Statistica Sinica 8(1998), 1219-1232

COVARIANCE FUNCTION OF A BIVARIATE DISTRIBUTION FUNCTION ESTIMATOR FOR LEFT TRUNCATED AND RIGHT CENSORED DATA

Irène Gijbels and Ülkü Gürler

Catholic University of Louvain and Bilkent University

Abstract: In left truncation and right censoring models one observes i.i.d. samples from the triplet (T, Z, δ) only if $T \leq Z$, where $Z = \min(Y, C)$ and δ is one if $Z = Y$ and zero otherwise. Here, Y is the variable of interest, T is the truncating variable and C is the censoring variable. Recently, Gürler and Gijbels (1996) proposed a nonparametric estimator for the bivariate distribution function when one of the components is subject to left truncation and right censoring. An asymptotic representation of this estimator as a mean of i.i.d. random variables with a negligible remainder term has been developed. This result establishes the convergence to a two time parameter Gaussian process. The covariance structure of the limiting process is quite complicated however, and is derived in this paper. We also consider the special case of censoring only. In this case the general expression for the variance function reduces to a simpler formula.

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

1. Introduction

In survival or reliability studies, the observed data is typically censored and/or truncated. Left truncation and right censoring (LTRC) together naturally occur in cohort follow-up studies. In a recent work, Gürler and Gijbels (1996) propose an estimator of the bivariate distribution function $F(y, x)$ of (Y, X) when the component Y is subject to LTRC. The variable of interest is the lifetime variable Y, but for several reasons one can observe samples of the random vector (T, Z, δ), only 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 assumed to be independent of (Y, X). Their distribution functions are denoted by G and H respectively. Let V_Z denote the distribution function of Z. Then $V_Z = 1 - (1 - F_Y)(1 - H)$, with F_Y being the marginal distribution function of Y. Without loss of generality we assume that all the random variables are nonnegative. The bivariate distribution function $F(y, x)$ is identifiable only if $a_G \leq a_{V_Z}$ and $b_G \leq b_{V_Z}$, where we denote $a_L = \inf\{t : L(t) > 0\}$ and $b_L = \inf\{t : L(t) = 1\}$ for any distribution function L.

This condition is similar to the one stated in Woodroofe (1985) for the univariate left truncated model.

Suppose 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 $W^1_{Z,X}(z,x)$ denote the bivariate sub-distribution function of the observed uncensored variables, i.e.

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

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

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

and

$$W^1_Z(z) = \alpha^{-1} \int_0^z G(u) \bar{H}(u-) F_Y(du), \tag{2}$$

where for any distribution function L we denote $\bar{L}(u) = 1 - L(u)$. Then,

$$W^1_{Z,X}(dz, dx) = \alpha^{-1} G(z) \bar{H}(z-) F(dz, dx), \tag{3}$$

which has the following marginal for Z,

$$W^1_Z(dz) = \alpha^{-1} G(z) \bar{H}(z-) F_Y(dz).$$

Denote by

$$W_T(t) = P(T \leq t | T \leq Z) \qquad \text{and} \qquad W_Z(z) = P(Z \leq z | T \leq Z)$$

the distribution function of the observed random variables T and Z respectively. Define

$$C(u) = P(T \leq u \leq Z | T \leq Z) = W_T(u) - W_Z(u-),$$

and note that

$$C(u) = \alpha^{-1} G(u) \bar{F}_Y(u-) \bar{H}(u-). \tag{4}$$

From (3) and (4) it follows that

$$F(dy, dx) = \frac{\bar{F}_Y(y-)}{C(y)} W^1_{Z,X}(dy, dx) \equiv A(y) W^1_{Z,X}(dy, dx). \tag{5}$$

Relation (5) motivates the following estimator for $F(y,x)$:

$$F_n(y,x) = \frac{1}{n} \sum_{i=1}^n \frac{\bar{F}_{Y,n}(Z_i-)}{C_n(Z_i)} I(Z_i \leq y, X_i \leq x, \delta_i = 1), \tag{6}$$

where

$$\bar{F}_{Y,n}(y) = \prod_{i:Z_i \le y} \Big[1 - \frac{s(Y_i)}{nC_n(Y_i)}\Big]^{\delta_i} \qquad \text{with} \qquad nC_n(u) = \#\{i : T_i \le u \le Z_i\},$$

and for $u > 0$, $s(u) = \#\{i : Z_i = u\}$. The estimator $F_n(y, x)$ is a bivariate distribution function and reduces to the univariate product limit estimator when $x \to \infty$.

Define

$$L_i(z) = \frac{I(Z_i \le z, \delta_i = 1)}{C(Z_i)} - \int_0^z \frac{I(T_i \le u \le Z_i)}{C^2(u)} W_Z^1(du) \text{ and } \bar{L}_n(z) = \frac{1}{n}\sum_{i=1}^n L_i(z).$$

Let

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

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

The following theorem of Gürler and Gijbels (1996) provides a strong i.i.d. representation for the estimator $F_n(y, x)$ given above. Such a representation for the univariate LTRC data was established in Gijbels and Wang (1993).

