

# Bivariate Distribution and Hazard Functions When a Component is Randomly Truncated

Ülkü Gürler

*Bilkent University, Ankara, Turkey*

In random truncation models one observes the i.i.d. pairs  $(T_i \leq Y_i)$ ,  $i = 1, \dots, n$ . If  $Y$  is the variable of interest, then  $T$  is another independent variable which prevents the complete observation of  $Y$  and random left truncation occurs. Such a type of incomplete data is encountered in medical studies as well as in economy, astronomy, and insurance applications. Let  $(Y, Y)$  be a bivariate vector of random variables with joint distribution function  $F(y, x)$  and suppose the variable  $Y$  is randomly truncated from the left. In this study, nonparametric estimators for the bivariate distribution and hazard functions are considered. A nonparametric estimator for  $F(y, x)$  is proposed and an a.s. representation is obtained. This representation is used to establish the consistency and the weak convergence of the empirical process. An expression for the variance of the asymptotic distribution is presented and an estimator is proposed. Bivariate “diverse-hazard” vector is introduced which captures the individual and joint failure behaviors of the random variables in opposite “time” directions. Estimators for this vector are presented and the large sample properties are discussed. Possible applications and a moderate size simulation study are also presented. © 1997 Academic Press

## 1. INTRODUCTION AND PRELIMINARIES

### 1.1. Introduction

In survival or reliability studies, incomplete data is frequently encountered. Random truncation and censoring are two common forms of such data. In random left truncation model, one observes the i.i.d. pairs  $(Y_i \geq T_i)$ ,  $i = 1, \dots, n$ , where  $Y$  is the variable of interest and  $T$  is another independent variable which prevents the complete observation of  $Y$ . Random right truncation is similarly defined by interchanging the roles of  $Y$  and  $T$ . One of the earliest applications of the left truncation model was given by Lynden–Bell (1971), where  $Y$  refers to the brightness of celestial objects, which is only partially observable due to a preventing variable  $T$ .

Received March 17, 1994; revised September 1995.

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

Key words and phrases: bivariate distribution, bivariate diverse-hazard, truncation, weak convergence, nonparametric estimation.

Random truncation models gained more interest in recent years because it is conveniently used to model several aspects of AIDS data, such as the distribution of the incubation time, the reporting lags from the detection of the incidence to the time it is reported to the officials, or the time from AIDS to death. Suppose for instance  $T_o$  is the time when the observation period starts, and let  $a$  be the time a person is diagnosed with AIDS and  $d$  is the time of death. If we set  $Y = d - a$  and  $T = T_o - a$ , then only those individuals are observed for whom  $T \leq Y$  and left truncation occurs. This selection bias becomes more important especially in the early stages of the epidemics, when sufficient historical data has not accumulated yet. Truncated data could also arise in insurance applications where a liability claim may arise due to an incidence but a delay occurs until it is reported to the insurance company. See e.g. Kalbfleisch and Lawless (1989) for such applications.

In the present study we consider the estimation of the bivariate distribution function (d.f.) and a different version of a bivariate hazard, namely diverse-hazard vector, when a component is randomly truncated. Bivariate d.f. is important in understanding the joint behavior of correlated random variables, as well as assessing the strength of such relations. Nonparametric or parametric regression, bivariate density estimation or developing tests of independence could be cited as some potential applications of the methods described in this study. Bivariate diverse-hazard vector is defined in analogy to the hazard vector introduced by Dabrowska (1988) for the censored observations. The diverse-hazard vector captures the immediate past and future failure characteristics of the individual variables as well as their joint failure behavior. Besides being of interest on its own, the functionals of this vector could be used to develop tests of independence as discussed in Section 3. Although the estimators presented in the paper can be used for arbitrary distributions, we assume the continuity of  $F$  in both components for establishing the large sample properties, which in turn inherit such assumptions from the results adopted from univariate case. Similarly, without loss of generality, the pair  $(Y, X)$  is taken to be nonnegative.

The paper is organized as follows: In the next subsection, main results for the univariate case are summarized. In Section 2, an estimator for  $F(y, z)$  is proposed and an almost sure (a.s) i.i.d. representation is obtained. Strong consistency and weak convergence are then established via this representation. The variance of the limiting distribution is presented and an estimator for it is suggested. In Section 3, the bivariate diverse-hazard vector is introduced and estimation procedures are discussed. A decomposition of an arbitrary bivariate d.f. in terms of the marginal distributions and a functional of the diverse-hazard vector is introduced in analogy to that of Dabrowska (1988). Two alternative methods to estimate the integrated diverse-hazard vector are discussed and their large sample equivalence is

established together with an i.i.d. representation. A discussion on how to utilize the results for developing tests of independence is also included. Finally in Section 4, some simulation results are presented which illustrate the performance of the bivariate d.f. estimator. In this section some remarks are also made concerning the possible extensions of the models considered. Proofs of most of the results are deferred to the Appendix.

## 1.2. Preliminaries

We now present some preliminary results for the univariate truncation model. Let  $Y$  be the variable of interest and  $T$  be the truncating variable with d.f.'s  $F$  and  $G$  respectively. The pairs  $(Y, T)$  are observed only if  $(T_i \leq Y_i)$ ,  $i = 1, \dots, n$ . Under this sampling scheme, Woodroffe (1985) points out that  $F$  and  $G$  can be estimated completely only if  $(F, G) \in R_o$ , where  $R_o = \{(F, G): a_G \leq a_F, b_G \leq b_F\}$  with  $a_W$  and  $b_W$  denoting the lower and upper end points of the support of any distribution function (d.f.)  $W$  respectively. Note that  $(F, G) \in R_o$  implies  $\alpha \equiv P(T \leq Y) > 0$ , which is assumed throughout. The nonparametric maximum likelihood estimator of  $F$  is first given by Lynden-Bell (1971) as

$$\bar{F}_n(y) = \prod_{i: Y_i \leq y} \left[ 1 - \frac{s(Y_i)}{nC_n(Y_i)} \right],$$

where

$$nC_n(u) = \# \{i: T_i \leq u \leq Y_i\} \quad s(u) = \# \{i: Y_i = u\}. \quad (1)$$

Consistency of  $F_n$  and its right truncation counterpart are studied by Woodroffe (1985) and Wang, Jewell and Tsai (1986). Chao and Lo (1988) derived a representation of  $(F_n - F)$  as i.i.d. mean processes. The order of the remainder term for this representation is improved by Stute (1993) and Gijbels and Wang (1993) (see Theorem 1 below). Kernel estimators of the hazard function for truncated/censored data are studied by Uzunoğulları and Wang (1992). Gürler, Stute and Wang (1993), Gu and Lai (1990), Lai and Ying (1991), and Gross and Huber-Carol (1992) extended the results for truncated/censored data in various directions. Keiding and Gill (1990) provided a Markov-Process approach to the model and derived similar results with martingale methods.

The following theorem summarizes the existing results concerning  $F_n(y)$ . Let  $(Y_i, T_i)$ ,  $i = 1, \dots, n$  denote the observed variables. Define

$$L_i(z) = \alpha \left\{ \frac{I(Y_i \leq z)}{G(Y_i) \bar{F}(Y_i)} - \int_0^z \frac{I(T_i \leq u \leq Y_i)}{G(u) \bar{F}^2(u)} dF(u) \right\}$$

and

$$\bar{L}_n(z) = \sum_{i=1}^n n^{-1} L_i(z).$$

**THEOREM 1.** *Suppose  $F$  is continuous and  $b < b_F$ . Then the following representation holds:*

$$\begin{aligned} F_n(z) - F(z) &= n^{-1} \bar{F}(z) \sum_{i=1}^n L_i(z) + \varepsilon_n(z) \\ &= \bar{F}(z) \bar{L}_n(z) + \varepsilon_n(z). \end{aligned}$$

(i) (Chao and Lo (1988)). Let  $a_G = a_F$ . If  $\lim_{x \rightarrow 0+} F(x)/G(x) = 0$  and  $\int dF/G < \infty$ , then

$$\sup_{0 \leq z \leq b} |\varepsilon_n(z)| = o(n^{-1/2}) \text{ a.s.}$$

(ii) (Gijbels and Wang (1993)). Let  $a_G < a_F$ . If  $\int dF/G < \infty$ , then

$$\sup_{0 \leq z \leq b} |\varepsilon_n(z)| = O(\log n/n) \text{ a.s.}$$

(iii) (Stute (1993a)). Let  $a_G \leq a_F$ . If  $\int dF/G^2 < \infty$ , then

$$\sup_{a_{F_y} \leq z \leq b} |\varepsilon_n(z)| = O(\log^3 n/n) \text{ a.s.}$$

**COROLLARY 1.** *Under the conditions of Theorem 1, the process  $W_n(z) = \sqrt{n}(F_n(z) - F(z))$  converges weakly to a zero mean Gaussian process on  $D[0, b]$ , with covariance structure:*

$$\text{Cov}(W_n(z_1), W_n(z_2)) = \alpha \bar{F}(z_1) \bar{F}(z_2) \int_0^{z_1 \wedge z_2} \frac{F(du)}{G(u) \bar{F}^2(u)}.$$

