Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

D.M. Bakker

Two nonparametric estimators of the survival function of bivariate right censored observations

Department of Operations Research, Statistics, and System Theory   Report BS-R9035   December

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Two Nonparametric Estimators of the Survival Function of Bivariate Right Censored Observations

Dick M. Bakker
Centre for Mathematics and Computer Science
P.O.Box 4079, 1009 AB Amsterdam
The Netherlands

A comparision is made of estimators for the bivariate survival function under random censoring. We consider an estimator based on the product integral and one based on the Volterra integral equation. Asymptotic root mean square errors are obtained through calculation of the influence curves of the estimators. We also compare the estimators with a pathwise estimator. A large number of simulation results are given.

(1980) Mathematics Subject Classification (1985 Revision): 62G05, 62P10, 65C05.
Key Words and Phrases: censored data, product-limit estimator, nonparametric estimation, volterra integral equation, survival analysis.

1. Introduction

The statistical analysis of univariate right censored data is well understood and the standard nonparametric estimator for the survival function is the Kaplan-Meier (1958) or product-limit estimator, which is actually the nonparametric maximum likelihood estimator (NPMLE). In this paper we study the extension to estimating the survival function of two dependent censored variables. This situation is quite common when a survival study is done on paired units, e.g. twins, brothers and sisters, married couples, both eyes of one patient, the response time of different treatments of a patient or the first experiences of certain disease symptoms.

We consider the following model: Let $T = (T_1, T_2)$ be a pair of nonnegative random variables, denoting unobservable survival or failure times with joint distribution function F and let $C = (C_1, C_2)$ be a pair of censoring times independent of T with joint distribution function G. The only observations available are (X, Δ) with

$$X = (X_1, X_2) = (T_1 \wedge C_1, T_2 \wedge C_2), \tag{1}$$

$$\Delta = (\Delta_1, \Delta_2) = (1_{\{T_1 \leq C_1\}}, 1_{\{T_2 \leq C_2\}}). \tag{2}$$

We have an i.i.d. sample $X_i = (X_{1i}, X_{2i})$, $\Delta_i = (\Delta_{1i}, \Delta_{2i})$, $i = 1, \ldots, n$, with each (X_i, Δ_i) distributed as (X, Δ).

One might hope that the NPMLE would again be a good estimator in this model. However it has some serious difficulties, and except for the following brief discussion we will not study it in this paper.

Only in some special cases of censoring is an NPMLE uniquely determined up to mass on certain regions, see e.g. Muñoz (1980), Campell (1981), Hanley and Parnes (1983). However

Report BS-R9035
Centre for Mathematics and Computer Science
P.O.Box 4079, 1009 AB Amsterdam, The Netherlands

in general there exists a kind of nonuniqueness of the NPMLE, as shown in the example below, that cannot be handled yet, because a closed form solution cannot be found, see Hanley and Parnes (1983). It is not clear to what NPMLE an iterative method like the EM-algorithm discribed by Dempster, Laird and Rubin (1977) will converge. To study this estimator, both empirically and theoretically, we need a faster algorithm to find and compare the various NPMLE's.

Example 1
This is a simple example for nonuniqueness of the NPMLE with four observations. Two observations (1 and 2) have the first failure time censored and two (3 and 4) have only the second one censored as in figure 1. The underlying uncensored observations lie on the lines drawn from the actual observations. Let p_i be the mass assigned to the intersection points. All solutions that satisfy the next equalities

$$\begin{aligned} p_1 + p_2 &= 1/2 \\ p_3 + p_4 &= 1/2 \\ p_1 + p_3 &= 1/2 \\ p_2 + p_4 &= 1/2 \end{aligned}$$

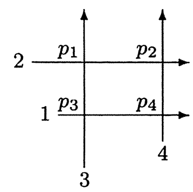

Figure 1. example 1 for non-uniqueness of the NPMLE.

are NPMLE solutions. For instance

$$p_1 = p_2 = p_3 = p_4 = 1/4$$

and

$$p_1 = p_4 = 0 \text{ and } p_2 = p_3 = 1/2$$

are both solutions.

Example 2
A more complicated example of nonuniqueness with six observations is given in figure 2. Again only the first or second failure time is censored. For the NPMLE the next equalities must hold.

$$\begin{aligned} p_1 + p_2 + p_3 &= 1/3 \\ p_4 + p_5 &= 1/3 \\ p_6 + p_7 + p_8 &= 1/3 \\ p_1 + p_6 &= 1/3 \\ p_2 + p_4 + p_7 &= 1/3 \\ p_3 + p_5 + p_8 &= 1/3 \end{aligned}$$

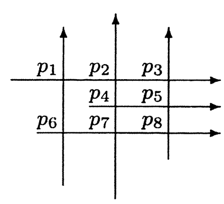

Figure 2. example 2 for non-uniqueness of the NPMLE.

For instance

$$p_2 = p_3 = p_5 = p_6 = p_7 = 0 \text{ and } p_1 = p_4 = p_8 = 1/3,$$

$$p_1 = p_2 = p_4 = p_5 = p_6 = p_8 = 1/6 \text{ and } p_3 = p_7 = 0$$

and

$$p_1 = p_3 = p_4 = p_7 = p_8 = 0 \text{ and } p_2 = p_5 = p_6 = 1/3$$

are all NPMLE solutions.

In this paper we abandon the NPMLE and compare the behaviour of the two-dimensional version of the product-limit estimator introduced by Dabrowska (1988,1989) and an estimator based on the Volterra representation of the survival function, an idea of P.J. Bickel. From now on we call them respectively "Dabrowska estimator" and "Volterra estimator". Also a short comparison is made with a pathwise product-limit estimator as introduced by Campell and Földes (1982). The comparison is mainly made by inspection of the influence functions of the estimators, i.e. we compare their asymptotic root mean square errors (R.M.S.E.). We also compare asymptotic and small sample (exact) R.M.S.E. by simulations.

2. Description of the estimators

2.1. Dabrowska estimator

Dabrowska (1988) introduced a bivariate Kaplan-Meier estimator, which is based on the product-integral representation of the survival function $\overline{F}(t) = \Pr(T_1 > t_1, T_2 > t_2)$, $t \in \mathbb{R}_+^2$

$$\overline{F}(t) = \prod_{u \le t_1} (1 - \Lambda_{10}(du, 0)) \prod_{v \le t_2} (1 - \Lambda_{01}(0, dv)) \prod_{\substack{u \le t_1 \\ v \le t_2}} (1 - L(du, dv)) \tag{3}$$

where $\prod (1 - X(du))$ denotes the product-integral as in Gill and Johansen (1990):

$$\prod_{u \le t} (1 - X(du)) = \lim_{\max|u_i - u_{i-1}| \to 0} \prod_i (1 - X((u_{i-1}, u_i])) \tag{4}$$

where $0 = u_0 < u_1 < \ldots < u_m = t$ is a partition of $(0, t]$ and

$$\prod_{\substack{u \le t_1 \\ v \le t_2}} (1 - X(du, dv)) = \lim_{\substack{\max|u_i - u_{i-1}| \to 0 \\ \max|v_j - v_{j-1}| \to 0}} \prod_{i,j} (1 - X((u_{i-1}, u_i] \times (v_{j-1}, v_j])) \tag{5}$$

where $0 = u_0 < u_1 < \ldots < u_{m_1} = t_1$ is a partition of $(0, t_1]$ and $0 = v_0 < v_1 < \ldots < v_{m_2} = t_2$ a partition of $(0, t_2]$ and the cumulative hazard function $\Lambda = (\Lambda_{10}, \Lambda_{01}, \Lambda_{11})$ is defined as

$$\Lambda_{10}(du, v) = \frac{-\overline{F}(du, v)}{\overline{F}(u-, v)} \tag{6}$$

$$\Lambda_{01}(u, dv) = \frac{-\overline{F}(u, dv)}{\overline{F}(u, v-)} \tag{7}$$

$$\Lambda_{11}(du, dv) = \frac{\overline{F}(du, dv)}{\overline{F}(u-, v-)} \tag{8}$$

with $\Lambda_{10}(0, v) = \Lambda_{01}(u, 0) = \Lambda_{11}(0, 0) = 0$ and $L(du, dv)$ is defined as

$$L(du,dv) = \frac{\Lambda_{10}(du,v-)\ \Lambda_{01}(u-,dv) - \Lambda_{11}(du,dv)}{(1-\Lambda_{10}(\Delta u,v-))\ (1-\Lambda_{01}(u-,\Delta v))}. \tag{9}$$