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:*

$$\begin{aligned} F_n(y, x) - F(y, x) &= \int_0^y A(u)[W^1_{Z,X,n}(du, x) - W^1_{Z,X}(du, x)] \\ &\quad - \int_0^y \Big\{\frac{A(u)}{C(u)}[C_n(u) - C(u)] + A(u)\bar{L}_n(u)\Big\} W^1_{Z,X}(du, x) \\ &\quad + R_n(y, x) \equiv \bar{\xi}_n(y, x) + R_n(y, x) \end{aligned} \tag{7}$$

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

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

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

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

The covariance structure of the limiting process is quite complicated, particularly due to both truncation and censoring effect. However, for the right truncation model with no censoring, this function takes a somewhat simpler form and is presented in Gürler (1996). In this paper, we first provide in Section 2 the covariances of the main processes involved. Using these we can then derive the general expression for the covariance function in Section 3. This expression is quite complicated and the special case of no truncation is treated separately. Outlines of the proofs of the presented results are deferred to Section 4. For details of the proofs see the technical report by Gijbels and Gürler (1996).

2. Covariances of the Main Processes

In this section we derive the covariance structure of the main processes involved in expression (7). Define the processes:

$$\tilde{F}_n(y,x) = \sqrt{n}[F_n(y,x) - F(y,x)] \qquad \tilde{W}_n(y,x) = \sqrt{n}[W^1_{Z,X,n}(y,x) - W^1_{Z,X}(y,x)]$$
$$\tilde{C}_n(y) = \sqrt{n}[C_n(y) - C(y)] \qquad \tilde{L}_n(y) = \sqrt{n}\bar{L}_n(y).$$

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

$$\begin{aligned}\tilde{F}_n(y,x) &= \tilde{W}_n(y,x)A(y) - \int_0^y \tilde{W}_n(s,x)A(ds)\\ &\quad - \int_0^y \frac{A(s)}{C(s)}\tilde{C}_n(s)W^1_{Z,X}(ds,x) - \int_0^y A(s)\tilde{L}_n(s)W^1_{Z,X}(ds,x) + R^*_n(y,x)\\ &\equiv \bar{\xi}^*_n(y,x) + R^*_n(y,x).\end{aligned}$$

We present below the covariance functions of the processes $\tilde{C}_n(y)$, $\tilde{W}_n(y,x)$ and $\tilde{L}_n(y)$, from which that of $\bar{\xi}^*_n(y,x)$ can be calculated. We first introduce some further notation. Let

$$a_1(t,x) = \int_0^t \frac{W^1_{Z,X}(dv,x)}{G(v)} \qquad b(t) = \int_0^t \frac{W^1_Z(dv)}{C^2(v)}$$
$$a_2(t,x) = \int_0^t \frac{G(v)}{C(v)}F(dv,x) \qquad h(t) = \int_0^t \frac{G(v)}{C^2(v)}W^1_Z(dv)$$
$$b_1(t,x) = \int_0^t \frac{W^1_{Z,X}(dv,x)}{C(v)} \qquad d(u,v,x) = \int_0^{u\wedge v}[a_1(u,x) - a_1(s,x)]h(ds).$$

Lemma 1. *Suppose* $\int F_Y(du)/G(u) < \infty$. *Then*

(i) $$\mathrm{Cov}\,(\tilde{C}_n(u),\tilde{C}_n(v)) = C(u\vee v)\frac{G(u\wedge v)}{G(u\vee v)} - C(u)C(v)$$

(ii) $$\mathrm{Cov}\,(\tilde{L}_n(u),\tilde{L}_n(v)) = \int_0^{u\wedge v}\frac{W^1_Z(dz)}{C^2(z)}$$

(iii) $$\mathrm{Cov}\,(\tilde{W}_n(u_1,u_2),\tilde{W}_n(v_1,v_2)) = W^1_{Z,X}(u_1\wedge v_1, u_2\wedge v_2) - W^1_{Z,X}(u_1,u_2)W^1_{Z,X}(v_1,v_2)$$

(iv) $$\mathrm{Cov}\,(\tilde{C}_n(u), \tilde{L}_n(v)) = -\frac{C(u)}{G(u)} \int_0^{u\wedge v} \frac{G(z)}{C^2(z)} W_Z^1(dz) = -\frac{C(u)}{G(u)} h(u\wedge v)$$

(v) $$\begin{aligned}\mathrm{Cov}\,(\tilde{C}_n(u), \tilde{W}_n(v,x)) &= G(u) \int_{u\wedge v}^{v} \frac{W_{Z,X}^1(dz,x)}{G(z)} - C(u) W_{Z,X}^1(v,x) \\ &= G(u)[a_1(v,x) - a_1(u\wedge v, x)] - C(u) W_{Z,X}^1(v,x)\end{aligned}$$