## 2. BIVARIATE DISTRIBUTION FUNCTION

### 2.1. Suggested Estimator

We now consider the bivariate truncation model, in which one observes the triplets  $(Y_i, X_i, T_i)$ ,  $i = 1, \dots, n$  only if  $(T_i \leq Y_i)$ . The purpose is to estimate the bivariate d.f.  $F(y, x)$  of  $(Y, X)$ . Here  $T$  is a nuisance random variable, which is assumed to be independent of  $(Y, X)$ , with d.f.  $G$ . All the variables are assumed to be continuous, nonnegative. The marginal d.f.'s of

$Y$  and  $X$  are denoted by  $F_Y$  and  $F_X$  respectively. To avoid the identifiability problems, it is assumed that  $a_{F_Y} \geq a_G$  and  $b_{F_Y} \geq b_G$ , as in the univariate model. Given this model, the observed triplets can be considered to arise from the following trivariate conditional distribution  $H$ :

$$\begin{aligned} H_{Y,X,T}(y, x, t) &= P(Y \leq y, X \leq x, T \leq t \mid Y \geq T) \\ &= \alpha^{-1} \int_0^x \int_0^y G(t \wedge u) dF(u, v); \end{aligned}$$

here  $\alpha$  is as defined before, and  $t \wedge u = \min(y, u)$ . The observed pairs then have the following distributions:

$$\begin{aligned} F_{Y,X}^*(y, x) &\equiv H_{Y,X,T}(y, x, \infty) = \alpha^{-1} \int_0^x \int_0^y G(u) dF(u, v) \\ H_{X,T}(x, t) &= \alpha^{-1} \int_0^x \int_0^\infty G(t \wedge u) dF(u, v) \quad (2) \\ H_{Y,T}(y, t) &= \alpha^{-1} \int_0^\infty \int_0^y G(t \wedge u) dF(u, v). \end{aligned}$$

The univariate marginals are obtained as

$$\begin{aligned} F_Y^*(y) &= \alpha^{-1} \int_0^\infty \int_0^y G(u) dF(u, v) \\ F_X^*(x) &= \alpha^{-1} \int_0^x \int_0^\infty G(u) dF(u, v) \\ G_T^*(t) &= H_{Y,T}(\infty, t) = \alpha^{-1} \int_0^\infty \int_0^\infty G(t \wedge u) dF(u, v). \end{aligned}$$

Assuming the existence of the densities (denoted in lowercase letters), we have

$$\begin{aligned} f_{Y,X}^*(y, x) &= \alpha^{-1} G(y) f_{Y,X}(y, x) \\ f_Y^*(y) &= \alpha^{-1} G(y) f_Y(y) \\ f_X^*(x) &= \alpha^{-1} \int_0^\infty G(u) f(u, x) du. \end{aligned}$$

The following function is of importance in truncation model, and the scaled empirical counterpart of it defined in (1) is the size of the 'risk set' at time  $z$ :

$$\begin{aligned} C(z) &= \alpha^{-1} G(z) \bar{F}_Y(z-) \\ &= G_T^*(z) - F_Y^*(z). \end{aligned} \quad (3)$$

It is interesting to note here that the conditional density of  $X$  given  $Y$  in the truncation model is the same as that of untruncated model. This implies that inference for the conditional distribution of  $X$  given  $Y$  can be based on the observed truncated sample. However the reverse is not true for  $Y$  given  $X$  and the procedures proposed in this study can be used to handle this case.

The estimator considered for  $F(y, x)$  is

$$F_n(y, x) = \frac{1}{n} \sum_i \frac{\bar{F}_{Y,n}(Y_i -)}{C_n(Y_i)} I(Y_i \leq y, X_i \leq x),$$

where  $\bar{F}_{Y,n}$  and  $C_n$  are as given in (1). This estimator is motivated by observing that

$$f(y, x) = [\alpha^{-1} G(y)]^{-1} f^*(y, x),$$

where

$$[\alpha^{-1} G(y)]^{-1} = \frac{\bar{F}_Y(y)}{C(y)},$$

which follows from (2) and (3).  $F_n(y, x)$  reduces to the product limit estimator (1) of  $F_Y(y)$  when  $x \rightarrow \infty$ . It can easily be verified that  $F_n(y, x)$  is a bivariate distribution function. Stute (1993, b) proposes an estimator for the censored case which is analogous to  $F_n$  and is a bivariate distribution function.

In the censored case when both components are subject to censoring, many of the existing estimators of the bivariate d.f. lack one or more of the requirements to be a proper d.f., as discussed in Dabrowska (1988). For the truncation model, there doesn't exist an estimator of  $F(y, x)$  when both components are truncated, in which case more delicate identifiability problems arise. Therefore it is not possible to present a direct comparison of bivariate censoring and truncation methods. However it will become apparent in the next section that there are obvious similarities in the structures of the bivariate hazard and d.f. estimators between the present case and those in the bivariate censored data. Therefore, even though the

sampling mechanisms of them are quite different, it is not unfair to say that the technical difficulties involved in singly (only a component) truncated bivariate data are comparable to those of bivariate censored data.

## 2.2. An Almost Sure Representation

Observe that  $F_n(y, x)$  is a weighted sum of i.i.d. variables, where the weights are the jump sizes of the truncation product limit estimator  $F_{Y,n}(y)$  at the data points. Therefore theoretical properties of this estimator are strongly related to those of  $F_{Y,n}(y)$ . The following theorem provides an i.i.d. representation of  $F_n(y, x)$ , the proof of which is given in the appendix.

**THEOREM 2.** *Assume  $F$  is continuous in both components,  $b < b_{F_Y}$  and let  $T_b = \{(y, x): 0 < y < b; 0 < x < \infty\}$ . Then  $F_n(y, x)$  admits the representation*

$$\begin{aligned} F_n(y, x) - F(y, x) &= \int_0^y \frac{\bar{F}(u)}{C(u)} [F_n^*(du, x) - F^*(du, x)] \\ &\quad + \int_0^y \frac{\bar{F}(u)}{C^2(u)} [C(u) - C_n(u) \\ &\quad + \bar{L}_n(u) C(u)] F^*(du, x) + R_n(y, x) \\ &\equiv \bar{\xi}_n(y, x) + R_n(y, x) \end{aligned} \quad (4)$$

and

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

$$\sup_{(y, x) \in T_b} |R_n(y, x)| = O(\log^2 n/n).$$

(ii) If  $a_G = a_{F_Y}$ , and  $\int G^{-2}(u) F_Y(du) < \infty$ , then

$$\sup_{(y, x) \in T_b} |R_n(y, x)| = O(\log^3 n/n) = o(n^{-1/2}).$$

Notice here that the order of the remainder term for part (i) is better than that of Chao and Lo (1988), but not as good as the result of Gijbels and Wang (1993). It may be another task to further improve this result to achieve a similar bound. The magnitude of the remainder term for part (ii) derives from the result of Stute (1993a) and therefore the integrability condition here is more restrictive than that of Chao and Lo (1988). Note however that, starting with the result of Chao and Lo (1988), one can obtain the same order of magnitude  $o(n^{-1/2})$  as theirs for the remainder term, by merely the assuming the conditions of part (ii) of Theorem 1.

Theorem 2 can also be utilized to establish the weak convergence of  $F_n(y, x)$ . Weak convergence of the empirical processes in multidimensional time is considered in Neuhaus (1971) and Straf (1972). Campbell (1981) used the results of Neuhaus (1971) to establish the weak convergence of the bivariate process in the censored case. For the present work also, the construction of Neuhaus (1971) is applicable and therefore only the result will be stated here. A detailed discussion can be found in the above article.

