} \\
&= -\frac{\bar{F}_Y}{C}\bar{L}_nW^1_{Z,X} - \frac{\bar{F}_Y}{C}\bar{L}_n\left(W^1_{Z,X,n} - W^1_{Z,X}\right) + \frac{\varepsilon_n'}{C}W^1_{Z,X,n} \\
&\equiv I_3 + R_{4,n} + R_{5,n}, \qquad \text{(A.6)}
\end{aligned}
$$

where we introduced the notation $\varepsilon_n' = -\bar{F}_Y\varepsilon_n$.

Combining (A.2) — (A.6) yields

$$
\frac{\bar{F}_{Y,n}}{C_n}W^1_{Z,X,n} - \frac{\bar{F}_Y}{C}W^1_{Z,X} = I_1 + I_2 + I_3 + \sum_{i=1}^{5} R_{i,n},
$$

and hence, by (A.1), leads to

$$
\begin{aligned}
F_n(y,x) - F(y,x) &= \int_0^y \frac{\bar{F}_{Y,n}(u-)}{C_n(u)}W^1_{Z,X,n}(du,x) - \int_0^y \frac{\bar{F}_Y(u-)}{C(u)}W^1_{Z,X}(du,x) \\
&= \int_0^y \frac{\bar{F}_Y(u-)}{C(u)}\left[W^1_{Z,X,n}(du,x) - W^1_{Z,X}(du,x)\right] \\
&\quad - \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)}\left[C_n(u) - C(u)\right]W^1_{Z,X}(du,x) \\
&\quad - \int_0^y \frac{\bar{F}_Y(u-)}{C(u)}\bar{L}_n(u)W^1_{Z,X}(du,x) + \sum_{i=1}^{5}\int_0^y R_{i,n} \\
&\equiv \bar{\xi}_n(y,x) + R_n(y,x),
\end{aligned}
$$

which is the representation given in (2.6) of Theorem 1.

The rest of the proof of Theorem 1 consists of dealing with the remainder term in the above representation. We start by looking at the first term in that remainder term given by (see (A.4))

$$
\int_0^y R_{1,n} = \int_0^y \frac{\bar{F}_{Y,n}(u-) - \bar{F}_Y(u-)}{C(u)\,C_n(u)}\left[C(u) - C_n(u)\right]W^1_{Z,X,n}(du,x),
$$

which can be bounded as follows

$$\left|\int_0^y R_{1,n}\right| \leq \sup_{u\leq y}|\bar{F}_{Y,n}(u-) - \bar{F}_Y(u-)| \sup_{Z_i\leq y}\frac{C(Z_i)}{C_n(Z_i)} \sup_{u\leq y}|C_n(u) - C(u)| \int_0^y \frac{1}{C^2(u)}W^1_{Z,X,n}(du,x).$$

A uniform bound over the set T_b defined in Theorem 1 is then given by

$$\begin{aligned}
&\sup_{(y,x)\in T_b}\left|\int_0^y R_{1,n}\right| \\
&\quad \leq \sup_{y\leq b}|\bar{F}_{Y,n}(y) - \bar{F}_Y(y)| \sup_{Z_i\leq b}\frac{C(Z_i)}{C_n(Z_i)} \\
&\quad\quad \times \sup_{y\leq b}|C_n(y) - C(y)| \int_0^b \frac{1}{C^2(u)}W^1_{Z,X,n}(du,x) \\
&\quad = O\left((n^{-1}\log n)^{1/2}\right) O(\log n)\, O\left((n^{-1}\log n)^{1/2}\right) \int_0^b \frac{1}{C^2(u)}W^1_{Z,X,n}(du,x).
\end{aligned}$$

Here we used the following result for the product-limit estimator:

$$\bar{F}_{Y,n}(y) - \bar{F}_Y(y) = -\bar{F}_Y(y)\bar{L}_n(y) + R_n(y),$$

with

$$\sup_{0\leq y\leq b}|R_n(y)| = O\left(n^{-1}\log n\right) \quad \text{a.s.}, \tag{A.7}$$

provided that

$$a_G < a_{V_Z},$$

and

$$\sup_{0\leq y\leq b}|R_n(y)| = O\left(n^{-1}\log^3 n\right) \quad \text{a.s.}, \tag{A.8}$$

provided that

$$\int G^{-2}(u)V_Z(du) < \infty.$$

See Gijbels and Wang (1993) and Stute (1993). We also used a result obtained by Stute (1993, Corollary 1.3, p.152):

$$\sup_{Z_i\leq b}\frac{C(Z_i)}{C_n(Z_i)} = O(\log n) \quad \text{a.s.},$$

for the case of left truncation only. The latter result can easily be extended for the right censoring and left truncation case.