L can be called the "odds ratio measure", see Gill (1990) and Bickel, Gill and Wellner (1991). Formula (6) actually defines a measure (in the u variable) for each v on $\mathbb{R}_+$. Λ_{10} as a function is then the corresponding family of distribution functions, similarly for Λ_{01}. Λ_{11} is the d.f. of the measure on $\mathbb{R}_+^2$defined by (8). For more details see Dabrowska (1988) and Gill (1990). By replacing the hazard function by its natural estimate $\widehat{\Lambda} = (\widehat{\Lambda}_{10}, \widehat{\Lambda}_{01}, \widehat{\Lambda}_{11})$ defined as

$$\widehat{\Lambda}_{10}(du,v) = \frac{-N_{10}(du,v)}{Y(u-,v)} \tag{10}$$

$$\widehat{\Lambda}_{01}(u,dv) = \frac{-N_{01}(u,dv)}{Y(u,v-)} \tag{11}$$

$$\widehat{\Lambda}_{11}(du,dv) = \frac{N_{11}(du,dv)}{Y(u-,v-)} \tag{12}$$

with

$$N_{10}(u,v) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{(X_i,\Delta_i):X_{1i}>u,X_{2i}>v,\Delta_{1i}=1\}} \tag{13}$$

$$N_{01}(u,v) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{(X_i,\Delta_i):X_{1i}>u,X_{2i}>v,\Delta_{2i}=1\}} \tag{14}$$

$$N_{11}(u,v) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{(X_i,\Delta_i):X_{1i}>u,X_{2i}>v,\Delta_{1i}=\Delta_{2i}=1\}} \tag{15}$$

$$Y(u,v) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{(X_i,\Delta_i):X_{1i}>u,X_{2i}>v\}} \tag{16}$$

we get a consistent estimator for the survival function, which unfortunately is not always monotone. Pruitt (1991) studied to which kind of points negative mass is assigned and showed that the total amount of negative mass does not disappear as $n \to \infty$.

2.2. Volterra estimator

Another way to write the survival function, now with t, u etc. vectors (t_1,t_2), $(u_1,u_2) \in \mathbb{R}_+^2$, is

$$\overline{F}(t) = \overline{F}(t_1,0) + \overline{F}(0,t_2) - 1 + F(t_1,t_2) \tag{17}$$

$$= \overline{F}(t_1,0) + \overline{F}(0,t_2) - 1 + \int\int_{u\le t} \overline{F}(u-)\Lambda_{11}(du)\ , \qquad t,u \in \mathbb{R}_+^2 \tag{18}$$

with $\Lambda_{11}(du) = \Lambda_{11}(du_1,du_2)$ the hazard function of the uncensored observations as in (8). Here $u \le t$ means $u_1 \le t_1$ and $u_2 \le t_2$, and $u < t$ means $u_1 < t_1$ and $u_2 < t_2$. It was an idea by P.J. Bickel to use this Volterra equation for an estimate of the survival function. The solution of this equation given by Gill and Johansen (1990) is:

$$\overline{F}(t) = \overline{F}(t_1,0) + \overline{F}(0,t_2) - 1 + \int\int_{s\le t} (\overline{F}(s_1-,0) + \overline{F}(0,s_2-) - 1)\Lambda_{11}(ds)P(s,t) \tag{19}$$

with the Peano-series

$$P(s,t) = 1 + \sum_{k=1}^{\infty} \int_{s<u_1<\ldots<u_k\le t} \!\!\!\ldots\int \Lambda_{11}(du_1)\ldots\Lambda_{11}(du_k)\,, \qquad s,t,u_i \in \mathbb{R}_+^2 \tag{20}$$

By replacing $\overline{F}(t_1,0)$ and $\overline{F}(0,t_2)$ by their marginal estimators, for instance the Kaplan-Meier estimator, and Λ_{11} by its cumulative empirical estimate $\widehat{\Lambda}_{11}$ as with the Dabrowska estimator, we have a consistent estimate of the survival function $\overline{F}$, which is again not always a survival function itself. Because $\widehat{\Lambda}_{11} \neq 0$ only in uncensored observation points, mass is only assigned to these points and to points at the upper/right border of the support of $\widehat{\overline{F}}$. It is not clear what happens at the border of the support of the observation space.

3. The influence function

To calculate or approximate the asymptotic variance of an estimator in a given (i.e. 'theoretical') situation, we can look at the average influence of separate observations. Therefore we determine a kind of first order Taylor approximation to the estimator in the point $(\mathrm{E}N, \mathrm{E}Y)$, where $N = (N_{10}, N_{01}, N_{11})$. The difference between the estimate and the real value is then written as a sum of values that each only depend on one separate observation. That is, we find a function Ψ of each observation such that

$$\widehat{\overline{F}}(t) - \overline{F}(t) \approx \frac{1}{n}\sum_{i=1}^{n} \Psi(X_i, \Delta_i; t, F, G) \tag{21}$$

where $\approx$ can be interpreted here and below as: $= \cdots + o_P(n^{-\frac{1}{2}})$; see Gill(1989). The variance of the estimator is then approximated by

$$\mathrm{var}(\widehat{\overline{F}} - \overline{F})(t) \approx \frac{1}{n}\mathrm{var}\,(\Psi(X,\Delta;t,F,G))\,, \tag{22}$$

where $\mathrm{var}(\Psi(\cdots))$ is the variance of the influence function for one generic observation. In some special cases an exact expression for this variance can be found. In other situations we have to substitute the known values for $\overline{F}$, Λ, L, $\mathrm{E}Y$ and P and estimate the variance of the influence function by Monte-Carlo (i.e. by simulation of i.i.d. copies of $\Psi(X,\Delta;t,F,G)$). Even this is not easily possible in every example, since Ψ can be rather complicated to calculate.

First we give the influence function for the hazard function

$$d\widehat{\Lambda} = \frac{dN}{Y}, \text{ about } d\Lambda = \frac{\mathrm{E}(dN)}{\mathrm{E}Y}.$$

Formal second order Taylor expansion gives

$$\begin{aligned} d\widehat{\Lambda} - d\Lambda &\approx \frac{dN - \mathrm{E}(dN)}{\mathrm{E}Y} - \frac{\mathrm{E}(dN)}{(\mathrm{E}Y)^2}(Y - \mathrm{E}Y) + \\ &\qquad \left(\frac{-1}{(\mathrm{E}Y)^2}(dN - \mathrm{E}(dN))\,(Y - \mathrm{E}Y) + \frac{\mathrm{E}(dN)}{(\mathrm{E}Y)^3}(Y - \mathrm{E}Y)^2\right) \\ &= \frac{dN - Y\,d\Lambda}{\mathrm{E}Y} - \frac{(Y - \mathrm{E}Y)(dN - Y\,d\Lambda)}{(\mathrm{E}Y)^2}; \end{aligned} \tag{23}$$

reducing to the first order part and substituting dN and Y by the one-dimensional analogues of (15) and (16) gives

$$d\widehat{\Lambda} - d\Lambda \approx \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1_{\{X_i\in ds,\Delta_i=1\}}}{\mathrm{E}Y(s)} - \frac{1_{\{X_i>s\}}\Lambda(ds)}{\mathrm{E}Y(s)}\right). \tag{24}$$

Integrating (or adding) over $(0,t]$ gives

$$\widehat{\Lambda}(t) - \Lambda(t) = \int\limits_{(0,t]}\left(d\widehat{\Lambda} - d\Lambda\right) \approx \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1_{\{X_i\le t,\Delta_i=1\}}}{\mathrm{E}Y(X_i)} - \int\limits_{(0,X_i\wedge t]}\frac{\Lambda(ds)}{\mathrm{E}Y(s)}\right). \tag{25}$$

From a mathematical point of view, the result (25) of these formal algebraic manipulations is precisely the first order Taylor expansion of $\widehat{\Lambda}$ considered as a *Hadamard differentiable functional* of the empirical distribution function, see Gill (1989) and Gill & Johansen (1990). The correctness of the algebra is shown in van der Laan (1990).