Define the processes:

$$\begin{aligned}\tilde{F}_n(y, x) &= \sqrt{n} [F_n(y, x) - F(y, x)] \\ \tilde{C}_n(y) &= \sqrt{n} [C_n(y) - C(y)] \\ \tilde{W}_n(y, x) &= \sqrt{n} [F_{Y,X,n}^*(y, x) - F_{Y,X,n}^*(y, x)] \\ \tilde{L}_n(y) &= \sqrt{n} \tilde{L}_n(y).\end{aligned}$$

The scaled version of the representation given in Theorem 1 can now be rewritten in the following form, which rends the covariance structure more visible. Let

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

Then

$$\begin{aligned}\tilde{F}_n(y, x) &= \tilde{W}_n(y, x) A(y) - \int_0^y \tilde{W}_n(s, x) A(ds) \\ &\quad - \int_0^y \frac{A(s)}{C(s)} \tilde{C}_n(s) F_{Y,X}^*(ds, x) \\ &\quad - \int_0^y A(s) \tilde{L}_n(y) F_{Y,X}^*(ds, x) + R_n^*(y, x) \\ &\equiv \tilde{\xi}_n^*(y, x) + R_n^*(y, x).\end{aligned}\tag{5}$$

We now present the covariance functions of the above processes:

LEMMA 1.

- (i)  $Cov(\tilde{C}_n(u), \tilde{C}_n(v)) = C(u \wedge v) \frac{\bar{F}_Y(u \vee v)}{\bar{F}_Y(u \wedge v)} - C(u) C(v)$
- (ii)  $Cov(\tilde{L}_n(u), \tilde{L}_n(v)) = \int^{u \wedge v} \frac{F_Y(dz)}{C(z) \bar{F}(z)} \equiv b(u \wedge v)$



$$(iii) \quad Cov(\tilde{W}_n(u_1, u_2), \tilde{W}_n(v_1, v_2)) = F_{Y,X}^*(u_1 \wedge v_1, u_2 \wedge v_2) \\ - F_{Y,X}^*(u_1, u_2) F_{Y,X}^*(v_1, v_2)$$

$$(iv) \quad Cov(\tilde{C}_n(u), \tilde{L}_n(v)) = -\bar{F}_Y(u) \frac{F_Y(u \wedge v)}{\bar{F}_Y(u \wedge v)}$$

$$(v) \quad Cov(\tilde{C}_n(u), \tilde{W}_n(v, x)) = \frac{C(u)}{\bar{F}_Y(u)} [F(v, x) - F(u \wedge v, x)] \\ - C(u) F^*(v, x)$$

$$(vi) \quad Cov(\tilde{L}_n(u), \tilde{W}_n(v, x)) = \frac{1}{\bar{F}_Y(u \wedge v)} [F(u \wedge v, x) \\ - F(v, x) F(u \wedge v)].$$

It follows from standard results that  $\tilde{L}_n(y)$  and  $\tilde{C}_n(y)$ , converge weakly to mean zero Gaussian processes on  $D[0, b]$  with covariance structures given above. The weak convergence of  $\tilde{W}_n(y, x)$  to a mean zero two-time parameter Gaussian process on the complete separable metric space  $(D_2, d)$  defined on  $[0, 1] \times [0, 1]$  described in Neuhaus (1971) follows from the arguments in that article. We therefore have the following result which is immediate from Theorem 2, SLLN and functional LIL.

**COROLLARY 2.** *Under the assumptions of part (ii) of Theorem 2,*

(a) *For  $(y, x) \in T_b$*

$$F_n(y, x) \rightarrow F(y, x) \quad a.s.$$

(b)  $\sup_{(y, x) \in T_b} |F_n(y, x) - F(y, x)| = O((\log n/n)^{1/2}).$

(c) *Suppose the conditions of part (ii) of Theorem 3 hold. Then for  $(y, x) \in T_b$ ,  $\tilde{F}_n(y, x)$  converges weakly to a mean zero, two dimensional time Gaussian process on  $(D_2, d)$ .*

It is hard to give a compact form for the general covariance function of the above limiting process. For special cases, it can be obtained from Corollary 1 and Lemma 1. However for practical purposes an expression for the limiting variance would be essential. We therefore provide below the asymptotic variance of the process  $\tilde{F}_n(y, x)$ , the proof of which could be found in Gürler (1995).

COROLLARY 3. *The variance of the limiting process is given below, provided that the integrals appearing exist:*

$$\begin{aligned}\sigma^2(y, x) &\equiv \text{Var}(\bar{\xi}_n^*(y, x)) \\ &= \int^y A(u) F(du, x) - 2 \int^y [F(y, x) - F(u, x)] \left[ \frac{1}{C(u)} - b(u) \right] F(du, x).\end{aligned}$$

It can be checked that the above variance reduces to  $F(y, x)(1 - F(y, x))$  when there is no truncation. This expression allows us to make some standard inferences, such as hypothesis testing and construction of confidence intervals. For such applications an estimator of the variance is needed and we provide below a natural nonparametric estimator for it. However, particularly for the more complicated general covariance function, bootstrapping can be another option. We can estimate  $A(u)$  by

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

Observe also that the jump size of  $F_{Y,X,n}(u, v)$  at  $(Y_i, x)$  is

$$\frac{\bar{F}_{Y,n}(Y_i-)}{C_n(Y_i)} I(X_{[i]} \leq x) = A_n(Y_i) I(X_i \leq x)$$

Let

$$V_{1,n}(y, x) = n^{-1} \sum_{i: Y_i \leq y, X_i \leq x} A_n^2(Y_i)$$

and

$$\begin{aligned}V_{2,n}(y, x) &= n^{-1} \sum_{i: Y_i \leq y, X_i \leq x} A_n(Y_i) [F_{Y,X,n}(y, x) \\ &\quad - F_{Y,X,n}(Y_i, x)] [1/C_n(Y_i) - b_n(Y_i)],\end{aligned}$$

where

$$b_n(u) = \sum_{i=1}^n \frac{I(Y_i \leq u)}{C_n^2(Y_i)}.$$

Then, an estimator of the asymptotic variace can be given as

$$\sigma_n^2(y, x) = V_{1,n}(y, x) - 2V_{2,n}(y, x).$$

### 3. BIVARIATE DIVERSE HAZARD

#### 3.1. Characterization

In this section “diverse hazard” function will be presented and the identifiability of a bivariate distribution function via the diverse hazard will be discussed. This development is motivated by the discussion given in Dabrowska (1988) for the censored observations. For the univariate data, the d.f. is expressed in terms of the cumulative d.f. in a unique way. For the bivariate case however, there is not a single definition of the cumulative hazard function (see e.g. Marshall (1975), Cox (1972), Johnson and Kotz (1975)). Dabrowska (1988) provides a nice representation of a bivariate survival function in terms of her cumulative hazard function which is a vector of three components that correspond to double and single failures. In what follows, a different version of the bivariate hazard vector, namely the diverse-hazard will be presented.

We first introduce the following notation. For a bivariate function  $\phi(u, v)$ , which is left-continuous in the first, right-continuous in the second component, let

$$\phi(\delta u, v) = \phi(u+, v) - \phi(u, v)$$

$$\phi(u, \delta v) = \phi(u, v) - \phi(u, v-)$$

$$\phi(\delta u, \delta v) = \phi(u+, v) - \phi(u, v) - \phi(u+, v-) + \phi(u, v-)$$

and define the sets

$$E_1(\phi) = \{(u, v): \phi(\delta u, v) = \phi(u, \delta v) = 0\}$$

$$E_2(\phi) = \{(u, v): \phi(\delta u, v) \neq 0, \phi(\delta u, \delta v) = 0\}$$

$$E_3(\phi) = \{(u, v): \phi(u, \delta v) \neq 0, \phi(\delta u, \delta v) = 0\}$$

$$E_4(\phi) = \{(u, v): \phi(\delta u, \delta v) \neq 0\}.$$

In the following definitions, the superscripts will refer to the components for which the partial derivatives are taken. Let

$$\phi(du, v) = \begin{cases} \phi_1(u, v) & (u, v) \in E_1(\phi) \cup E_3(\phi) \\ \phi(\delta u, v) & (u, v) \in E_2(\phi) \cup E_4(\phi) \end{cases}.$$

$\phi(u, dv)$  is similarly defined, so that we have

$$\phi(du, dv) = \begin{cases} \phi^{1,2}(u, v) & (u, v) \in E_1(\phi) \\ \phi^2(\delta u, v) & (u, v) \in E_2(\phi) \\ \phi^1(u, \delta v) & (u, v) \in E_3(\phi) \\ \phi^{1,2}(\delta u, \delta v) & (u, v) \in E_4(\phi) \end{cases}.$$

To avoid introducing more notation, the integral of the above functions will be denoted by simply inserting the integral sign, since the distinction will be clear from the context. For example it will be understood that

$$\int_{E_s(\phi)} \phi(du, dv) = \int_u \sum_v \phi^1(u, \delta v) du.$$

Let  $\bar{F}(u, v) = P(Y \geq u, X \leq v)$ . We define the bivariate ‘diverse hazard’ vector  $\tilde{A}(u, v)$  as

$$\tilde{A}(u, v) = [A_{12}(du, dv), A_1(du, v), A_2(u, dv)],$$

where, with some abuse of the above notation for the ease of the following presentation, we define

$$A_{12}(du, dv) \equiv - \frac{\bar{F}(du, dv)}{\bar{F}(u, v)},$$

$$A_1(du, v) \equiv - \frac{\bar{F}(du, v)}{\bar{F}(v, v)},$$

and

$$A_2(u, dv) \equiv \frac{\bar{F}(u, dv)}{\bar{F}(u, v)}.$$

Note that the first member of  $\tilde{A}(u, v)$  corresponds to the failures of both components at  $(u, v-)$ , given that the first one is still alive at  $u-$ , while the second is known to have failed at  $v$ . In other words, it describes the conditional probability of double failures, the first in the immediate present and the second in the immediate past. The other two components have similar interpretations, which explains the term “diverse” hazard. This diverse-hazard vector is analogous to the bivariate hazard vector given in Dabrowska (1988) and following her lines, a bivariate distribution function will be presented in terms of this vector and the marginal distributions of  $Y$  and  $X$ .