Applying the strong law of large numbers we get

$$\int_0^b \frac{1}{C^2(u)}W^1_{Z,X,n}(du,x) = O(1) \quad \text{a.s.},$$

since indeed

$$\begin{aligned}
E\left\{\int_0^b \frac{1}{C^2(u)}W^1_{Z,X,n}(du,x)\right\} &= E\left\{\frac{1}{n}\sum_{i=1}^n \frac{1}{C^2(Z_i)}I(Z_i\leq b, X_i\leq x, \delta_i = 1)\right\} \\
&= \int_0^b \frac{1}{C^2(u)}W^1_{Z,X}(du,x),
\end{aligned}$$

which is finite under the condition that $a_G < a_{V_Z}$ as well as under the conditions in part (ii) of the theorem.

The remainder terms involving $R_{3,n}$ and $R_{5,n}$ can be treated similarly as $R_{1,n}$ and hence we will discuss these next.

For the remainder term involving $R_{3,n}$, defined as in (A.5), we have to look at

$$\int_0^y R_{3,n} = \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)C_n(u)} \left[C(u) - C_n(u)\right]^2 W^1_{Z,X,n}(du, x),$$

which can be bounded uniformly as follows

$$\begin{aligned}
\sup_{(y,x)\in T_b} \left|\int_0^y R_{3,n}\right| &\leq \sup_{y\leq b} \bar{F}_Y(y) \,.\, \left[\sup_{y\leq b} |C_n(y) - C(y)|\right]^2 \\
&\quad \times \sup_{Z_i\leq b} \frac{C(Z_i)}{C_n(Z_i)} \,.\, \int_0^b \frac{1}{C^3(u)} W^1_{Z,X,n}(du, x) \\
&= O\left(n^{-1}\log n\right) O(\log n) \int_0^b \frac{1}{C^3(u)} W^1_{Z,X,n}(du, x) \\
&= O\left(n^{-1}\log^2 n\right), \quad \text{a.s.}
\end{aligned}$$

since the last factor is again $O(1)$ a.s. by the strong law of large numbers, under the conditions of the theorem.

For the remainder term involving $R_{5,n}$, defined in (A.6), we obtain a uniform bound

$$\begin{aligned}
\sup_{(y,x)\in T_b} \left|\int_0^y R_{5,n}\right| &\leq \sup_{(y,x)\in T_b} \left|\frac{\varepsilon_n'(u)}{C(u)} W^1_{Z,X,n}(du, x)\right| \\
&\leq \sup_{y\leq b} |\varepsilon_n'(y)| \,.\, \int_0^b \frac{1}{C(u)} W^1_{Z,X,n}(du, x) \\
&= \sup_{y\leq b} |\varepsilon_n'(y)| \,.\, O(1) \quad \text{a.s.},
\end{aligned}$$

which is $O\left(n^{-1}\log n\right)$ a.s.; under the assumption $a_G < a_{V_Z}$, and is $O\left(n^{-1}\log^3 n\right)$ a.s. under the assumptions of part (ii) of the theorem. See (A.7) and (A.8).

It remains to check the remainder terms involving $R_{2,n}$ and $R_{4,n}$.