(vi) $$\begin{aligned}\mathrm{Cov}\,(\tilde{L}_n(u), \tilde{W}_n(v,x)) &= \int_0^{u\wedge v} \frac{W_{Z,X}^1(dz,x)}{C(z)} \\ &\quad - \int_0^{u\wedge v} \frac{G(s)}{C^2(s)} \int_s^v \frac{W_{Z,X}^1(dz,x)}{G(z)} W_Z^1(ds) \\ &= b_1(u\wedge v, x) - \int_0^{u\wedge v} [a_1(v,x) - a_1(s,x)] h(ds).\end{aligned}$$

The concise proofs of items (i), (iv) and (v) of Lemma 1 are provided in Section 4.1. The proofs of the other items are quite similar and are not given. Further details of the proof can be found in Gijbels and Gürler (1996).

3. Covariance of the Bivariate Distribution Function Estimator

Starting from the covariance structures provided in Lemma 1 of Section 2, we can now derive the covariance function for the bivariate estimator defined in (6). Since $A(z)W_{Z,X}^1(dz,x) = F(dz,x)$, we can write

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

Also, $E[\bar{\xi}_n^*(y,x)] = 0$, implies

$$\mathrm{Cov}\,(y_1,y_2,x_1,x_2) \equiv \mathrm{Cov}\,(\bar{\xi}_n^*(y_1,x_1), \bar{\xi}_n^*(y_2,x_2)) = E[\bar{\xi}_n^*(y_1,x_1)\bar{\xi}_n^*(y_2,x_2)].$$

In order to give the expression for the covariance function we need some further notation: Let

$$\begin{aligned}\mathcal{T}_1(y_1,y_2,x_1,x_2) &= -a_2(y_2,x_2)\int_{y_1\wedge y_2}^{y_1} A(u)a_1(du,x_1) - \int_0^{y_1\wedge y_2} A(u)a_2(u,x_2)a_1(du,x_1) \\ \mathcal{T}_2(y_1,y_2,x_1,x_2) &= -\int_0^{y_1\wedge y_2} [F(y_2,x_2) - F(u,x_2)]A(u)b_1(du,x_1) \\ \mathcal{T}_3(y_1,y_2,x_1,x_2) &= \int_0^{y_1}\int_0^{y_2} [d(y_1,v,x_1) - d(u,v,x_1)]A(du)F(dv,x_2) \\ \mathcal{T}(y_1,y_2,x_1,x_2) &= \mathcal{T}_1(y_1,y_2,x_1,x_2) + \mathcal{T}_2(y_1,y_2,x_1,x_2) + \mathcal{T}_3(y_1,y_2,x_1,x_2)\end{aligned}$$

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

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

Theorem 2. *Suppose* $\int F_Y(du)/G(u) < \infty$. *Then*

$$\begin{aligned}\mathrm{Cov}\,(y_1,y_2,x_1,x_2) = &\int_0^{y_1\wedge y_2} A^2(u)W^1_{Z,X}(du,x_1\wedge x_2)\\ &+\mathcal{T}(y_1,y_2,x_1,x_2)+\mathcal{T}(y_2,y_1,x_2,x_1)+\mathcal{X}(y_1,y_2,x_1,x_2). \quad (8)\end{aligned}$$

Proof. See Section 4.2.

For applications, the variance function of the bivariate distribution function estimator (6) is of special interest. We therefore explicitly present it below.

Corollary 1. *Under the condition of Theorem* 2*,*

$$\mathrm{Var}\,(y,x) = \mathrm{Cov}\,(y,y,x,x) = \int_0^y A^2(u)W^1_{Z,X}(du,x) + 2\mathcal{T}(y,x) + \mathcal{X}(y,x), \quad (9)$$

where

$$\mathcal{X}(y,x) = \int_0^y\int_0^y K(u,v)F(du,x)F(dv,x),$$

and

$$\begin{aligned}\mathcal{T}(y,x) = &-\int_0^y A(u)a_2(u,x)a_1(du,x) - \int_0^y [F(y,x)-F(u,x)]A(u)b_1(du,x)\\ &+\int_0^y\int_0^y [d(y,v,x)-d(u,v,x)]A(du)F(dv,x).\end{aligned}$$

The variance function in (9) should of course reduce to the variance function found for the censoring only case. In the special case of no truncation, $\alpha = 1$ and $G(x) = 1$ for all x. This leads to simplifications of all quantities involved. Straightforward calculations yield the result in Corollary 2. Note that if there is only censoring, the integrability condition of Theorem 2 always holds.

Corollary 2. *For the right censoring model*

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

This expression is similar to the expression obtained by Gürler (1997) in the case of truncation only (see Corollary 3 in that paper) with appropriate replacements for the definitions of the quantities $A(u)$, $C(v)$ and $b(v)$.