Let

$$R(y, x) = \log \bar{F}(y, x).$$

Then we can write

$$\int_x^\infty \int_0^y R(du, dv) = \sum_{i=1}^4 R_i(u, v),$$

where

$$\begin{aligned}
 R_1(u, v) &= \int_v^\infty \int_0^u I[(s, t) \in E_1(R)] R(ds, dt) \\
 R_2(u, v) &= \int_v^\infty \sum_{s \leq u} I[(s, t) \in E_2(R)] R(ds, dt) \\
 R_3(u, v) &= \sum_{t > v} \int_0^u I[(s, t) \in E_3(R)] R(ds, dt) \\
 R_4(u, v) &= \sum_{t > v} \sum_{s \leq u} I[(s, t) \in E_4(R)] R(ds, dt).
 \end{aligned}$$

The following identities will be used to calculate  $R_i(u, v)$ ,  $i = 1, \dots, 4$ .

$$\begin{aligned}
 \frac{\bar{F}(u+, v)}{\bar{F}(u, v)} &= 1 - A_1(du, v) \\
 \frac{\bar{F}(u, v-)}{\bar{F}(u, v)} &= 1 - A_2(u, dv) \\
 \frac{\bar{F}(u+, v-)}{\bar{F}(u, v)} &= 1 - A_1(du, v) - A_2(u, dv) + A_{12}(du, dv).
 \end{aligned}$$

We then have

$$\begin{aligned}
 R_1(u, v) &= \int_v^\infty \int_0^u I[(s, t) \in E_1] [A_1(du, v) A_2(u, dv) - A_{12}(du, dv)] \\
 R_2(u, v) &= \int_v^\infty \sum_{s \leq u} \frac{I[(s, t) \in E_2]}{[1 - A_1(du, v)]} [A_1(du, v) A_2(u, dv) - A_{12}(du, dv)] \\
 R_3(u, v) &= \sum_{t > v} \int_0^u \frac{I[(s, t) \in E_3]}{[1 - A_2(u, dv)]} [A_1(du, v) A_2(u, dv) - A_{12}(du, dv)] \\
 R_4(u, v) &= \sum_{t > v} \sum_{s \leq u} I[(s, t) \in E_4] \\
 &\quad - \log \left[ 1 - \frac{A_1(du, v) A_2(u, dv) - A_{12}(du, dv)}{(1 - A_1(u, dv))(1 - A_2(u, dv))} \right].
 \end{aligned}$$

Define the function  $\Gamma(u, v)$  via

$$\Gamma(du, dv) = \frac{A_1(du, v) A_2(u, dv) - A_{12}(du, dv)}{[1 - A_1(du, v)][1 - A_2(u, dv)]}$$

and note that by the definition of the sets  $E_i(R)$ ,  $A_1(du, v) = 0$  for  $(u, v) \in E_1(R) \cup E_3(R)$  and  $A_2(u, dv) = 0$  for  $(u, v) \in E_1(R) \cup E_2(R)$ . Therefore we have the following unified representation.

**PROPOSITION.** *For  $(y, x)$  such that  $\bar{F}(y, x) > 0$ , it holds that*

$$\bar{F}(y, x) = \bar{F}_Y(y) F_X(x) \prod_{i=1}^4 A_i(y, x)$$

and

$$F(y, x) = F_X(x) \left\{ 1 - \bar{F}_Y(y) \prod_{i=1}^4 A_i(y+, x) \right\},$$

where for  $i = 1, 2, 3$ ,

$$A_i(y, x) = \exp \left\{ - \int_x^y \int I[(u, v) \in E_i] \Gamma(du, dv) \right\}$$

and

$$A_4(y, x) = \prod_{v > x} \prod_{u \leq y} [1 - \Gamma(du, dv)].$$

### 3.2. Identifiability and Applications with Truncated Data

The representation of the bivariate distribution function as a functional of the bivariate diverse-hazard is given in the previous section as a general result. For the truncation model, the identifiability of this hazard vector follows from the arguments below:

Define

$$C_2(y, x) = H_{T,X}(y, x) - F^*(y-, x) \quad (6)$$

$$= \alpha^{-1} G(y) \bar{F}(y, x) \quad (7)$$

and recall that

$$F_{Y,X}^*(y, x) \equiv H_{Y,X,T}(y, x, \infty) = \alpha^{-1} \int_0^x \int_0^y G(u) F(du, dv)$$

with

$$F_{Y,X}^*(dy, x) = \alpha^{-1} \int_0^x G(y) F(y, dv)$$

$$f_{Y,X}^*(y, x) = \alpha^{-1} G(y) f_{Y,X}(y, x).$$

Then we can write

$$\tilde{A}(y, x) = \left\{ \frac{F^*(du, dv)}{C_2(u, v)}, \frac{F^*(du, v)}{C_2(u, v)}, \frac{C_2(u, dv)}{C_2(u, v)} \right\}.$$

Note that the quantities above correspond to the observed random variables with the natural estimates

$$A_{12,n}(du, dv) = \frac{F_n^*(du, dv)}{C_{2,n}(u, v)}$$

$$A_{1,n}(du, v) = \frac{F_n^*(du, v)}{C_{2,n}(u, v)}$$

$$A_{2,n}(u, dv) = \frac{C_n(u, dv)}{C_{2,n}(u, v)},$$

where

$$nC_{2,n}(u, v) = \# \{i: T_i \leq u \leq Y_i, X_i \leq v\},$$

where  $nC_{2,n}(y, x)$  is the size of the risk set at  $(y, x)$  w.r.t. diverse-hazard set-up and  $F_n^*(y, x)$  is the empirical d.f. of observed  $(Y, X)$  pairs.

The representation of the proposition above could in principal be used to define alternative estimators for the bivariate distribution function by replacing the marginals and the bivariate hazard with their estimators. This was done in Dabrowska (1988) when both  $(Y, X)$  were censored. Such an estimator for the truncated observations would have a similar structure with that of her estimator. However this approach is not immediately available with the truncated data since there does not exist a consistent estimator in the literature for the marginal distribution of  $X$ . The  $X$  marginal of  $F_n(y, x)$  involves the integration of  $F_{Y,X,n}(y, x)$  w.r.t.  $y$  over an infinite region and this creates problems to establish the consistency. To remedy this situation, a smoothed version with a compact support kernel could be used, albeit at the cost of slower rates of convergence. We will not further pursue this idea here but suggest another possible application below. Note that from the proposition we have

$$Q(y, z) \equiv \frac{\bar{F}(y, x)}{\bar{F}_Y(y) F_X(x)} = \prod_{i=1}^4 A_i(y, x).$$

If  $X$  and  $Y$  are independent, the L.H.S. is unity, and we can estimate the R.H.S., which then could be used to test the independence of  $Y$  and  $X$ . Two alternative estimators for the L.H.S. are presented below.