Note that the term involving $R_{2,n}$, which is defined in (A.5), can be rewritten as follows

$$\begin{aligned}
\int_0^y R_{2,n} &= \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} \left[C(u) - C_n(u)\right] \left[W^1_{Z,X,n}(du, x) - W^1_{Z,X}(du, x)\right] \\
&= \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X,n}(du, x) - \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X,n}(du, x) \\
&\quad - \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du, x) + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X}(du, x). \qquad \text{(A.9)}
\end{aligned}$$

The second term in the above expression can be handled by using Lemma 1, formulated and proved in Appendix A.2, and which reveals that

$$\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X,n}(du, x)$$

$$= \frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du,x) + \frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X,n}(du,x)$$

$$+\frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X}(du,x) - F(y,x) + \frac{1}{n}F(y,x) + E^1_n(y,x), \quad \text{(A.10)}$$

with

$$\sup |E^1_n(y,x)| = o\left(n^{-1}(\log n)^{\delta}\right) \quad \text{a.s.} \qquad \text{for } \delta > 1/2. \qquad \text{(A.11)}$$

Substituting (A.10) into (A.9) we find

$$\int_0^y R_{2,n} \;=\; -\frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du,x) + \frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X,n}(du,x)$$

$$+\frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X}(du,x) - \frac{1}{n}F(y,x) - E^1_n(y,x),$$

and hence

$$\int_0^y R_{2,n} = -E^1_n(y,x) + O\left(n^{-1}\right) \quad \text{a.s.},$$

with the a.s. behavior of the first term as given in (A.11).

A similar result can be proven for the term involving $R_{4,n}$. The basic result establishing the behavior of this remainder term is provided in Lemma 2. □

A.2. Lemmas.

The proof of Theorem 1 uses two lemmas which are formulated and proved below. The proofs of the lemmas rely on exploiting the U-statistic structure of the terms.

Lemma 1:

Let

$$U_n(y,x) = \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X,n}(du,x),$$

and denote by $\hat{U}_n(y,x)$ the Hajek projection (see Hajek (1968)) of $U_n(y,x)$. Then,

$$U_n(y,x) = \hat{U}_n(y,x) + E^1_n(y,x),$$

with

$$\hat{U}_n(y,x) \;=\; \frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du,x) + \frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X,n}(du,x)$$

$$+\frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X}(du,x) - F(y,x) + \frac{1}{n}F(y,x),$$

and

$$\sup |E^1_n(y,x)| = o\left(n^{-1}(\log n)^{\delta}\right) \quad \text{a.s.} \qquad \text{for } \delta > 1/2.$$

Proof of Lemma 1: Note first that $U_n(y,x)$ can be rewritten in the following form

$$U_n(y,x) \;=\; \frac{1}{n}\sum_{i=1}^n \frac{\bar{F}_Y(Z_i-)}{C^2(Z_i)} C_n(Z_i) I(Z_i \le y, X_i \le x, \delta_i = 1)$$

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{\substack{i,j=1\\ i\neq j}}^{n}\frac{\bar{F}_Y(Z_i-)}{C^2(Z_i)}I(T_j\le Z_i\le Z_i)I(Z_i\le y, X_i\le x,\delta_i=1)\\
&\quad+\frac{1}{n^2}\sum_{i=1}^{n}\frac{\bar{F}_Y(Z_i-)}{C^2(Z_i)}I(Z_i\le y, X_i\le x,\delta_i=1)\\
&\equiv (I)_a+(II)_a. \qquad\qquad (A.12)
\end{aligned}
$$

Introducing the notations

$$D_j = (Z_j, X_j, T_j, \delta_j)$$

$$h(D_i, D_j; y, x) = \frac{\bar{F}_Y(Z_i-)}{C^2(Z_i)}I(T_j\le Z_i\le Z_j)I(Z_i\le y, X_i\le x,\delta_i=1)$$

$$h^*(D_i, D_j; y, x) = \frac{1}{2}\left[h(D_i, D_j; y, x)+h(D_j, D_i; y, x)\right],$$

where $h^*(D_i, D_j; y, x)$ is the symmetrized version of $h(D_i, D_j; y, x)$, we can write the term $(I)_a$ as

$$
\begin{aligned}
(I)_a &= \frac{1}{n^2}\sum_{1\le i<j\le n}\left[h(D_i, D_j; y, x)+h(D_j, D_i; y, x)\right]\\
&= \frac{2}{n^2}\sum_{1\le i<j\le n}h^*(D_i, D_j; y, x)\\
&= \frac{n-1}{n}\left[\binom{n}{2}\right]^{-1}\sum_{1\le i<j\le n}h^*(D_i, D_j; y, x)\\
&\equiv \frac{n-1}{n}U_n^*(y,x), \qquad\qquad (A.13)
\end{aligned}
$$

with $U_n^*(y,x)$ a U-statistic of degree 2 with symmetric kernel $h^*(\cdot,\cdot;y,x)$.