In the case of no censoring and no truncation, we have in addition to the previous simplifications that $H(x) = 0$ for all finite x. Straightforward calculations lead to the well-known expression for the variance function. (See Gijbels and Gürler (1996) for details.)

4. Proofs

4.1. Proof of Lemma 1

(i). Denoting $C_i(u) = I(T_i \le u \le Z_i)$ we can write

$$\begin{aligned}\mathrm{Cov}\,(\tilde{C}_n(u), \tilde{C}_n(v)) &= E\,[C_i(u)C_i(v)] - C(u)C(v)\\ &= P\,(T \le u \wedge v, Z \ge u \vee v | T \le Z) - C(u)C(v)\\ &= C(u \vee v)\frac{G(u \wedge v)}{G(u \vee v)} - C(u)C(v).\end{aligned}$$

(iv). Since $E[\tilde{L}_n(v)] = \sqrt{n}E[\bar{L}_n(v)] = 0$ and $E[\tilde{C}_n(u)] = 0$ we have

$$\begin{aligned}\mathrm{Cov}\,(\tilde{C}_n(u), \tilde{L}_n(v)) &= E\left[\tilde{C}_n(u)\tilde{L}_n(v)\right] = E\,[C_i(u)L_i(v)]\\ &= \int_u^v \int_0^u \frac{1}{C(z)} W^1_{Z,T}(dz, dt)\\ &\quad - \int_0^v P(T \le u \wedge t, Z \ge u \vee t | T \le Z)\frac{1}{C^2(t)} W^1_Z(dt)\\ &= (I) - (II). \end{aligned} \qquad (10)$$

We deal with these two terms separately. From (1) and (4) it is easily obtained that

$$\begin{aligned}(I) &= \int_u^v \int_0^u \frac{1}{\alpha^{-1}G(z)\bar{F}_Y(z-)\bar{H}(z-)} \alpha^{-1}\bar{H}(z-)F_Y(dz)G(dt)\\ &= G(u)\int_u^v \frac{1}{G(z)\bar{F}_Y(z-)} F_Y(dz), \end{aligned} \qquad (11)$$

provided $u < v$. When $u \ge v$ it is obvious from (10) that $(I) = 0$. For the second term in expression (10) note that $P(T \le u \wedge t, Z \ge u \vee t | T \le Z) = \alpha^{-1}G(u \wedge t)\bar{H}((u \vee t)-)\bar{F}_Y((u \vee t)-)$ and therefore

$$(II) = \int_0^v \frac{\alpha^{-1}G(u \wedge t)\bar{H}((u \vee t)-)\bar{F}_Y((u \vee t)-)}{C^2(t)} W^1_Z(dt). \qquad (12)$$

For $u < v$ this leads to

$$(II) = \int_0^u \frac{\alpha^{-1}G(t)\bar{H}(u-)\bar{F}_Y(u-)}{C^2(t)} W^1_Z(dt) + \int_u^v \frac{\alpha^{-1}G(u)\bar{H}(t-)\bar{F}_Y(t-)}{C^2(t)} W^1_Z(dt), \qquad (13)$$

where the first term in the above expression equals $[C(u)/G(u)]\int_0^u[G(t)/C^2(t)]\, W_Z^1(dt)$. Using (2) it is easily seen that the second term in (13) can be written as follows

$$\alpha^{-1}G(u)\int_u^v \frac{\bar{H}(t-)\bar{F}_Y(t-)\alpha^{-1}G(t)\bar{H}(t-)}{C^2(t)}F_Y(dt)=G(u)\int_u^v \frac{1}{G(t)\bar{F}_Y(t-)}F_Y(dt). \tag{14}$$

Combining (10), (11), (13) and (14) we get that for the case $u < v$

$$\mathrm{Cov}\,(\tilde{C}_n(u), \tilde{L}_n(v)) = -\frac{C(u)}{G(u)}\int_0^u \frac{G(t)}{C^2(t)}W_Z^1(dt).$$

If $u \geq v$ then $(I) = 0$ and moreover, from (12), we find that $(II) = [C(u)/G(u)]\, \int_0^v[G(t)/C^2(t)]W_Z^1(dt)$. Hence in general we have

$$\mathrm{Cov}\,(\tilde{C}_n(u), \tilde{L}_n(v)) = -\frac{C(u)}{G(u)}\int_0^{u\wedge v} \frac{G(t)}{C^2(t)}W_Z^1(dt),$$

which is the stated result.