If  $F(y, x)$  is continuous in both components we can write

$$\begin{aligned} Q(y, x) &= A_1(y, x) \\ &= \exp \left\{ - \left[ \int_x^\infty \int_0^y A_1(du, v) A_2(u, dv) - A_{12}(du, dv) \right] \right\} \\ &\equiv \exp \{ - A(y, x) \}. \end{aligned}$$

Similarly, for  $F(y, x)$  discrete in both components we have

$$\begin{aligned} Q(y, x) &= A_4(y, x) \\ &= \prod_{v > x} \prod_{u \leq y} [1 - \Gamma(du, dv)]. \end{aligned}$$

These two cases induce two possible estimators for  $\bar{F}(y, x)/\bar{F}_Y(y) F_X(x)$ ,

$$Q_{1,n}(y, x) = \exp \{ - A_n(y, x) \},$$

where

$$A_n(y, x) = \sum_{u \leq y} \sum_{v > x} [A_{1,n}(du, v) A_{2,n}(u, dv) - A_{12,n}(du, dv)] \quad (8)$$

and

$$Q_{2,n}(y, x) = \prod_{u \leq y; v > x} [1 - \Gamma_n(du, dv)]$$

with

$$\Gamma_n(du, dv) = \frac{A_{1,n}(du, v) A_{2,n}(u, dv) - A_{12,n}(du, dv)}{[1 - A_{1,n}(du, v)][1 - A_{2,n}(u, dv)]}.$$

The large sample equivalence of these two estimators are established in the next section.

### 3.3. Large Sample Results

LEMMA 2. Let  $a > a_{F_X}$   $b > b_{F_Y}$  be such that  $\bar{F}(b, a) > 0$  and  $T_{a,b} = \{(y, x): 0 < y \leq b, x \geq a\}$ . If  $\int F(dy)/G^2(y) < \infty$ , then

$$\begin{aligned} \text{(i)} \quad A_n(y, x) - A(y, x) &= \int_0^y \int_x^\infty \left\{ \frac{C_2^2(u, v) - C_{2,n}^2(u, v)}{C_2^4(u, v)} F^*(du, v) C(u, dv) \right. \\ &\quad \left. + \frac{C_2(u, v) - C_{2,n}(u, v)}{C_2^2(u, v)} F^*(u, v) \right\} + R_{1,n}(y, x) \\ &\equiv \bar{\xi}_{A,n}(y, x) + R_{1,n}(y, x) \end{aligned}$$

$$\text{(ii)} \quad Q_{1,n}(y, x) - Q(y, x) = -\bar{\xi}_{A,n}(y, x) + R_{2,n}(y, x),$$



where for  $i = 1, 2$

$$\sup_{(y, x) \in T_{a,b}} |R_{i,n}(y, x)| = O(\log^3 n/n).$$

LEMMA 3. *Let  $D_n(y, x) = Q_{2,n}(y, x) - Q_{1,n}(y, x)$ . Then, under the conditions of Lemma 1,*

$$\sup_{(y, x) \in T_{a,b}} |D_n(y, x)| = O(\log^3 n/n).$$

The weak convergence of  $Q_{i,n}(y, x)$  could be studied similar to the previous discussion relating to  $F_n(y, x)$  but will not be further elaborated here. This limiting distribution can then be used for hypothesis testing purposes. However such an application clearly requires more work in terms of assessing the asymptotic variance and investigating the power properties. These issues will be addressed elsewhere.

#### 4. SIMULATIONS AND CONCLUDING REMARKS

Large sample results for  $F_n(y, x)$  are presented in Section 2. Here, the results of a moderate size simulation study will be reported to provide some practical insight. As mentioned by a referee, in regression applications with censored data, correlation of the covariate and the censoring variables may create significant problems. To get an idea for the impact of such dependencies, a case for correlated  $(X, T)$  is also included in the simulations. Let  $U_1$ ,  $U_2$  and  $U_{12}$  be independent Exponential(exp) random variables with mean one. The following cases are simulated

- (i)  $(Y, X)$ : Independent Exponential with means one,  $T: \exp(\mu)$ , independent of  $(Y, X)$
- (ii)  $(Y, X)$ : Bivariate Exponential,  $Y = \min(U_1, U_{12})$ ;  $X = \min(U_2, U_{12})$   $T$  as in part (i).
- (iii)  $(Y, X)$  as in (i),  $(X, T)$ : Bivariate Exponential;  $T = \min(U_2, U_3)$  with  $U_3: \exp(\tau)$ ,  $X$ : as in (ii)
- (iv)  $(Y, X)$  as in (ii),  $(X, T)$ : as in (ii);  $T = \min(U_2, U_3)$ ,  $U_3: \exp(\tau)$

The parameters  $\tau$  and  $\mu$  are adjusted to obtain light, moderate and heavy truncation, with corresponding  $\alpha$  values of approximately .75 .50 and .25. The results are displayed in Figures 1 and 2, where the horizontal axis denotes the average proportion of observed (untruncated) samples.

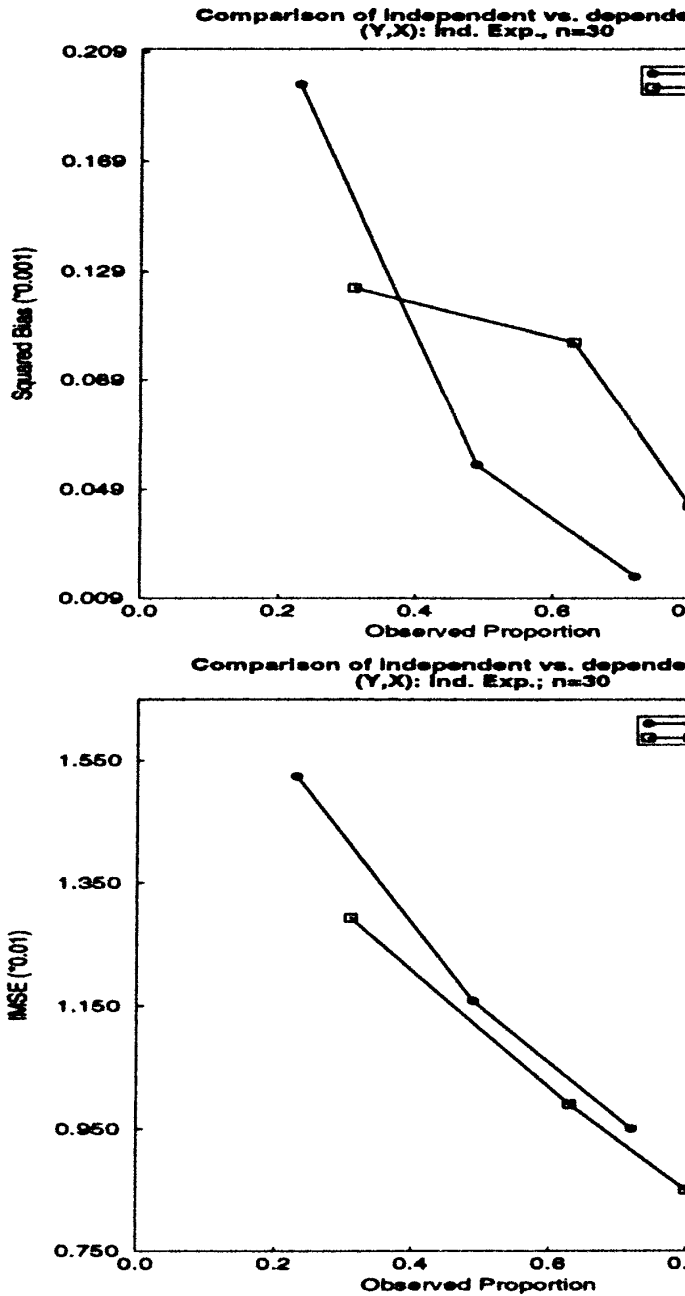


FIG. 1. Dependent vs independent (Y, X, T), n = 30.