Now we know from Theorem 5.3.3 in Serfling (1980, p. 189) that

$$U_n^*(y,x) = \hat{U}_n^*(y,x) + R_n^*(y,x), \qquad (A.14)$$

with

$$R_n^*(y,x) = o\left(n^{-1}(\log n)^{\delta}\right) \quad \text{a.s.} \qquad \text{for } \delta > 1/2, \qquad (A.15)$$

and where

$$\hat{U}_n^*(y,x) = \sum_{i=1}^{n}E\left\{U_n^*(y,x)|D_i\right\} - (n-1)E\left\{h^*(D_1, D_2; y, x)\right\}. \qquad (A.16)$$

It can be checked that the conditions needed for (A.15) are satisfied under the assumptions of part (i) and part (ii) of the theorem.

Calculation of

$$E\left\{U_n^*(y,x)|D_i\right\} = \left[\binom{n}{2}\right]^{-1}\sum_{1\le k<\ell\le n}E\left\{h^*(D_k, D_\ell; y, x)|D_i\right\} \qquad (A.17)$$

is rather straightforward and is outlined below. For the index i fixed this involves looking at three cases, namely, Case 1: $i = k < \ell$; Case 2: $k < \ell = i$, and Case 3: $k < \ell, i \neq k$ and $i \neq \ell$. For example for Case 1, one obtains

$$\begin{aligned} E\left\{h^*(D_i, D_\ell; y, x)|D_i\right\} &= \frac{1}{2}\left[E\left\{h^*(D_i, D_\ell; y, x)|D_i\right\} + E\left\{h^*(D_\ell, D_i; y, x)|D_i\right\}\right] \\ &= \frac{1}{2}\left[\frac{\bar{F}_Y(Z_i-)}{C(Z_i)} I(Z_i \leq y, X_i \leq x, \delta_i = 1)\right. \\ &\quad \left. + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} I(T_i \leq u \leq Z_i) W^1_{Z,X}(du, x)\right] \end{aligned} \tag{A.18}$$

Case 2 is treated in a similar way. For Case 3 we proceed as follows:

$$\begin{aligned} E\left\{h^*(D_k, D_\ell; y, x)|D_i\right\} &= E\left\{h^*(D_k, D_\ell; y, x)\right\} \\ &= E\left\{\left\{h^*(D_k, D_\ell; y, x)|D_k\right\}\right\} \\ &= E\left\{\frac{1}{2}\left[\frac{\bar{F}_Y(Z_k-)}{C(Z_k)} I(Z_k \leq y, X_k \leq x, \delta_k = 1)\right.\right. \\ &\quad \left.\left. + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} I(T_k \leq u \leq Z_k) W^1_{Z,X}(du, x)\right]\right\} \\ &= \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du, x). \end{aligned} \tag{A.19}$$

From (A.17), (A.18) and (A.19) we get

$$\begin{aligned} \sum_{i=1}^n E\left\{U_n^*(y, x)|D_i\right\} &= \left[\binom{n}{2}\right]^{-1} \left\{\frac{n-1}{2} \sum_{i=1}^n \left[\frac{\bar{F}_Y(Z_i-)}{C(Z_i)} I(Z_i \leq y, X_i \leq x, \delta_i = 1)\right.\right. \\ &\quad \left. + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} I(T_i \leq u \leq Z_i) W^1_{Z,X}(du, x)\right] \\ &\quad \left. + \frac{(n-1)(n-2)}{2} \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du, x)\right\} \\ &= \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X,n}(du, x) + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X}(du, x) \\ &\quad + (n-2) \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du, x). \end{aligned} \tag{A.20}$$