For the bivariate estimators we can do a similar kind of heuristic derivation. To make it more readable we write: $d\Lambda_{10} = \Lambda_{10}(ds_1, s_2-)$, $d\Lambda_{01} = \Lambda_{01}(s_1-, ds_2)$, $d\Lambda = d\Lambda_{11} = \Lambda_{11}(ds_1, ds_2)$, $d\Lambda_1 = \Lambda_{10}(ds_1, 0)$, $d\Lambda_2 = \Lambda_{01}(0, ds_2)$, $dN_1 = N_{10}(ds_1, 0)$, $dN_2 = N_{01}(0, ds_2)$, $Y_1 = Y(s_1, 0)$, $Y_2 = Y(0, s_2)$, $\overline{F}_1 = \overline{F}(s_1, 0)$ and $\overline{F}_2 = \overline{F}(0, s_2)$.

3.1. Influence function of the Dabrowska estimator

To simplify we suppose the "true" F is continuous. Then for the Dabrowska estimator we write (9) as

$$dL = \frac{d\Lambda_{10}\,d\Lambda_{01} - d\Lambda_{11}}{(1-\Delta\Lambda_{10})\,(1-\Delta\Lambda_{01})}, \tag{26}$$

and first order Taylor expansion gives

$$\begin{aligned} d\widehat{L} - dL &\approx \frac{d\Lambda_{01}\,(d\widehat{\Lambda}_{10} - d\Lambda_{10}) + d\Lambda_{10}\,(d\widehat{\Lambda}_{01} - d\Lambda_{01}) - (d\widehat{\Lambda}_{11} - d\Lambda_{11})}{(1-\Delta\Lambda_{10})\,(1-\Delta\Lambda_{01})} \\ &\approx \frac{dN_{10}\,d\Lambda_{01}}{\mathrm{E}Y} + \frac{dN_{01}\,d\Lambda_{10}}{\mathrm{E}Y} - 2\frac{Y d\Lambda_{10}\,d\Lambda_{01}}{\mathrm{E}Y} - \frac{dN_{11} - Y\,d\Lambda_{11}}{\mathrm{E}Y} \end{aligned} \tag{27}$$

Here we use that $(1-\Delta\Lambda) = 1$ at the true Λ. Then we have from (3) using again continuity of F

$$\begin{aligned} \widehat{\overline{F}} - \overline{F} &\approx -\overline{F}(t_1,t_2)\Bigg\{\int\limits_{(0,t_1]}\left(d\widehat{\Lambda}_1 - d\Lambda_1\right) + \int\limits_{(0,t_2]}\left(d\widehat{\Lambda}_2 - d\Lambda_2\right) + \iint\limits_{(0,t]}\left(d\widehat{L} - dL\right)\Bigg\} \\ &\approx -\overline{F}(t_1,t_2)\Bigg\{\int\limits_{(0,t_1]}\frac{dN_1 - Y_1\,d\Lambda_1}{\mathrm{E}Y_1} + \int\limits_{(0,t_2]}\frac{dN_2 - Y_2\,d\Lambda_2}{\mathrm{E}Y_2} \\ &\qquad + \iint\limits_{(0,t]}\frac{dN_{10}\,d\Lambda_{01} + dN_{01}\,d\Lambda_{10} - 2Yd\Lambda_{10}\,d\Lambda_{01} - dN_{11} + Y\,d\Lambda_{11}}{\mathrm{E}Y}\Bigg\} \end{aligned} \tag{28}$$

The heuristic used here is to think of the product integral $\mathcal{\pi}$ just as an ordinary finite product $\prod$, just as before we thought of the integral $\int$ as a finite sum $\sum$. Hence the influence function for the Dabrowska estimator is

$$
\begin{aligned}
IC_{dabrowska}&(t_1,t_2) = \\
-\overline{F}(t_1,t_2)\Bigg\{ & \frac{1_{\{X_1\le t_1,\Delta_1=1\}}}{\mathrm{E}Y(X_1,0)} - \int_{(0,X_1\wedge t_1]} \frac{\Lambda_{10}(ds_1,0)}{\mathrm{E}Y(s_1,0)} \\
& + \frac{1_{\{X_2\le t_2,\Delta_2=1\}}}{\mathrm{E}Y(0,X_2)} - \int_{(0,X_2\wedge t_2]} \frac{\Lambda_{01}(0,ds_2)}{\mathrm{E}Y(0,s_2)} \\
& + 1_{\{X_1\le t_1,\Delta_1=1\}} \int_{(0,X_2\wedge t_2]} \frac{\Lambda_{01}(X_1,ds_2)}{\mathrm{E}Y(X_1,s_2)} \\
& + 1_{\{X_2\le t_2,\Delta_2=1\}} \int_{(0,X_1\wedge t_1]} \frac{\Lambda_{10}(ds_1,X_2)}{\mathrm{E}Y(s_1,X_2)} \\
& - 2 \iint_{(0,X\wedge t]} \frac{\Lambda_{10}(ds_1,s_2)\,\Lambda_{01}(s_1,ds_2)}{\mathrm{E}Y(s_1,s_2)} \\
& - \frac{1_{\{X_1\le t_1,\,\Delta_1=1,\,X_2\le t_2,\,\Delta_2=1\}}}{\mathrm{E}Y(X_1,X_2)} + \iint_{(0,X\wedge t]} \frac{\Lambda_{11}(ds_1,ds_2)}{\mathrm{E}Y(s_1,s_2)} \Bigg\}
\end{aligned}
\tag{29}
$$

When $\overline{F}$ and $\overline{G}$ can be written as products of their marginal distributions, i.e. in case of independent margins, we can write the influence function of the Dabrowska estimator in terms of the influence functions for estimating the margins. Independence gives $\overline{F} = \overline{F}_1\overline{F}_2$, $\overline{G} = \overline{G}_1\overline{G}_2$, $\Lambda_{10} = \Lambda_1$, $\Lambda_{01} = \Lambda_2$, $\Lambda_{11} = \Lambda_1\Lambda_2$. Define

$$
IC_k = -\overline{F}_k\left\{ \frac{1_{\{X_k\le t_k,\Delta_k=1\}}}{\mathrm{E}Y_k(X_k)} - \int_{(0,X_k\wedge t_k]} \frac{\Lambda_k(ds_k)}{\mathrm{E}Y(s_k)} \right\}, \qquad k=1,2. \tag{30}
$$

This is the IC of the univariate Kaplan-Meier estimator, see Gill and Johansen (1990). Then

$$
IC_{dabrowska} = IC_1\, IC_2 + \overline{F}_2\, IC_1 + \overline{F}_1\, IC_2 \tag{31}
$$

and the variance is given by

$$
VAR_{dabrowska} = \mathrm{var}(IC_1)\,\mathrm{var}(IC_2) + \overline{F}_2^2\,\mathrm{var}(IC_1) + \overline{F}_1^2\,\mathrm{var}(IC_2) \tag{32}
$$

because IC_1 and IC_2 are independent. It is known that

$$
\mathrm{var}(IC_k) = \overline{F}_k^2 \int_{(0,t_k]} \frac{d\Lambda_k}{\overline{F}_k\,\overline{G}_k}, \qquad k=1,2. \tag{33}
$$