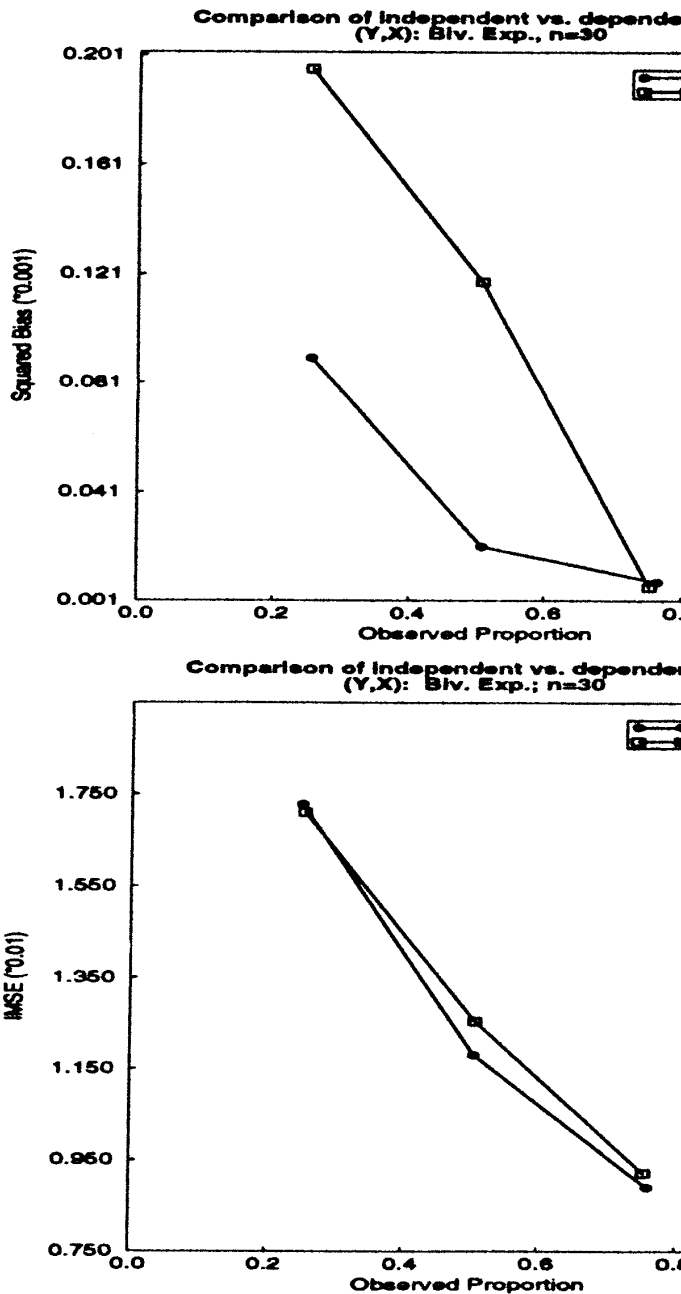


FIG. 1.—(Continued)

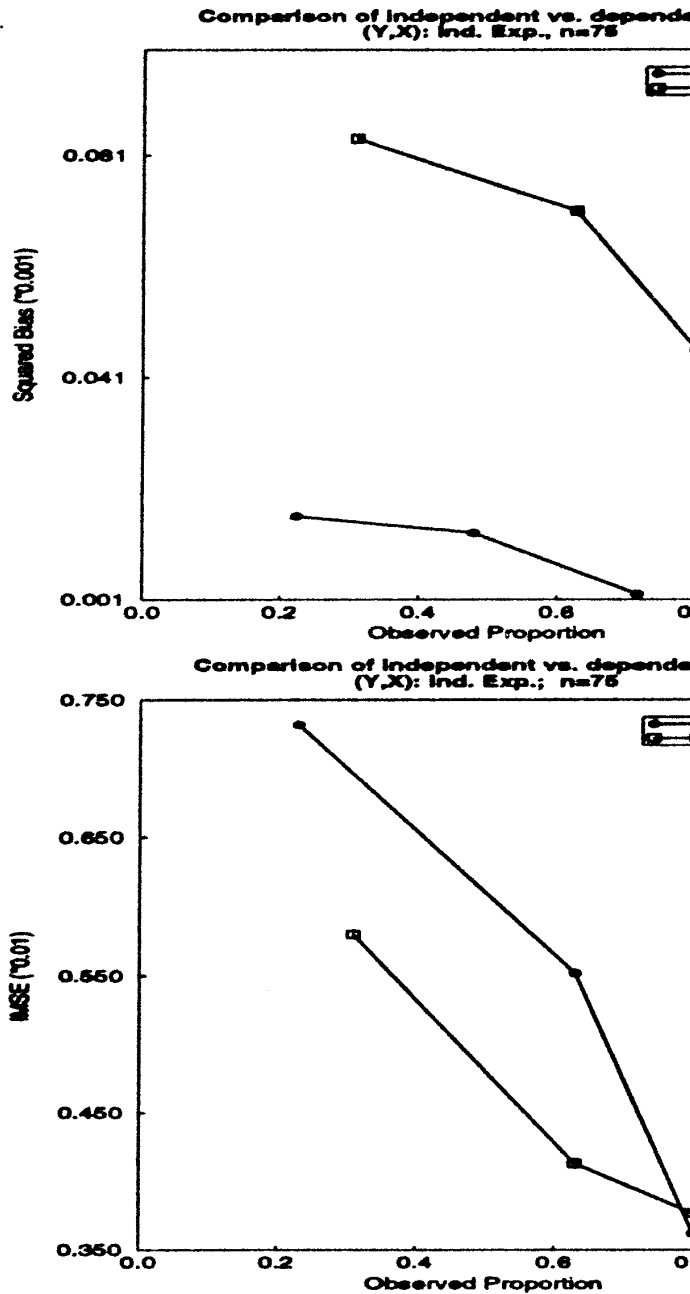


FIG. 2. Dependent vs independent (Y, X, T), n = 75.

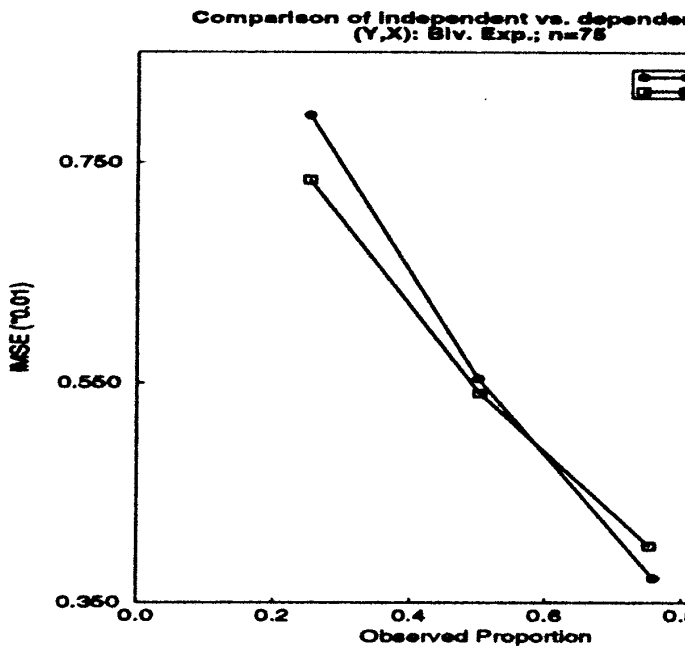
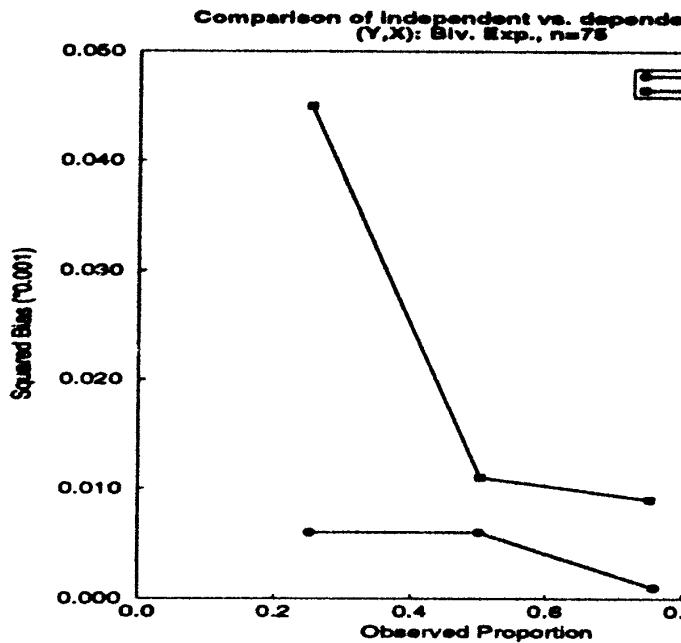


FIG. 2.—(Continued)

Samples of size  $n = 30, 75$  are used in 1000 replications. The performance measure was the estimated Integrated Mean Squared Error (IMSE). This integral is evaluated over a region covering about %95 of the total probability. The simulation results suggest that the mean contribution to the IMSE comes from the variance term, where the squared bias term in the worst case was in the order of  $10^{-4}$  which corresponds to  $\alpha = 0.25$  approximately. The results are displayed as the profile graphics to render an easier interpretation. From Figure 1 and 2, it is seen that IMSE of  $F_n(y, x)$  is slightly smaller for the independent  $(Y, X)$  case. The impact of correlated  $(X, T)$  pairs is seen more clearly on the bias terms where this correlation creates considerable bias in comparison the the independent case which is observed better in Figure 2 for  $n = 75$ . This difference however disappears in the IMSE since the contribution of the bias to this term is negligible. The increased bias for the correlated variables would of course seriously effect the estimates in the regression. These limited results already suggest that further research is needed to develop methods to handle the case where truncating variable is not independent (see also remark 4) of the bivariate vector  $(Y, X)$ . Both the bias and the variance reduce with increasing  $\alpha$  as would be expected. We conclude this section by pointing out some extensions of the proposed methods.