(v). In order to calculate the covariance between $\tilde{C}_n(u)$ and $\tilde{W}_n(v,x)$ we first derive the joint distribution function of the observed uncensored observations, i.e.

$$\begin{aligned}
&W_{Z,X,T}^1(z,x,t)\\
&= P(Z \leq z, X \leq x, T \leq t, \delta = 1|T \leq Z)\\
&= \alpha^{-1}\int_0^{+\infty}\int_0^{c\wedge z}\int_0^x G(y\wedge t)F(dy,dx)H(dc)\\
&= \alpha^{-1}\left[\int_0^z\int_0^c G(y\wedge t)F(dy,x)H(dc) + \int_z^{+\infty}\int_0^z G(y\wedge t)F(dy,x)H(dc)\right]\\
&= \alpha^{-1}\left[\int_0^z G(y\wedge t)[H(z)-H(y-)]F(dy,x) + \int_0^z G(y\wedge t)\bar{H}(z)F(dy,x)\right]\\
&= \alpha^{-1}\int_0^z G(y\wedge t)\bar{H}(y-)F(dy,x)\\
&= \alpha^{-1}\int_0^{z\wedge t} G(y)\bar{H}(y-)F(dy,x) + \alpha^{-1}G(t)\int_{z\wedge t}^z \bar{H}(y-)F(dy,x).
\end{aligned}$$

Using the above expression we find

$$\begin{aligned}
&\mathrm{Cov}\,(\tilde{C}_n(u), \tilde{W}_n(v,x))\\
&= E\,\{I(T_i \leq u \leq Z_i, Z_i \leq v, X_i \leq x, \delta_i = 1)\} - C(u)W_{Z,X}^1(v,x)\\
&= W_{Z,X,T}^1(v,x,u) - W_{Z,X,T}^1(u-,x,u) - C(u)W_{Z,X}^1(v,x)\\
&= \alpha^{-1}G(u)\int_{u\wedge v}^v \bar{H}(y-)F(dy,x) - C(u)W_{Z,X}^1(v,x)\\
&= G(u)\int_{u\wedge v}^v \frac{1}{G(y)}W_{Z,X}^1(dy,x) - C(u)W_{Z,X}^1(v,x),
\end{aligned}$$

which proves the stated result.

4.2. Proof of Theorem 2

We can write

$$\mathrm{Cov}\,(y_1, y_2, x_1, x_2) = \sum_{i=1}^{16} E[T_i(y_1, y_2, x_1, x_2)],$$

where

$$
\begin{aligned}
T_1(y_1, y_2, x_1, x_2) &= A(y_1)A(y_2)\tilde{W}_n(y_1, x_1)\tilde{W}_n(y_2, x_2) \\
T_2(y_1, y_2, x_1, x_2) &= -A(y_1)\int_0^{y_2} \tilde{W}_n(y_1, x_1)\tilde{W}_n(u, x_2)A(du) \\
T_3(y_1, y_2, x_1, x_2) &= -A(y_1)\int_0^{y_2} \tilde{W}_n(y_1, x_1)\frac{\tilde{C}_n(v)}{C(v)}F(dv, x_2) \\
T_4(y_1, y_2, x_1, x_2) &= -A(y_1)\int_0^{y_2} \tilde{W}_n(y_1, x_1)\tilde{L}_n(v)F(dv, x_2) \\
T_5(y_1, y_2, x_1, x_2) &= -A(y_2)\int_0^{y_1} \tilde{W}_n(y_2, x_2)\tilde{W}_n(u, x_1)A(du) \\
T_6(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \tilde{W}_n(u, x_1)\tilde{W}_n(v, x_2)A(du)A(dv) \\
T_7(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \tilde{W}_n(u, x_1)\frac{\tilde{C}_n(v)}{C(v)}A(du)F(dv, x_2) \\
T_8(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \tilde{L}_n(v)\tilde{W}_n(u, x_1)A(du)F(dv, x_2) \\
T_9(y_1, y_2, x_1, x_2) &= -A(y_2)\int_0^{y_1} \tilde{W}_n(y_2, x_2)\frac{\tilde{C}_n(u)}{C(u)}F(du, x_1) \\
T_{10}(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \tilde{W}_n(v, x_2)\frac{\tilde{C}_n(u)}{C(u)}A(dv)F(du, x_1) \\
T_{11}(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \frac{\tilde{C}_n(v)}{C(v)}\frac{\tilde{C}_n(u)}{C(u)}F(du, x_1)F(dv, x_2) \\
T_{12}(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \frac{\tilde{C}_n(u)}{C(u)}\tilde{L}_n(v)F(du, x_1)F(dv, x_2) \\
T_{13}(y_1, y_2, x_1, x_2) &= -A(y_2)\int_0^{y_1} \tilde{W}_n(y_2, x_2)\tilde{L}_n(u)F(du, x_1) \\
T_{14}(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \tilde{L}_n(u)\tilde{W}_n(v, x_2)A(dv)F(du, x_1) \\
T_{15}(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \frac{\tilde{C}_n(v)}{C(v)}\tilde{L}_n(u)F(du, x_1)F(dv, x_2) \\
T_{16}(y_1, y_2, x_1, x_2) &= \int_0^{y_1}\int_0^{y_2} \tilde{L}_n(v)\tilde{L}_n(u)F(du, x_1)F(dv, x_2).
\end{aligned}
$$