It can be shown that the Dabrowska estimator is efficient in this case, see van der Laan (1990). When $\overline{F}$ and $\overline{G}$ are both *uniform (0,1)* on the plane we have

$$
\mathrm{var}(IC_k) = \frac{1}{2}(1-(1-t_k)^2), \qquad k=1,2. \tag{34}
$$

3.2. Influence function of the Volterra estimator

For the Volterra estimator we need

$$\widehat{P}(s,t) - P(s,t) \approx \int\int_{s<u\leq t} P(s,u-)\,(\widehat{\Lambda}-\Lambda)(du)P(u,t)\ , \qquad s,u,t \in \mathbb{R}_+^2 \tag{35}$$

which we get from Duhamel's equation, theorem 6 in Gill and Johansen (1990). Now we have

$$\begin{aligned}
\widehat{\overline{F}} - \overline{F} \approx\ & -\overline{F}_1(t_1)(\widehat{\Lambda}_1(t_1)-\Lambda_1(t_1)) - \overline{F}_2(\widehat{\Lambda}_2(t_2)-\Lambda_2(t_2)) && (36)\\
& -\int\int_{s\leq t}(\overline{F}_1(s_1)(\widehat{\Lambda}_1(s_1)-\Lambda_1(s_1)) + \overline{F}_2(s_2)(\widehat{\Lambda}_2(s_2)-\Lambda_2(s_2)))\Lambda(ds)P(s,t) && (37)\\
& +\int\int_{s\leq t}(\overline{F}_1(s_1)+\overline{F}_2(s_2)-1)(\widehat{\Lambda}-\Lambda)(ds)P(s,t) && (38)\\
& +\int\int_{s\leq t}\int\int_{s<u\leq t}(\overline{F}_1(s_1)+\overline{F}_2(s_2)-1)\Lambda(ds)P(s,u-)(\widehat{\Lambda}-\Lambda)(du)P(u,t) && (39)\\
=\ & -\overline{F}_1(t_1)(\widehat{\Lambda}_1(t_1)-\Lambda_1(t_1)) - \overline{F}_2(\widehat{\Lambda}_2(t_2)-\Lambda_2(t_2))\\
& -\int\int_{s\leq t}(\overline{F}_1(s_1)(\widehat{\Lambda}_1(s_1)-\Lambda_1(s_1)) + \overline{F}_2(s_2)(\widehat{\Lambda}_2(s_2)-\Lambda_2(s_2)))\Lambda(ds)P(s,t)\\
& +\int\int_{s\leq t}\overline{F}(s)(\widehat{\Lambda}-\Lambda)(ds)P(s,t) && (40)
\end{aligned}$$

One gets (36) and (37) by differentiating with respect to Λ_1 and Λ_2, (38) w.r.t. Λ and (39) w.r.t. P. Replacing $(\widehat{\Lambda}-\Lambda)$ by its first order approximation we get for the influence function

$$\begin{aligned}
& IC_{volterra}(t_1,t_2) = \\
& -\overline{F}_1(t_1)\left\{\frac{1_{\{X_1\leq t_1,\Delta_1=1\}}}{\mathrm{E}Y(X_1,0)} - \int_{(0,X_1\wedge t_1]}\frac{\Lambda_{10}(ds_1,0)}{\mathrm{E}Y(s_1,0)}\right\}\\
& -\overline{F}_2(t_2)\left\{\frac{1_{\{X_2\leq t_2,\Delta_2=1\}}}{\mathrm{E}Y(0,X_2)} - \int_{(0,X_2\wedge t_2]}\frac{\Lambda_{01}(0,ds_2)}{\mathrm{E}Y(0,s_2)}\right\}\\
& -\int\int_{s\leq t}\left\{\overline{F}_1(s_1)\left(\frac{1_{\{X_1\leq s_1,\Delta_1=1\}}}{\mathrm{E}Y(X_1,0)} - \int_{(0,X_1\wedge s_1]}\frac{\Lambda_{10}(du_1,0)}{\mathrm{E}Y(u_1,0)}\right)\right.\\
& \qquad\left. +\overline{F}_2(s_2)\left(\frac{1_{\{X_2\leq s_2,\Delta_2=1\}}}{\mathrm{E}Y(0,X_2)} - \int_{(0,X_2\wedge s_2]}\frac{\Lambda_{01}(0,du_2)}{\mathrm{E}Y(0,u_2)}\right)\right\}\Lambda(ds_1,ds_2)P(s,t)\\
& +1_{\{X_1\leq t_1,\,\Delta_1=1,\,X_2\leq t_2,\,\Delta_2=1\}}\frac{\overline{F}(X)P(X,t)}{\mathrm{E}Y(X)}\\
& -\int\int_{(0,X\wedge t]}\overline{F}(s)\frac{\Lambda(ds)}{\mathrm{E}Y(s)}P(s,t) \qquad\qquad (41)
\end{aligned}$$

Unfortunately this formula is difficult to handle in concrete examples.

In case of independent failure times we have

$$P(s,t) = \sum_{k=0}^{\infty} \left(\frac{1}{k!}\right)^2 (\Lambda_1(t_1) - \Lambda_1(s_1))^k (\Lambda_2(t_2) - \Lambda_2(s_2))^k \tag{42}$$

Even with this expression it is still difficult to find an exact expression for the variance of the influence function. We only can do Monte-Carlo estimation in some special cases, because only the first few terms of the summation are important.

3.3. Pathwise estimator

In some examples we will compare these estimators with the "pathwise estimator" as defined by Campell and Földes (1982). We only consider the asymptotic variance of this estimator and we will not do any simulation with it. This estimator is based on the property that

$$\Pr(T_1 > t_1, T_2 > t_2) = \Pr(T_2 > t_2 | T_1 > t_1)\ \Pr(T_1 > t_1) \tag{43}$$

One can then estimate both probabilities by their separate Kaplan-Meier estimates. The influence function is given by

$$\begin{aligned} IC_{pathwise}(t_1,t_2) \;=\; -\,\overline{F}\Bigg\{ & \frac{1_{\{X_1 \le t, \Delta_1 = 1\}}}{\mathrm{E}Y(X_1,0)} - \int_{(0,X_1 \wedge t_1]} \frac{\Lambda_{10}(ds_1,0)}{\mathrm{E}Y(s_1,0)} \\ & + \frac{1_{\{X_2 \le t_2, \Delta_2 = 1, X_1 > t_1\}}}{\mathrm{E}Y(X_1,X_2)} - 1_{\{X_1 > t_1\}} \int_{(0,X_2 \wedge t_2]} \frac{\Lambda_{01}(t_1,ds_2)}{\mathrm{E}Y(t_1,s_2)} \Bigg\} \end{aligned} \tag{44}$$

The variance of this influence function is given by

$$VAR_{pathwise}(t_1,t_2) = \overline{F}(t_1,t_2)^2 \left\{ \int_{(0,t_1]} \frac{\Lambda_{10}(ds_1,0)}{\mathrm{E}Y(s_1,0)} + \int_{(0,t_2]} \frac{d\Lambda_{01}(t_1,ds_2)}{\mathrm{E}Y(t_1,s_2)} \right\} \tag{45}$$

For the pathwise estimator there are two different paths dependent on which coördinate is taken first. In appendix B we give results for this variance in different situations.