*Remarks.* (1) (Extension to  $k$  covariates). If we observe  $(Y_i, X_{1,i}, \dots, X_{k,i}, T_i)$ ,  $i = 1, \dots, n$  only if  $(T_i \leq Y_i)$ , an extension of  $F_n(\cdot)$  could be obtained by modifying the indicator function. Such an extension for the hazard representation of the multivariate d.f. is also possible but would be quite messy. Interested readers could refer to Dabrowska (1988).

(2) (Extension to double truncation). Let  $\bar{Y} = (Y_1, Y_2)$   $\bar{T} = (T_1, T_2)$  be random bivariate vectors and suppose we observe  $(\bar{Y}_i, \bar{T}_i)$  only if  $\bar{Y}_i \geq \bar{T}_i$  componentwise, for  $i = 1, \dots, n$ . The results of the present paper are not directly applicable for this truncation scheme due to further identifiability restrictions. This problem will be addressed elsewhere.

(3) (Right truncation). The estimators suggested in this paper could be extended for right truncation model in a natural way, with some modifications. For a related discussion see Gürlér (1996).

(4) (Correlated variables). In this paper it is assumed that the truncating variable  $T$  is independent of the vector  $(Y, X)$ . If this assumption is not true, there is not a straightforward generalization of the methods presented here, neither in the literature to the best of our knowledge. For the censored observations, Leurgans (1987) states that the correlation of the censoring variable and the covariate creates significant problems in the analysis of linear models. There she suggests grouping w.r.t. the covariate values, which clearly is also applicable to the present truncation set-up if the nature of the correlation suggests such a grouping. Another standard

assumption for truncation model is the independence of  $Y$  and  $T$ . For the regression problem with right truncated observations Kalbfleisch and Lawless (1991) relax this requirement somewhat by assuming that the response  $Y$  and  $T$  are conditionally independent given the covariate  $X$ . In particular, they assume that  $F(y | t, x) = F(y | x)/F(t | x)$  for  $y \leq t$ . However these approaches can only provide a partial solution to the general model of correlated  $(Y, X, T)$  and as mentioned earlier this area is open to more research.

## APPENDIX

For the proof of the lemmas and theorems of Section 3, the following lemmas will be useful:

LEMMA A1.

$$(i) \quad \sup_{y \leq b} \frac{(C_n(y) - C(y))^2}{C(y)} = O(\log n/n)$$

$$(ii) \quad \sup_{(y, x) \in T_{n,b}} \frac{(C_{2,n}(y, x) - C_2(y, x))^2}{C_2(y, x)} = O(\log n/n).$$

*Proof.* Part (i) is Lemma A2 of Chao and Lo (1988) and part (ii) could be obtained similarly.

LEMMA A2.

$$(i) \quad \sup_i \frac{C(Y_i)}{C_n(Y_i)} = O(\log n)$$

$$(ii) \quad \sup_i \frac{C_2(Y_i, X_i)}{C_{2,n}(Y_i, X_i)} = O(\log n).$$

*Proof.* Part (i) is Corollary 1.3 of Stute (1991) and for part (ii) a similar approach can be used, by defining the process,

$$\bar{H}_N(y, x, t) = N^{-1} \sum I(Y_i \leq y, X_i \leq x, T_i \leq t),$$

and showing that  $\bar{N}_N(y, x, t)/F(y, x)G(t)$  is a reverse submartingale.

*Proof of Theorem 2.* To simplify the notation, the arguments of  $\bar{F}_Y(u)$ ,  $\bar{F}_{Y,n}(u)$ ,  $C(u)$ , and  $C_n(u)$  will be suppressed and the notation  $\int_{y[x]}^x$  will be used to denote the double integral  $\int_0^y \int_0^x$ :

$$\begin{aligned}
& F_n(y, x) - F(y, x) \\
&= \int_{y]x]} \left\{ \frac{\bar{F}_{Y,n}}{C_n} F_n^*(du, dv) - \frac{\bar{F}_Y}{C} F^*(du, dv) \right\} \\
&= \int_{y]x]} \left\{ \frac{\bar{F}_Y}{C} (F_n^*(du, dv) - F^*(du, dv)) + \frac{\bar{F}_Y(C - C_n)}{C^2} F^*(du, dv) \right. \\
&\quad + \frac{\bar{F}_{Y,n} - \bar{F}_Y}{C} F_n^*(du, dv) + \frac{\bar{F}_Y(C - C_n)}{C^2} (F_n^*(du, dv) - F^*(du, dv)) \\
&\quad \left. + \frac{\bar{F}_Y(C - C_n)^2}{C^2 C_n} F_n^*(du, dv) + \frac{(\bar{F}_{Y,n} - \bar{F}_Y)(C - C_n)}{C C_n} F_n^*(du, dv) \right\} \\
&\equiv I + II + III + R_{2,n} + R_{3,n} + R_{4,n}.
\end{aligned}$$

From the representation in Theorem 1, the term *III* above is written as

$$\begin{aligned}
& \int_{y]x]} \left\{ \frac{\bar{F}_Y \bar{L}_n}{C} F^*(du, dv) + \frac{\bar{F}_{Y,n} - \bar{F}_Y}{C} (F_n^*(du, dv) - F^*(du, dv)) \right\} + \varepsilon'_n(y) \\
&\equiv \int_{y]x]} \frac{\bar{F}_Y \bar{L}_n}{C} F^*(du, dv) + R_{1,n} + \varepsilon'_n(y),
\end{aligned}$$

where

$$\sup_{(0 < y < b)} |\varepsilon'_n(y)| = O(|\varepsilon_n(y)|).$$

Hence after evaluating the above integrals w.r.t.  $v$ , we obtain the representation of Theorem 2, with

$$R_n(y, x) = \sum_{i=1}^4 R_{i,n}(y, x) + O(|\varepsilon_n(y)|).$$

The orders of  $R_{i,n}$ , ( $i=1, \dots, 4$ ) can now be obtained from Theorem 1, Lemma A1 (i), Lemma A2 (i) and the following facts:

(i) For  $a_G \leq a_{F_Y}$ ,  $b_G \leq b_{F_Y}$ ,

$$\sup_{0 \leq y < \infty} |C_n(y) - C(y)| = O((\log n/n)^{1/2})$$



since  $C_n$  is a difference of empirical d.f.'s.

(ii) For  $(y, x) \in [0, \infty) \times [0, \infty)$ ,

$$\sup |F_n^*(y, x) - F^*(y, x)| = O(\log n/n)^{1/2}.$$

(iii) For  $(y, x) \in [0, \infty) \times [0, \infty)$ ,

$$\int_{y]x]} \frac{\bar{F}}{C^2} F^*(du, dv) \leq \alpha \int_0^y \frac{F_Y(du)}{G(u)}.$$

As an illustration, consider

$$R_{3,n}(y, x) = \int^y \frac{\bar{F}_Y(C - C_n)^2}{C^2 C_n} F_n^*(du, x).$$

We have

$$\begin{aligned} \sup |R_{3,n}(y, x)| &\leq \sup \frac{(C - C_n)^2}{|C|} \int^y \frac{\bar{F}_Y}{C \cdot C_n} F_n^*(du, x) \\ &\leq O(\log n/n) \sup_i \frac{C(Y_i)}{C_n(Y_i)} \int^y \frac{\bar{F}_Y}{C^2} F_n^*(du, x) \\ &= O(\log n/n) O(\log n) O(1) = O(\log^2 n/n), \end{aligned}$$

by Lemma A1(i) and A2(i).

*Proof of Lemma 2.* In the following proof, the double integral  $\int_0^y \int_x^\infty$  will be denoted by  $\int_{y]x]}$  and the arguments of  $C_2(y, x)$  and its empirical counterpart will be dropped to avoid a messy presentation.