The second term in (A.16) yields

$$(n-1) E\left\{E\left[h^*(D_1, D_2; y, x)|D_1\right]\right\} = (n-1) \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du, x), \tag{A.21}$$

and hence combination of (A.13), (A.14), (A.16), (A.20) and (A.21) leads to the following expression for the first term in (A.12)

$$\begin{aligned} (I)_a &= \frac{n-1}{n}\left\{\int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X,n}(du, x) + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X}(du, x)\right. \\ &\quad \left. - \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du, x)\right\} + o\left(n^{-1}(\log n)^\delta\right) \quad \text{a.s.} \end{aligned} \tag{A.22}$$

It remains to consider the second term in (A.12). Note that

$$(II)_a = \frac{1}{n} \cdot \frac{1}{n} \sum_{i=1}^{n} \tilde{h}(D_i; y, x) \equiv \frac{1}{n} \tilde{U}_n(y, x),$$

with

$$\tilde{h}(D_i; y, x) = \frac{\bar{F}_Y(Z_i-)}{C^2(Z_i)} I(Z_i \le y, X_i \le x, \delta_i = 1),$$

and where $\tilde{U}_n(y,x)$ is a U-statistic of degree one. Applying once again Theorem 5.3.3. in Serfling (1980, p. 189) we can write

$$\tilde{U}_n(y, x) = \hat{\tilde{U}}_n(y, x) + \tilde{R}_n(y, x),$$

with

$$\tilde{R}_n(y, x) = o\left(n^{-1} (\log n)^{\delta}\right) \quad \text{a.s.} \qquad \text{for } \delta > 1/2, \tag{A.23}$$

and where

$$\hat{\tilde{U}}_n(y, x) = \sum_{i=1}^{n} E\left\{\tilde{U}_n(y, x) | D_i\right\} - (n-1) E\left\{\tilde{h}(D_1; y, x)\right\}.$$

Again it can be shown that the conditions for validity of (A.23) are satisfied here.

Calculations similar to these for term $(I)_a$ lead to

$$\begin{aligned}
E\left\{\tilde{U}_n(y, x) | D_i\right\} &= \frac{1}{n} \sum_{j=1}^{n} E\left\{\tilde{h}(D_j; y, x) | D_i\right\} \\
&= \frac{n-1}{n} E\left\{\tilde{h}(D_1; y, x)\right\} + \frac{1}{n} \tilde{h}(D_i; y, x) \\
&= \frac{n-1}{n} \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X}(du, x) \\
&\quad + \frac{1}{n} \frac{\bar{F}_Y(Z_i-)}{C^2(Z_i)} I(Z_i \le y, X_i \le x, \delta_i = 1),
\end{aligned}$$

and hence

$$\sum_{i=1}^{n} E\left\{\tilde{U}_n(y, x) | D_i\right\} = (n-1) \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X}(du, x) + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du, x). \tag{A.24}$$

Further we get

$$E\left\{\tilde{h}(D_1; y, x)\right\} = \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X}(du, x), \tag{A.25}$$

and hence combining (A.24) and (A.25) we find

$$\begin{aligned}
\frac{1}{n} \hat{\tilde{U}}_n(y, x) &= \frac{n-1}{n} \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X}(du, x) + \frac{1}{n} \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du, x) \\
&\quad - \frac{n-1}{n} \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X}(du, x) \\
&= \frac{1}{n} \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du, x).
\end{aligned}$$

Now from (A.12), (A.22) and the above expression one obtains

$$\begin{aligned} U_n(y,x) &= (I)_a + (II)_a \\ &= \frac{n-1}{n}\left\{\int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X,n}(du,x) + \int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} C_n(u) W^1_{Z,X}(du,x)\right\} \\ &\quad + \frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du,x) - \frac{n-1}{n} F(y,x) + o\left(n^{-1}(\log n)^\delta\right) \text{ a.s.,} \end{aligned}$$

where we used that

$$F(y,x) = \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} W^1_{Z,X}(du,x),$$

which is easily seen from (2.1) and (2.2) . This completes the proof of Lemma 1. □