4. Algorithms and Simulations

For both the Dabrowska and the Volterra estimator the mass is concentrated on the set of points

$$\begin{aligned} S \;=\; & \{(X_1,X_2) : X_1 = X_{1i}, \delta_{1i} = 1, X_2 = X_{2j}, \delta_{2j} = 1 \text{ for some } 1 \le i,j \le n\} \\ & \cup \{(X_1, X_{2max}) : X_1 = X_{1i} \text{ for some } 1 \le i \le n\} \\ & \cup \{(X_{1max}, X_2) : X_2 = X_{2j} \text{ for some } 1 \le j \le n\} \end{aligned} \tag{46}$$

with (X_{1max}, X_{2max}) the upper bound of the support of the observation space of X. Hence we can restrict the calculations to an $(n+1) \times (n+1)$ grid that is defined by the pairs (X_{1i}, X_{2j}), (X_{1i}, X_{2max}), (X_{1max}, X_{2j}), for $1 \le i,j \le n$. Define $0 = Y_{10} < \ldots < Y_{1n} < Y_{1n+1} = X_{1max}$ and $0 = Y_{20} < \ldots < Y_{2n} < Y_{2n+1} = X_{2max}$ as the ordered observation times of X_{1i} and X_{2i} respectively. The Dabrowska estimator on this grid of points is

$$\widehat{\widehat{F}}_{kl} = \widehat{\widehat{F}}(Y_{1k}, Y_{2l}) = \prod_{i\le k}(1-\Lambda_{10_{i0}})\prod_{j\le l}(1-\Lambda_{01_{0j}})\prod_{\substack{i\le k\\ j\le l}}(1-L_{ij}) \tag{47}$$

with $\Lambda_{10_{ij}} = \widehat{\Lambda}_{10}(dY_i, Y_j)$, $\Lambda_{01_{ij}} = \widehat{\Lambda}_{01}(Y_i, dY_j)$ and $L_{ij} = \widehat{L}(dY_i, dY_j)$. To calculate this estimator we first order the observation points from right to left and from above downwards. Then we have a logical order to calculate N_{10}, N_{01}, N_{11} and Y. Next the estimator is determined from left to right and from below upwards. For that purpose we use the following relation

$$\widehat{\widehat{F}}_{kl} = \widehat{\widehat{F}}_{k,l-1}\,(1-\Lambda_{01_{0l}})\prod_{i\le k}(1-L_{il}) = \widehat{\widehat{F}}_{k-1,l}\,(1-\Lambda_{10_{k0}})\prod_{j\le l}(1-L_{kj}) \tag{48}$$

for $k = 1, \ldots, n$ and $l = 1, \ldots, n$.

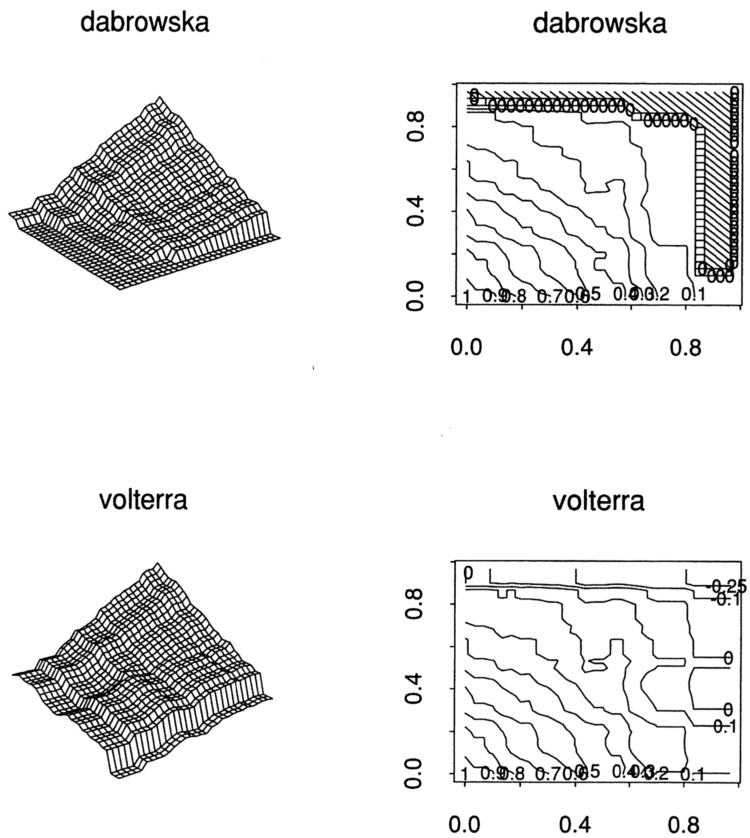

Figure 3. Dabrowska and Volterra estimator for the survival function for 100 observations and both failure time and censor *uniform (0,1)* on the plane.

The Volterra estimator on the grid is

$$\widehat{\widehat{F}}_{kl} = \widehat{\widehat{F}}_{k0} + \widehat{\widehat{F}}_{0l} - 1 + \sum_{i\le k}\sum_{j\le l} (\widehat{\widehat{F}}_{i-1,0} + \widehat{\widehat{F}}_{0,j-1} - 1)\ \Lambda_{ij}$$

$$\left(1 + \sum_{m\ge 1} \sum_{\substack{i<k_1<\ldots<k_m\le k \\ j<l_1<\ldots<l_m\le l}} \cdots \sum \Lambda_{k_1 l_1} \ldots \Lambda_{k_m l_m}\right); \qquad (49)$$

here we write Λ_{ij} for $\Lambda_{11_{ij}} = \widehat{\Lambda}_{11}(dY_i, dY_j)$. For the Volterra estimator we use the iterative scheme

$$\widehat{\widehat{F}}_{kl} = \widehat{\widehat{F}}_{k0} + \widehat{\widehat{F}}_{0l} - 1 + \sum_{\substack{i\le k \\ j\le l}} \widehat{\Lambda}_{ij} \widehat{\widehat{F}}_{i-1,j-1} \qquad (50)$$

again for $k = 1, \ldots, n$ and $l = 1, \ldots, n$. For the marginal survival functions we use the one dimensional Kaplan-Meier estimator.

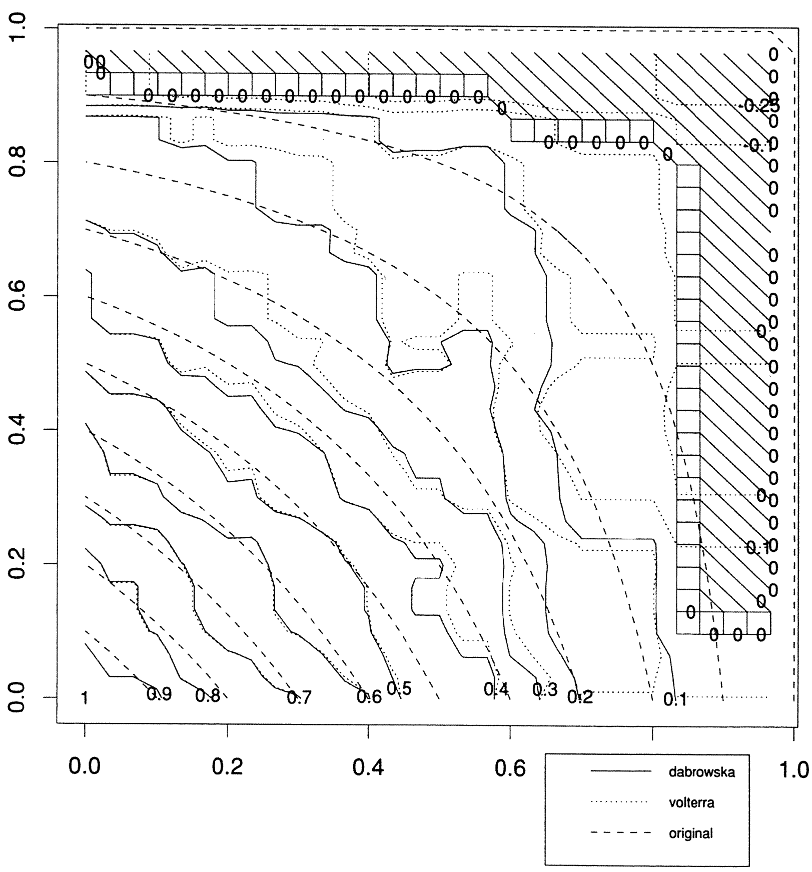

Figure 4. Combination of the contour plot of the Dabrowska and the Volterra estimator for 100 observations, compared with the original survival function with both failure time and censor *uniform (0,1)* on the plane.