(i) The representation is easily obtained with the remainder term

$$\begin{aligned} R_{1,n}(y, x) &= \int_{y]x]} \left\{ \frac{C_2 - C_{2,n}}{C_2^2} [F_n^*(du, dv) - F^*(du, dv)] \right. \\ &\quad + \frac{(C_2 - C_{2,n})^2}{C_2^2 C_{2,n}} F_n^*(du, dv) \\ &\quad + \frac{C_2^2 - C_{2,n}^2}{C_2^4} [F_n^*(du, v) C_{2,n}(u, dv) - F^*(du, v) C_2(u, dv)] \\ &\quad \left. + \frac{(C_2^2 - C_{2,n}^2)^2}{C_2^4 C_{2,n}^2} F_n^*(du, v) C_{2,n}(u, dv) \right\} \\ &\equiv \sum_{i=1}^4 R_{1,i}(y, x), \end{aligned} \tag{9}$$

where, for  $(y, x) \in T_{a,b}$ , it follows from L.I.L. and S.L.L.N. that

$$\sup |R_{1,1}(y, x)| = O(\log n/n).$$

From Lemma A1 and A2 parts (ii), we have

$$\sup |R_{1,2}(y, x)| = O(\log^2 n/n)$$

and

$$\begin{aligned} \sup |R_{1,3}(y, x)| &\leq \text{cons.} \sup |C_2 - C_{2,n}| \\ &\quad \times \left| \int_{y] [x} \frac{[F_n^*(du, v) C_{2,n}(u, dv) - F^*(du, v) C_2(u, dv)]}{C_2^4} \right| \\ &\leq O((\log n/n)^{1/2}) O((\log n/n)^{1/2}). \end{aligned}$$

The last inequality above is obtained by applying the LIL to the sum of i.i.d. variables represented by the integral there and the relations

$$F^*(du, v) = \alpha^{-1} G(u) F(du, v)$$

and

$$C_2(u, dv) = \alpha^{-1} [G(u) F_X(dv) - G(u) F(u, dv)];$$

hence,

$$\begin{aligned} &\int_{y] [x} C_2^{-4} F^*(du, v) C_2(u, dv) \\ &= \int_{y] [x} \alpha \frac{F(du, v) F_X(dv)}{G(u) \bar{F}^3(u, v)} - \int_{y] [x} \alpha \frac{F(du, v) F(u, dv)}{G(u) \bar{F}^3(u, v)} \end{aligned}$$

and it is in absolute value less than

$$\alpha \bar{F}^{-4}(a, b) \int F_Y(du)/G^2(u) = O(1).$$

For  $R_{1,4}(y, x)$ , note that by Lemma A2 (ii), and A1 (ii)

$$\begin{aligned} \frac{(C_2^2 - C_{2,n}^2)^2}{C_2^4 C_{2,n}^2} &\leq \sup \left| \frac{(C_2 - C_{2,n})^2}{C_2} \cdot \frac{(C_2 + C_{2,n})^2}{C_2^3 C_{2,n}^2} \right| \\ &\leq O(\log n/n) \cdot 2 \max \left( \frac{1}{C_2^3}, \frac{1}{C_{2,n}^2 C_2} \right) \\ &= O(\log n/n) O(\log^2 n) \frac{1}{C_2^3} \\ &= O(\log^3 n/n) \frac{1}{C_2^3}. \end{aligned}$$

The bound for  $R_{1,4}(y, x)$  is now obtained from the SLLN, since the remaining integral represents a sum of i.i.d. random variables with finite mean, bounded similar to term  $R_{1,3}(y, x)$  (ii) Follows from part (i)

*Proof of Lemma 3.* After some algebra we can write

$$\begin{aligned}
|D_n(y, x)| &= |Q_{2,n}(y, x) - Q_{1,n}(y, x)| \\
&\leq |\log Q_{2,n}(y, x) + A_n(y, x)| \\
&= \left| \sum_{u \leq y} \sum_{v > x} \log[1 - \Gamma_n(\delta u, \delta v)] \right. \\
&\quad \left. + A_{1,n}(\delta u, v) A_{2,n}(u, \delta v) - A_{12,n}(\delta, \delta v) \right| \\
&= \left| \sum_{u \leq y} \sum_{v > x} \log \left( \frac{C_{2,n}(u, v-) C_{2,n}(u, v) + C_{2,n}(u, v-) F_n^*(\delta u, v)}{C_{2,n}(u, v-) C_{2,n}(u, v) + C_{2,n}(u, v) F_n^*(\delta u, v-)} \right) \right. \\
&\quad \left. - \left( \frac{C_{2,n}(u, v-) F_n^*(\delta u, v) - C_{2,n}(u, v) F_n^*(\delta u, v-)}{C_{2,n}^2(u, v)} \right) \right|.
\end{aligned}$$

Applying a two term Taylor expansion to the logarithm term and rearranging the terms, the above expression can be written (up to a term of a smaller order) as

$$\begin{aligned}
&\left| \sum_{u \leq y} \sum_{v > x} C_{2,n}^{-2}(u, v) F_n^*(\delta u, \delta v) [F_n^*(\delta u, v) + F_n^*(\delta u, v-) + C_{2,n}(u, \delta v)] \right. \\
&\quad \left. - \frac{F_n^*(\delta u, v-) C_{2,n}(u, \delta v)}{C_{2,n}^2(u, v) C_{2,n}(u, v-)} [C_{2,n}(u, \delta v) + F_n^*(\delta u, v-)] \right|.
\end{aligned}$$

Now observing that

$$\frac{F_n^*(\delta u, v-)}{C_{2,n}(u, v-)} \leq \frac{F_{Y,n}(\delta u)}{C_n(u)},$$

and using Lemma A1, we obtain

$$\begin{aligned}
\sup_{(y, x) \in T_{a,b}} |D_n(y, x)| &\leq O(\log^2 n/n) \int \frac{F_n^*(\delta u, \delta v)}{C_2(u, v)} \\
&\quad + O(\log^3 n/n) \int \frac{F_{Y,n}(\delta u) C_{2,n}(u, \delta v)}{C_2^3(u, v)} \\
&= O(\log^3 n/n).
\end{aligned}$$

## REFERENCES

- [1] Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika* **68** 417–442.
- [2] Cox, D. R. (1972). Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220.
- [3] Chao, M. T., and Lo, S. H. (1968). Some representations of the nonparametric maximum likelihood estimators with truncated data. *Ann. Statist.* **16** 661–668.
- [4] Dabrowska, D. M. (1988). Kaplan–Meier estimate on the plane. *Ann. Statist.* **16** 1485–1489.
- [5] 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.
- [6] Gross, S. T., and Huber-Carol, C. (1992). Regression models for truncated survival data. *Scand. J. Statist.* **19** 193–213.
- [7] Gu, M. G., and Lai, T. L. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann. Probab.* **18** 160–189.
- [8] Gürlér, Ü. (1996). Bivariate estimation with right truncated data. *J. Amer. Statist. Assoc.* **91** 1152–1165.
- [9] Gürlér, Ü. (1995). “Variance of the estimator for the bivariate distribution function when a component is truncated,” IEOR-9515, Bilkent University, Dept. of Indust. Eng., 06533 Ankara, Turkey.
- [10] Gürlér, Ü., Stute, W., and Wang, J.-L. (1993). Weak and strong quantile representations for randomly truncated data with applications. *Statist. Probab. Lett.* **17** 139–148.
- [11] Kalbfleisch, J. D., and Lawless, J. F. (1989). Inference based on retrospective ascertainment: An analysis of the data on transfusion-related AIDS. *J. Amer. Statist. Assoc.* **84** 360–372.
- [12] Lai, T. L., and Ying, Z. (1991). Estimating a distribution function with truncated and censored data. *Ann. Statist.* **19** 417–442.
- [13] Leurgans, S. (1987). Linear models, random censoring and synthetic data. *Biometrika* **74** 301–309.
- [14] Lo, S. H., and Wang, J. L. (1989). I.I.D. representations for the bivariate product limit estimators and the bootstrap versions. *J. Multivariate Anal.* **28** 211–226.
- [15] Lynden-Bell, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Mon. Nat. Roy. Astr. Soc.* **155** 95–118.
- [16] Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.
- [17] Stute, W. (1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **21** 146–156.
- [18] Stute, W. (1993). Consistent estimation under random censorship when covariables are present. *J. Multivariate Anal.* **45** 89–103.
- [19] Uzunogullari, Ü., and Wang, J.-L. (1992). A comparison of the hazard rate estimators for left truncated and right censored data. *Biometrika* **79** 297–310.
- [20] van der Laan, M. J. (1992). “Efficient estimator of the bivariate survival function for right censored data,” Tech. rep., Department of Mathematics, University of Utrecht, The Netherlands.
- [21] Wang, M. C., Jewell, N. P., and Tsai, W. Y. (1986). Asymptotic properties of the product limit estimate under random truncation. *Ann. Statist.* **14** 1597–1605.
- [21] Wang, M. C., Jewell, N. P., and Tsai, W. Y. (1986). Asymptotic properties of the product limit estimate under random truncation. *Ann. Statist.* **14** 1597–1605.
- [22] Woodroffe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13** 163–177.