From the covariance structures provided in Lemma 1 of Section 2, we can calculate the expectations of the above terms and obtain

$$\begin{aligned}
E[T_1(y_1,y_2,x_1,x_2)] &= A(y_1)A(y_2)\Big[W^1_{Z,X}(y_1\wedge y_2,x_1\wedge x_2)\\
&\qquad\qquad -W^1_{Z,X}(y_1,x_1)W^1_{Z,X}(y_2,x_2)\Big]\\
&\equiv (1.1)+(1.2);\\
E[T_2(y_1,y_2,x_1,x_2)] &= -A(y_1)A(y_2)W^1_{Z,X}(y_1\wedge y_2,x_1\wedge x_2)\\
&\quad +A(y_1)F(y_1\wedge y_2,x_1\wedge x_2)\\
&\quad +A(y_1)A(y_2)W^1_{Z,X}(y_1,x_1)W^1_{Z,X}(y_2,x_2)\\
&\quad -A(y_1)W^1_{Z,X}(y_1,x_1)F(y_2,x_2)\\
&\equiv (2.1)+(2.2)+(2.3)+(2.4);\\
E[T_3(y_1,y_2,x_1,x_2)] &= -A(y_1)\int_0^{y_2}[a_1(y_1,x_1)-a_1(v\wedge y_1,x_1)]a_2(dv,x_2)\\
&\quad +A(y_1)W^1_{Z,X}(y_1,x_1)F(y_2,x_2)\equiv (3.1)+(3.2);\\
E[T_4(y_1,y_2,x_1,x_2)] &= -A(y_1)\int_0^{y_2} b_1(v\wedge y_1,x_1)F(dv,x_2)\\
&\quad +A(y_1)\int_0^{y_2}\int_0^{v\wedge y_1}[a_1(y_1,x_1)-a_1(s,x_1)]h(ds)F(dv,x_2)\\
&\equiv (4.1)+(4.2);\\
E[T_5(y_1,y_2,x_1,x_2)] &= -A(y_1)A(y_2)W^1_{Z,X}(y_1\wedge y_2,x_1\wedge x_2)\\
&\quad +A(y_1)A(y_2)W^1_{Z,X}(y_1,x_1)W^1_{Z,X}(y_2,x_2)\\
&\quad +A(y_2)F(y_1\wedge y_2,x_1\wedge x_2)-A(y_2)W^1_{Z,X}(y_2,x_2)F(y_1,x_1)\\
&\equiv (5.1)+(5.2)+(5.3)+(5.4);\\
E[T_6(y_1,y_2,x_1,x_2)] &= A(y_1\wedge y_2)A(y_1\vee y_2)W^1_{Z,X}(y_1\wedge y_2,x_1\wedge x_2)\\
&\quad -A(y_1)F(y_1\wedge y_2,x_1\wedge x_2)-A(y_2)F(y_1\wedge y_2,x_1\wedge x_2)\\
&\quad +\int_0^{y_1\wedge y_2} A^2(u)W^1_{Z,X}(du,x_1\wedge x_2)\\
&\quad -A(y_1)A(y_2)W^1_{Z,X}(y_1,x_1)W^1_{Z,X}(y_2,x_2)\\
&\quad +A(y_1)W^1_{Z,X}(y_1,x_1)F(y_2,x_2)+A(y_2)W^1_{Z,X}(y_2,x_2)F(y_1,x_1)\\
&\quad -F(y_1,x_1)F(y_2,x_2)\equiv \sum_{k=1}^{8}(6.k);\\
E[T_7(y_1,y_2,x_1,x_2)] &= \int_0^{y_1}\int_0^{y_2}[a_1(u,x_1)-a_1(u\wedge v,x_1)]A(du)a_2(dv,x_2)\\
&\quad -W^1_{Z,X}(y_1,x_1)A(y_1)F(y_2,x_2)+F(y_1,x_1)F(y_2,x_2)\\
&\equiv (7.1)+(7.2)+(7.3);
\end{aligned}$$