Figure 3 shows an example of these two estimators for 100 observations and both the failure time and censor *uniform (0,1)* on the plane. In figure 4 we can see that the difference between the Dabrowska and the Volterra estimator is small left below and large in the upper right region.

5. Results

To find out what the behaviour of the estimators is in different situations and to compare the large sample (influence function) and the small sample behaviour at $n = 100$, we did a large number of simulations with different survival and censoring distributions.

For simulation of the data we used the new S language, see Becker, Chambers and Wilks (1988). The estimators themselves were programmed separately in C.

First we introduce a family of survival distributions which gives a nice collection for simulations, varying from completely dependent to totally independent margins.

5.1. The Pareto distribution

To get a set of combinations of distribution functions we use the bivariate Pareto distribution with each margin transformed to the uniform distribution. For the survival function $\overline{F}$ we take a reflected bivariate Pareto distribution and for the distribution function G the standard form Pareto distribution itself. They are defined by

$$\overline{F}_a(u,v) = ((1-u)^{-1/a} + (1-v)^{-1/a} - 1)^{-a} \tag{51}$$

and

$$G_b(u,v) = (u^{-1/b} + v^{-1/b} - 1)^{-b} \tag{52}$$

Note that the densities of F and G are mirrors of each other in the line $y = 1 - x$, see figure 5. These distribution functions and how to simulate them are described by Johnson (1987). When $a \to \infty$ and $b \to \infty$ we get independence; $\overline{F}_\infty(u,v) = (1-u)(1-v)$ and $G_\infty(u,v) = uv$. When $a \to 0$ and $b \to 0$ we get complete dependent margins; $\overline{F}_0(u,v) = \min\{1-u, 1-v\}$ and $G_0(u,v) = \min\{u,v\}$.

When a and b are 0, 1 or ∞ we can rather simply determine the asymptotic variances of the pathwise estimator. First we will handle these nine situations.

In the relevant tables, see appendix A, the asymptotic R.M.S.E. is given for the two pathwise estimators and a sample R.M.S.E. based on 100 replicates of the estimators and an asymptotic R.M.S.E. is given for the Dabrowska and the Volterra estimator at $n = 100$. The asymptotic values are used as approximation for the R.M.S.E. at $n = 100$. For the asymptotic R.M.S.E. of the Dabrowska and the Volterra estimator there are three possibilities: the variance of the influence curve is determined exactly; the R.M.S.E. is approximated by Monte-Carlo (based on 10,000 replicates of the influence curve); or we cannot handle the influence curve at all. In the last cases the simulations were done also for $n = 200$ observations.

When we compare the different R.M.S.E.'s, we have to be aware of the different sources of differences. First we have the sample variability of the R.M.S.E. (for 100 replicates of the estimators), not the true value.

Secondly the asymptotic properties do not exactly hold for estimators based on a small sample size ($n = 100$).

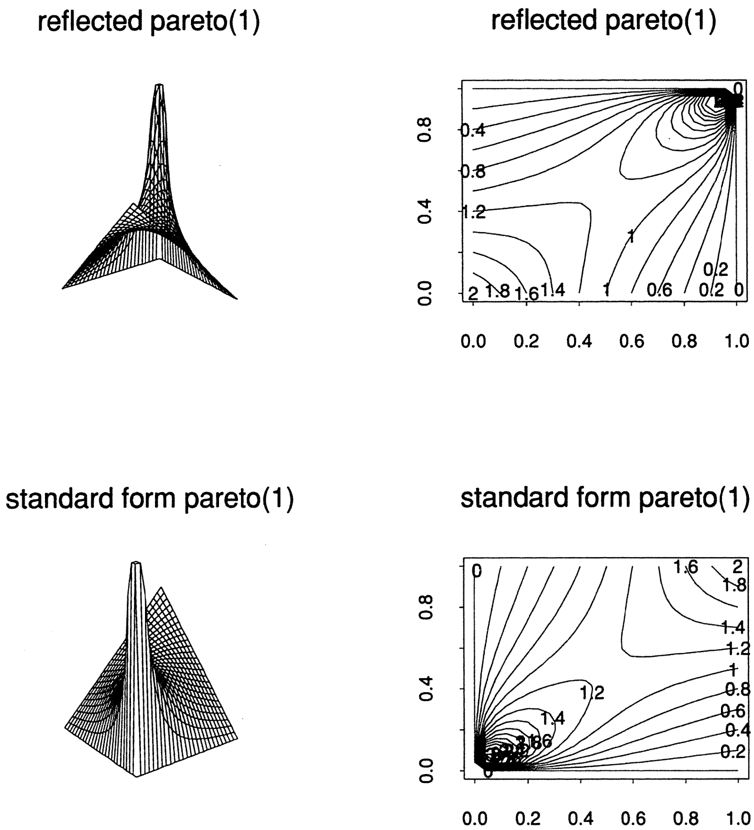

Figure 5. densities of *Pareto(1)* survival functions.

Finally, the asymptotic R.M.S.E. we get via the influence function is sometimes approximated by Monte-Carlo. This also gives some variation.

5.1.1. Pareto $a = 0$, $b = 0$

In this case both the failure and the censoring times are completely dependently distributed; we have uniform distributions on the diagonal. The pathwise, Dabrowska and the Volterra estimator are the same (though calculated in rather different ways!) and equal to the Kaplan-Meier estimate on the diagonal. The asymptotic variance is known.

$$VAR(t_1, t_2) = \frac{1}{2}\left(1 - (1 - t_1 \vee t_2)^2\right) \tag{53}$$

From table 1 we see that at a sample size of $n = 100$ the asymptotic R.M.S.E. is a good approximation to the sample R.M.S.E. based on 100 replicates of the estimator, but we have to be careful with conclusions only based on these values.

5.1.2. Pareto $a = 0,\ b = 1$

The failure times have completely dependent margins and the censor times are related in a certain way characterized by the *pareto(1)* distribution function

$$\overline{G}_1(t_1,t_2) = 1 - t_1 - t_2 + (t_1^{-1} + t_2^{-1} - 1)^{-1} = \frac{(t_1+t_2)(1-t_1)(1-t_2)}{t_1+t_2-t_1t_2} \tag{54}$$

Because

$$\Lambda_{01}(t_1, dt_2) = 0 \text{ for } t_1 > t_2, \tag{55}$$

$$\begin{aligned}\int\limits_{(0,t_2]} \frac{\Lambda_{01}(t_1,ds_2)}{\mathrm{E}Y(t_1,s_2)} &= \int\limits_{(t_1,t_2]} \frac{t_1+s_2-t_1s_2}{(t_1+s_2)(1-t_1)(1-s_2)^3} ds_2 \\ &= \frac{1}{2\,(1-t_1)(1+t_1)}\left(\frac{1}{(1-t_2)^2} - \frac{1}{(1-t_1)^2}\right) \\ &\quad + \frac{t_1^2}{(1-t_1)(1+t_1)^2}\left(\frac{1}{1-t_2} - \frac{1}{1-t_1}\right) \\ &\quad + \frac{t_1^2}{(1-t_1)(1+t_1)^3}\left(\log\frac{t_1+t_2}{2t_1} - \log\frac{1-t_2}{1-t_1}\right) \text{ for } t_1 \le t_2\,,\end{aligned} \tag{56}$$

and

$$\begin{aligned}\iint\limits_{s\le t} \frac{\Lambda_{11}(ds)}{\mathrm{E}Y(s)} &= \int\limits_{(0,t_1\wedge t_2]} \frac{ds}{(1-s)^2\frac{2s(1-s)^2}{2s-s^2}} \\ &= \frac{1}{2}\left(\frac{1}{3(1-t_1\wedge t_2)^3} - \frac{1}{3} + \frac{1}{2(1-t_1\wedge t_2)^2} - \frac{1}{2}\right),\end{aligned} \tag{57}$$