Lemma 2: Let

$$V_n(y,x) = \int_0^y \frac{\bar{F}_Y(u-)}{C(u)} \bar{L}_n(u) W^1_{Z,X,n}(du,x),$$

and denote by $\hat{V}_n(y,x)$ the Hajek projection of $U_n(y,x)$. Then,

$$V_n(y,x) = \hat{V}_n(y,x) + E_n^2(y,x),$$

with

$$\begin{aligned} \hat{V}_n(y,x) &= \frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du,x) \\ &\quad -\frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C(u)}\left[\int_0^u \frac{1}{C^2(s)} W^1_Z(ds)\right] W^1_{Z,X,T,n}(du,x,u) \\ &\quad +\frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C(u)}\left[\int_0^u \frac{1}{C(s)} W^1_{Z,n}(ds)\right] W^1_{Z,X}(du,x) \\ &\quad -\frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C(u)}\left[\int_0^u \frac{C_n(s)}{C^2(s)} W^1_Z(ds)\right] W^1_{Z,X}(du,x), \end{aligned}$$

and

$$\sup|E_n^2(y,x)| = o\left(n^{-1}(\log n)^\delta\right) \quad \text{a.s.} \qquad \text{for } \delta > 1/2.$$

Proof of Lemma 2: The proof is based on similar techniques as those used for proving Lemma 1, and for brevity we ignore details here. Note that, using the definition of $\bar{L}_n(\cdot)$ as given in (2.5), one can rewrite the quantity $V_n(y,x)$ as follows

$$\begin{aligned} V_n(y,x) &= \frac{1}{n^2}\sum_{\substack{i,j=1\\ i\neq j}}^n \frac{\bar{F}_Y(Z_i-)}{C(Z_i)}\left\{\frac{I(Z_j\le Z_i,\delta_j=1)}{C(Z_j)} - \int_0^{Z_i}\frac{I(T_j\le u\le Z_j)}{C^2(u)} W^1_Z(du)\right\} \\ &\qquad \times I(Z_i\le y, X_i\le x, \delta_i=1) \\ &\quad +\frac{1}{n^2}\sum_{i=1}^n \frac{\bar{F}_Y(Z_i-)}{C(Z_i)}\left\{\frac{1}{C(Z_i)} - \int_0^{Z_i}\frac{I(T_i\le u\le Z_i)}{C^2(u)} W^1_Z(du)\right\} I(Z_i\le y, X_i\le x, \delta_i=1) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{\substack{i,j=1\\ i\neq j}}^{n} \frac{\bar F_Y(Z_i-)}{C(Z_i)}\left\{\frac{I(Z_j\le Z_i,\delta_j=1)}{C(Z_j)} - \int_0^{Z_i}\frac{I(T_j\le u\le Z_j)}{C^2(u)}W_Z^1(du)\right\}\\
&\qquad\qquad \times I(Z_i\le y, X_i\le x,\delta_i=1)\\
&\quad + \frac{1}{n^2}\sum_{i=1}^{n}\frac{\bar F_Y(Z_i-)}{C(Z_i)}\left\{\frac{1}{C(Z_i)} - \int_0^{y}\frac{I(T_i\le u\le Z_i)}{C^2(u)}W_Z^1(du)\right\} I(Z_i\le y, X_i\le x,\delta_i=1)\\
&\equiv (I)_b + (II)_b. \qquad\qquad (A.26)
\end{aligned}
$$

As in Lemma 1 we write

$$
(I)_b = \frac{n-1}{n}\left[\begin{pmatrix} n\\ 2\end{pmatrix}\right]^{-1}\sum_{1\le i<j\le n} h^*(D_i,D_j;y,x) \equiv \frac{n-1}{n}V_n^*(y,x),
$$

where now

$$
h^*(D_i,D_j;y,x) = \frac{1}{2}\left[h(D_i,D_j;y,x) + h(D_j,D_i;y,x)\right],
$$

with

$$
\begin{aligned}
h(D_i,D_j;y,x) &= \frac{\bar F_Y(Z_i-)}{C(Z_i)}\left\{\frac{I(Z_j\le Z_i,\delta_i=1)}{C(Z_j)} - \int_0^{Z_i}\frac{I(T_j\le u\le Z_j)}{C^2(u)}W_Z^1(du)\right\}\\
&\qquad \times I(Z_i\le y, X_i\le x,\delta_i=1).
\end{aligned}
$$