$$\begin{aligned}
E[T_8(y_1,y_2,x_1,x_2)] &= \int_0^{y_1}\int_0^{y_2} b_1(u\wedge v,x_1)A(du)F(dv,x_2)\\
&\quad -\int_0^{y_1}\int_0^{y_2}\int_0^{u\wedge v}[a_1(u,x_1)-a_1(s,x_1)]h(ds)A(du)F(dv,x_2)\\
&\equiv (8.1)+(8.2);\\
E[T_9(y_1,y_2,x_1,x_2)] &= -A(y_2)\int_0^{y_1}[a_1(y_2,x_2)-a_1(v\wedge y_2,x_2)]a_2(dv,x_1)\\
&\quad +A(y_2)W^1_{Z,X}(y_2,x_2)F(y_1,x_1)\equiv (9.1)+(9.2);\\
E[T_{10}(y_1,y_2,x_1,x_2)] &= \int_0^{y_1}\int_0^{y_2}[a_1(v,x_2)-a_1(u\wedge v,x_2)]A(dv)a_2(du,x_1)\\
&\quad -W^1_{Z,X}(y_2,x_2)A(y_2)F(y_1,x_1)+F(y_1,x_1)F(y_2,x_2)\\
&\equiv (10.1)+(10.2)+(10.3);\\
E[T_{11}(y_1,y_2,x_1,x_2)] &= \int_0^{y_1}\int_0^{y_2}\frac{C(u\vee v)G(u\wedge v)}{C(u)C(v)G(u\vee v)}F(du,x_1)F(dv,x_2)\\
&\quad -F(y_1,x_1)F(y_2,x_2)\\
&\equiv (11.1)+(11.2);\\
E[T_{12}(y_1,y_2,x_1,x_2)] &= -\int_0^{y_1}\int_0^{y_2}\frac{h(u\wedge v)}{G(u)}F(du,x_1)F(dv,x_2)\equiv (12);\\
E[T_{13}(y_1,y_2,x_1,x_2)] &= -A(y_2)\int_0^{y_1} b_1(v\wedge y_2,x_2)F(dv,x_1)\\
&\quad +A(y_2)\int_0^{y_1}\int_0^{v\wedge y_2}[a_1(y_2,x_2)-a_1(s,x_2)]h(ds)F(dv,x_1)\\
&\equiv (13.1)+(13.2);\\
E[T_{14}(y_1,y_2,x_1,x_2)] &= \int_0^{y_1}\int_0^{y_2} b_1(u\wedge v,x_2)A(du)F(dv,x_1)\\
&\quad -\int_0^{y_1}\int_0^{y_2}\int_0^{u\wedge v}[a_1(v,x_2)-a_1(s,x_2)]h(ds)A(dv)F(du,x_1)\\
&\equiv (14.1)+(14.2);\\
E[T_{15}(y_1,y_2,x_1,x_2)] &= -\int_0^{y_1}\int_0^{y_2}\frac{h(u\wedge v)}{G(v)}F(du,x_1)F(dv,x_2)\equiv (15);\\
E[T_{16}(y_1,y_2,x_1,x_2)] &= \int_0^{y_1}\int_0^{y_2} b(u\wedge v)F(du,x_1)F(dv,x_2)\equiv (16).
\end{aligned}$$

Observe now that the following terms cancel:

(1.1) $\longrightarrow$ (2.1)	(2.4) $\longrightarrow$ (3.2)	(5.3) $\longrightarrow$ (6.3)	(6.8) $\longrightarrow$ (7.3)
(1.2) $\longrightarrow$ (2.3)	(5.1) $\longrightarrow$ (6.1)	(5.4) $\longrightarrow$ (6.7)	(9.2) $\longrightarrow$ (10.2)
(2.2) $\longrightarrow$ (6.2)	(5.2) $\longrightarrow$ (6.5)	(6.6) $\longrightarrow$ (7.2)	(10.3) $\longrightarrow$ (11.2)

Also observe here that, among the remaining terms, expressions (3.1), (4.1), (4.2), (7.1), (8.1) and (8.2) are similar to (9.1), (13.1), (13.2), (10.1), (14.1) and (14.2) respectively, except that y_1 and y_2 are interchanged as well as x_1 and x_2. In the following part, we will therefore consider only the first group of terms in detail. Before getting to these terms, which are more complicated, we observe that

$$(11.1)+(12)+(15)+(16) = \int_0^{y_1}\int_0^{y_2} K(u,v)F(du,x_1)F(dv,x_2) \equiv \mathcal{X}(y_1,y_2,x_1,x_2).$$

Now, consider (3.1). It is easy to see that

$$\begin{aligned}(3.1) &= -A(y_1)\int_0^{y_1\wedge y_2}[a_1(y_1,x_1)-a_1(v,x_1)]a_2(dv,x_2)\\ &= -A(y_1)a_1(y_1,x_1)a_2(y_1\wedge y_2,x_2) + A(y_1)\int_0^{y_1\wedge y_2} a_1(v,x_1)a_2(dv,x_2).\end{aligned}$$

A similar calculation yields

$$\begin{aligned}(7.1) = a_2(y_2,x_2)&\int_0^{y_1} a_1(u,x_1)A(du) - a_2(y_2,x_2)\int_0^{y_1\wedge y_2} a_1(u,x_1)A(du)\\ &+\int_0^{y_1\wedge y_2} a_1(u,x_1)d[A(u)a_2(u,x_2)] - A(y_1)\int_0^{y_1\wedge y_2} a_1(v,x_1)a_2(dv,x_2).\end{aligned}$$