the influence curve of the Dabrowska estimator is

$$\begin{aligned} IC_{dabrowska}(t_1,t_2) = \\ -(1-t_1\vee t_2)\Bigg\{ &\frac{1_{\{X_1\le t_1,\Delta_1=1\}}}{(1-X_1)^2} - \frac{1}{2}\left(\frac{1}{(1-X_1\wedge t_1)^2} - 1\right) \\ &+ \frac{1_{\{X_2\le t_2,\Delta_2=1\}}}{(1-X_2)^2} - \frac{1}{2}\left(\frac{1}{(1-X_2\wedge t_2)^2} - 1\right) \\ &+ 1_{\{X_1\le t_1\wedge X_2\wedge t_2,\Delta_1=1\}}\Bigg\{\frac{1}{2\,(1-X_1)(1+X_1)}\left(\frac{1}{(1-X_2\wedge t_2)^2} - \frac{1}{(1-X_1)^2}\right) \\ &\quad + \frac{X_1^2}{(1-X_1)(1+X_1)^2}\left(\frac{1}{1-X_2\wedge t_2} - \frac{1}{1-X_1}\right) \\ &\quad + \frac{X_1^2}{(1-X_1)(1+X_1)^3}\left(\log\frac{X_1+X_2\wedge t_2}{2X_1} - \log\frac{1-X_2\wedge t_2}{1-X_1}\right)\Bigg\} \\ &+ 1_{\{X_2\le t_2\wedge X_1\wedge t_1,\Delta_2=1\}}\Bigg\{\frac{1}{2\,(1-X_2)(1+X_2)}\left(\frac{1}{(1-X_1\wedge t_1)^2} - \frac{1}{(1-X_2)^2}\right)\end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad + \frac{X_2^2}{(1-X_2)(1+X_2)^2}\left(\frac{1}{1-X_1\wedge t_1}-\frac{1}{1-X_2}\right)\\
&\qquad\qquad + \frac{X_2^2}{(1-X_2)(1+X_2)^3}\left(\log\frac{X_2+X_1\wedge t_1}{2X_2}-\log\frac{1-X_1\wedge t_1}{1-X_2}\right)\Bigg\}\\
&-0\\
&-\frac{1_{\{X_1\le t_1,\Delta_1=1,X_2\le t_2,\Delta_2=1\}}(X_1+X_2-X_1X_2)}{(1-X_1\vee X_2)(X_1+X_2)(1-X_1)(1-X_2)}\\
&+\frac{1}{2}\left(\frac{1}{3(1-X_1\wedge t_1\wedge X_2\wedge t_2)^3}-\frac{1}{3}+\frac{1}{2(1-X_1\wedge t_1\wedge X_2\wedge t_2)^2}-\frac{1}{2}\right)\Bigg\} \qquad (58)
\end{aligned}
$$

For Volterra it is also possible to determine the influence curve

$$
\begin{aligned}
&IC_{volterra}(t_1,t_2)=\\
&-(1-t_1)\left\{\frac{1_{\{X_1\le t_1,\Delta_1=1\}}}{(1-X_1)^2}-\frac{1}{2}\left(\frac{1}{(1-X_1\wedge t_1)^2}-1\right)\right\}\\
&-(1-t_2)\left\{\frac{1_{\{X_2\le t_2,\Delta_2=1\}}}{(1-X_2)^2}-\frac{1}{2}\left(\frac{1}{(1-X_2\wedge t_2)^2}-1\right)\right\}\\
&-\frac{1}{2\,(1-t_1\wedge t_2)}\Bigg\{\log(1-X_1\wedge t_1\wedge t_2)-\frac{1}{2}\left((1-t_1\wedge t_2)^2-1\right)\\
&\qquad\qquad -1_{\{X_1\le t_1\wedge t_2\}}\left(1_{\{\Delta_1=1\}}-\frac{1}{2}\right)\left(\frac{(1-t_1\wedge t_2)^2}{(1-X_1)^2}-1\right)\Bigg\}\\
&-\frac{1}{2\,(1-t_1\wedge t_2)}\Bigg\{\log(1-t_1\wedge X_2\wedge t_2)-\frac{1}{2}\left((1-t_1\wedge t_2)^2-1\right)\\
&\qquad\qquad -1_{\{X_2\le t_2\wedge t_1\}}\left(1_{\{\Delta_2=1\}}-\frac{1}{2}\right)\left(\frac{(1-t_2\wedge t_1)^2}{(1-X_2)^2}-1\right)\Bigg\}\\
&+1_{\{X_1\le t_1,\Delta_1=1,X_2\le t_2,\Delta_2=1\}}\frac{(1-X_1\vee X_2)(X_1+X_2-X_1X_2)}{(1-t_1\wedge t_2)(X_1+X_2)(1-X_1)(1-X_2)}\\
&-\frac{1}{2(1-t_1\wedge t_2)}\left\{-\log(1-X_1\wedge t_1\wedge X_2\wedge t_2)+\frac{1}{1-X_1\wedge t_1\wedge X_2\wedge t_2}-1\right\} \qquad (59)
\end{aligned}
$$

We used these expressions for the influence curves for Monte-Carlo approximation of the asymptotic R.M.S.E.'s. Comparing these and the asymptotic R.M.S.E. of the pathwise estimator in table 2 we see that the difference between the pathwise and the Volterra estimator is not to the advantage of one of them. For the points left below in the support Volterra seems better, partly dependent on which of the two pathwise estimators is taken. As we will see in other cases also the Volterra estimator seems very bad in points that are close to the upper/right border of the support. The Dabrowska estimator is the best one as long as we do not take points close to this border. There the difference between the pathwise and the Dabrowska estimator is unclear. The asymptotic R.M.S.E. based on 10,000 replicates of the influence curve is not reliable here.

5.1.3. Pareto $a = 0$, $b = \infty$

We can compare this situation with one failure time and two censoring times. The variances of the estimators are expected to be smaller than in for instance the independent case. The influence curve of the Dabrowska estimator is

$$
\begin{aligned}
&IC_{dabrowska}(t_1, t_2) = \\
&-(1 - t_1 \vee t_2)\Bigg\{ \frac{1_{\{X_1 \le t_1, \Delta_1 = 1\}}}{(1 - X_1)^2} - \frac{1}{2}\left(\frac{1}{(1 - X_1 \wedge t_1)^2} - 1\right) \\
&\qquad + \frac{1_{\{X_2 \le t_2, \Delta_2 = 1\}}}{(1 - X_2)^2} - \frac{1}{2}\left(\frac{1}{(1 - X_2 \wedge t_2)^2} - 1\right) \\
&\qquad + \frac{1_{\{X_1 \le t_1, \Delta_1 = 1\}} 1_{\{X_1 \le X_2 \wedge t_2\}}}{2\,(1 - X_1)}\left(\frac{1}{(1 - X_2 \wedge t_2)^2} - \frac{1}{(1 - X_1)^2}\right) \\
&\qquad + \frac{1_{\{X_2 \le t_2, \Delta_2 = 1\}} 1_{\{X_2 \le X_1 \wedge t_1\}}}{2\,(1 - X_2)}\left(\frac{1}{(1 - X_1 \wedge t_1)^2} - \frac{1}{(1 - X_2)^2}\right) \\
&\qquad - 0 \\
&\qquad - \frac{1_{\{X_1 \le t_1, \Delta_1 = 1, X_2 \le t_2, \Delta_2 = 1\}}}{(1 - X_1 \vee X_2)(1 - X_1)(1 - X_2)} + \frac{1}{3}\left(\frac{1}{1 - X_1 \wedge t_1 \wedge X_2 \wedge t_2)^3} - 1\right)\Bigg\}
\end{aligned} \tag{60}
$$