As before we search for the Hajek projection $\hat V_n^*(y,x)$ of $V_n^*(y,x)$:

$$
\hat V_n^*(y,x) = \sum_{i=1}^{n} E\left\{V_n^*(y,x)|D_i\right\} - (n-1)E\left\{h^*(D_1,D_2;y,x)\right\}.
$$

Via similar arguments as before we find that

$$
\begin{aligned}
\sum_{i=1}^{n} E\left\{V_n^*(y,x)|D_i\right\} &= \int_0^y \frac{\bar F_Y(v-)}{C(v)}\left[\int_0^y \frac{1}{C(u)}W_{Z,n}^1(du)\right]W_{Z,X}^1(dv,x)\\
&\quad - \int_0^y \frac{\bar F_Y(v-)}{C(v)}\left[\int_0^y \frac{C_n(u)}{C^2(u)}W_Z^1(du)\right]W_{Z,X}^1(dv,x),
\end{aligned}
$$

and for the first term in (A.26) we obtain

$$
\begin{aligned}
(I)_b &= \frac{n-1}{n}\int_0^y \frac{\bar F_Y(v-)}{C(v)}\left[\int_0^y \frac{1}{C(u)}W_{Z,n}^1(du)\right]W_{Z,X}^1(dv,x)\\
&\quad - \frac{n-1}{n}\int_0^y \frac{\bar F_Y(v-)}{C(v)}\left[\int_0^y \frac{C_n(u)}{C^2(u)}W_Z^1(du)\right]W_{Z,X}^1(dv,x) + o\left(n^{-1}(\log n)^{\delta}\right) \quad \text{a.s.}
\end{aligned}
\qquad (A.27)
$$

To deal with the second term in (A.26) we rewrite this term as

$$
\begin{aligned}
(II)_b &= \frac{1}{n}.\frac{1}{n}\sum_{i=1}^{n}\frac{\bar F_Y(Z_i-)}{C^2(Z_i)}I(Z_i\le y, X_i\le x,\delta_i=1)\\
&\quad - \frac{1}{n}.\frac{1}{n}\sum_{i=1}^{n}\frac{\bar F_Y(Z_i-)}{C(Z_i)}\int_0^y \frac{I(T_i\le u\le Z_i)}{C^2(u)}W_Z^1(du)\ .\ I(Z_i\le y, X_i\le x,\delta_i=1)\\
&= \frac{1}{n}\int_0^y \frac{\bar F_Y(u-)}{C^2(u)}W_{Z,X,n}^1(du,x)\\
&\quad - \frac{1}{n}\int_0^y \frac{\bar F_Y(v-)}{C(v)}\int_0^v \frac{1}{C^2(u)}W_Z^1(du)W_{Z,X,T,n}^1(dv,x,u). \qquad (A.28)
\end{aligned}
$$

Hence from (A.27) and (A.28) we obtain that the main term is given by

$$\begin{aligned}
&\frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(v-)}{C(v)}\left[\int_0^y \frac{1}{C(u)} W^1_{Z,n}(du)\right] W^1_{Z,X}(dv,x)\\
&-\frac{n-1}{n}\int_0^y \frac{\bar{F}_Y(v-)}{C(v)}\left[\int_0^y \frac{C_n(u)}{C^2(u)} W^1_{Z}(du)\right] W^1_{Z,X}(dv,x)\\
&+\frac{1}{n}\int_0^y \frac{\bar{F}_Y(u-)}{C^2(u)} W^1_{Z,X,n}(du,x) - \frac{1}{n}\int_0^y \frac{\bar{F}_Y(v-)}{C(v)}\int_0^v \frac{1}{C^2(u)} W^1_Z(du) W^1_{Z,X,T,n}(dv,x,u),
\end{aligned}$$

which completes the proof of Lemma 2. □