Then we can write

$$\begin{aligned}(3.1)+(7.1) = a_2(y_2,x_2)&\int_{y_1\wedge y_2}^{y_1} a_1(u,x_1)A(du)+\int_0^{y_1\wedge y_2} a_1(u,x_1)d[A(u)a_2(u,x_2)]\\ &-A(y_1)a_1(y_1,x_1)a_2(y_1\wedge y_2,x_2).\end{aligned}$$

Observe that for $y_1 < y_2$ the above equation reduces to

$$-\int_0^{y_1} A(u)a_2(u,x_2)a_1(du,x_1),$$

and for $y_2 < y_1$

$$\begin{aligned}(3.1)+(7.1) &= a_2(y_2,x_2)\int_{y_2}^{y_1} a_1(u,x_1)A(du) + \int_0^{y_2} a_1(u,x_1)d[A(u)a_2(u,x_2)]\\ &\quad -A(y_1)a_1(y_1,x_1)a_2(y_2,x_2)\\ &= -a_2(y_2,x_2)\int_{y_2}^{y_1} A(u)a_1(du,x_1) - \int_0^{y_2} A(u)a_2(u,x_2)a_1(du,x_1).\end{aligned}$$

So that, in general we have

$$(3.1)+(7.1) = \mathcal{T}_1(y_1,y_2,x_1,x_2).$$

For the term (4.1) we can write

$$\begin{aligned}(4.1) = &-A(y_1)\int_0^{y_1\wedge y_2} b_1(v,x_1)F(dv,x_2)\\ &-A(y_1)b_1(y_1\wedge y_2,x_1)[F(y_2,x_2)-F(y_1\wedge y_2,x_2)],\end{aligned}$$

and for the term (8.1) we find

$$\begin{aligned}(8.1) &= \int_0^{y_1\wedge y_2}[F(y_2,x_2)-F(u,x_2)]b_1(u,x_1)A(du)\\ &\quad+\int_0^{y_1\wedge y_2}[A(y_1)-A(u)]b_1(u,x_1)F(du,x_2).\end{aligned}$$

Summing the above two terms we obtain

$$\begin{aligned}(4.1)+(8.1) &= -A(y_1)b_1(y_1\wedge y_2,x_1)[F(y_2,x_2)-F(y_1\wedge y_2,x_2)]\\ &\quad+\int_0^{y_1\wedge y_2}[F(y_2,x_2)-F(u,x_2)]b_1(u,x_1)A(du)\\ &\quad-\int_0^{y_1\wedge y_2}A(u)b_1(u,x_1)F(du,x_2).\end{aligned}$$

Applying integration by parts and after some simplification we get

$$(4.1)+(8.1) = -\int_0^{y_1\wedge y_2}[F(y_2,x_2)-F(u,x_2)]A(u)b_1(du,x_1) = \mathcal{T}_2(y_1,y_2,x_1,x_2).$$

The terms (4.2) and (8.2) are more messy and we write their sum in the following compact form

$$\begin{aligned}(4.2)+(8.2) &= \int_0^{y_1}\int_0^{y_2}[d(y_1,v,x_1)-d(u,v,x_1)]A(du)F(dv,x_2)\\ &= \mathcal{T}_3(y_1,y_2,x_1,x_2).\end{aligned}$$

Now observe that

$$(3.1)+(4.1)+(4.2)+(7.1)+(8.1)+(8.2)\equiv\mathcal{T}(y_1,y_2,x_1,x_2)$$

$$(9.1)+(13.1)+(13.2)+(10.1)+(14.1)+(14.2)\equiv\mathcal{T}(y_2,y_1,x_2,x_1).$$

Then adding the remaining term (6.4), we obtain expression (8).

Acknowledgements

This research was supported by NATO Collaborative Research Grant CRG 950271. The first author was supported by ‘Projet d’Actions de Recherche Concertées’ (No. 93/98 - 164) and by an FNRS-grant (No. 1.5.001.95F) from the National Science Foundation (FNRS), Belgium. The authors thank the associate editor and the referees for their valuable comments which led to an improvement of the paper.

References

Gijbels, I. and Gürler, Ü. (1996). Covariance function of a bivariate distribution function estimator for left truncated and right censored data. *Discussion Paper* #9703, Institute of Statistics, Catholic University of Louvain, Louvain-la-Neuve, Belgium.

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.

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

Gürler, Ü. (1997). Bivariate distribution and hazard functions when a component is randomly truncated. *J. Multivariate Anal.* **60**, 20-47.

Gürler, Ü. and Gijbels, I. (1996). A bivariate distribution function estimator and its variance under left truncation and right censoring. *Discussion Paper* #9702, Institute of Statistics, Catholic University of Louvain, Louvain-la-Neuve, Belgium.

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

Institute of Statistics, Catholic University of Louvain, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium.

E-mail: gijbels@stat.ucl.ac.be

Department of Industrial Engineering, Bilkent University, 06533 Bilkent, Ankara, Turkey.

E-mail: ulku@bilkent.edu.tr

(Received November 1996; accepted November 1997)