For Volterra we have

$$
P(s, t) = \frac{1 - s_1 \vee s_2}{1 - t_1 \wedge t_2} \tag{61}
$$

and this gives for the influence curve

$$
\begin{aligned}
&IC_{volterra}(t_1, t_2) = \\
&\quad - (1 - t_1)\left\{ \frac{1_{\{X_1 \le t_1, \Delta_1 = 1\}}}{(1 - X_1)^2} - \frac{1}{2}\left(\frac{1}{(1 - X_1 \wedge t_1)^2} - 1\right)\right\} \\
&\quad - (1 - t_2)\left\{ \frac{1_{\{X_2 \le t_2, \Delta_2 = 1\}}}{(1 - X_2)^2} - \frac{1}{2}\left(\frac{1}{(1 - X_2 \wedge t_2)^2} - 1\right)\right\} \\
&\quad - \frac{1}{2\,(1 - t_1 \wedge t_2)}\Bigg\{ \log(1 - X_1 \wedge t_1 \wedge t_2) - \frac{1}{2}\left((1 - t_1 \wedge t_2)^2 - 1\right) \\
&\qquad\qquad -1_{\{X_1 \le t_1 \wedge t_2\}}\left(1_{\{\Delta_1 = 1\}} - \frac{1}{2}\right)\left(\frac{(1 - t_1 \wedge t_2)^2}{(1 - X_1)^2} - 1\right)\Bigg\} \\
&\quad - \frac{1}{2\,(1 - t_1 \wedge t_2)}\Bigg\{ \log(1 - t_1 \wedge X_2 \wedge t_2) - \frac{1}{2}\left((1 - t_1 \wedge t_2)^2 - 1\right) \\
&\qquad\qquad - 1_{\{X_2 \le t_2 \wedge t_1\}}\left(1_{\{\Delta_2 = 1\}} - \frac{1}{2}\right)\left(\frac{(1 - t_2 \wedge t_1)^2}{(1 - X_2)^2} - 1\right)\Bigg\} \\
&\quad + 1_{\{X_1 \le t_1, \Delta_1 = 1, X_2 \le t_2, \Delta_2 = 1\}} \frac{1 - X_1 \vee X_2}{(1 - t_1 \wedge t_2)(1 - X_1)(1 - X_2)} \\
&\quad - \frac{1}{1 - t_1 \wedge t_2}\left(\frac{1}{1 - X_1 \wedge t_1 \wedge X_2 \wedge t_2} - 1\right)
\end{aligned} \tag{62}
$$

We used these expressions for Monte-Carlo approximation of the respective variances of the influence curves. From table 3 we see that in most points the sample R.M.S.E. of the Dabrowska estimator is smaller than the asymptotic R.M.S.E. of the pathwise estimator. From the point $(0.75, 0.75)$ the pathwise estimator has a smaller asymptotic R.M.S.E. than the Dabrowska estimator. This tells us nothing about which estimator of these two is asymptotically the best one, because at the upper/right border of the support there are relatively too few points to get a reliable value for the variance of the influence curve.

5.1.4. Pareto $a = 1$, $b = 0$

We have

$$\overline{F}(t_1, t_2) = \left((1-t_1)^{-1} + (1-t_2)^{-1} - 1\right)^{-1} = \frac{(1-t_1)(1-t_2)}{(1-t_1t_2)} \tag{63}$$

and

$$d\Lambda_{11} = \frac{2\, dt_1\, dt_2}{(1-t_1t_2)^2} \tag{64}$$

Because of the rather complex structure of double integrals with Λ_{11}, we were not able to find an expression for the influence curves of the Dabrowska and the Volterra estimator. In particular for the influence curve of the Dabrowska estimator the integral

$$\iint_{s \leq t} \frac{d\Lambda_{11}}{\mathrm{E}Y} = \iint_{s \leq t} \frac{2\, dt_1\, dt_2}{(1-t_1t_2)(1-t_1)(1-t_2)(1-t_1 \vee t_2)} \tag{65}$$

is difficult to handle. In general the influence curve of the Volterra estimator is even more complicated. Hence we only compare for $n = 100$ the sample R.M.S.E. of the Dabrowska and the Volterra estimator and the asymptotic R.M.S.E. of the pathwise estimator, see table 4. Apart from the point (0.1,0.1) the Dabrowska estimator has the smallest R.M.S.E.. The difference in the point (0.1,0.1) between the Dabrowska and the pathwise estimator is too small to conclude that the Dabrowska estimator is worse in this point.

Table 12 gives the results of simulations for $n = 200$ observations for each estimator. The pathwise estimator is worse than the Dabrowska estimator, except in the point $(0.1, 0.9)$. This can be a sample error. The difference between the Dabrowska and the Volterra estimator is small for the region left below of the support, but the overal impression is that the Dabrowska estimator is better.

5.1.5. Pareto $a = 1$, $b = 1$

Because of the more complicated structure of the survival function of the censor, we expect the same kind of problems for determining the influence curves as in the case with $a = 1$ and $b = 0$. In table 5 for $n = 100$ there is no uniform difference between the sample R.M.S.E. of the Dabrowska estimator and the asymptotic R.M.S.E. of the pathwise. In some points the pathwise seems better, in other points the Dabrowska estimator. $n = 100$ seems too small to make a good comparision.

Table 13 shows that for the estimators based on $n = 200$ observations the sample R.M.S.E. of the Dabrowska estimator is in most points smaller than the asymptotic R.M.S.E. of the pathwise estimator. The difference between the pathwise and the Volterra estimator is still unclear.

5.1.6. Pareto $a = 1$, $b = \infty$

Here the same kind of problems occur as in the case with $a = 1$ and $b = 0$. Even the same conclusions hold for table 6. For estimators based on $n = 200$ observations table 14 gives the Dabrowska estimator as the best one, except in the point $(0.1, 0.1)$. The difference between the pathwise and the Volterra estimator depends on the part of the support. Left below the Volterra estimator seems the second best, and as usual at the upper/right border of the support the Volterra is the worst.

5.1.7. Pareto $a = \infty$, $b = 0$

We have one censoring time for both failure times. This situation is quite common when you observe two survival variables of one object, for instance both eyes of a patient. The influence curve of the Dabrowska estimator can be determined. The variance is too complicated. We have

$$
\begin{aligned}
&IC_{dabrowska}(t_1, t_2) = \\
&\quad -(1-t_1)(1-t_2)\Bigg\{\frac{1_{\{X_1 \le t_1, \Delta_1 = 1\}}}{(1-X_1)^2} - \frac{1}{2}\left(\frac{1}{(1-X_1 \wedge t_1)^2} - 1\right) \\
&\qquad + \frac{1_{\{X_2 \le t_2, \Delta_2 = 1\}}}{(1-X_2)^2} - \frac{1}{2}\left(\frac{1}{(1-X_2 \wedge t_2)^2} - 1\right) \\
&\qquad + 1_{\{X_1 \le t_1, \Delta_1 = 1\}}\Bigg\{\frac{1}{(1-X_1)^2}\left(\frac{1}{1-X_1 \wedge X_2 \wedge t_2} - 1\right) \\
&\qquad\qquad + \frac{1_{\{X_1 > X_2 \wedge t_2\}}}{2\,(1-X_1)}\left(\frac{1}{1-X_2 \wedge t_2)^2} - \frac{1}{(1-X_1)^2}\right)\Bigg\} \\
&\qquad + 1_{\{X_2 \le t_2, \Delta_2 = 1\}}\Bigg\{\frac{1}{(1-X_2)^2}\left(\frac{1}{1-X_1 \wedge t_1 \wedge X_2} - 1\right) \\
&\qquad\qquad + \frac{1_{\{X_2 > X_1 \wedge t_1\}}}{2\,(1-X_2)}\left(\frac{1}{1-X_1 \wedge t_1)^2} - \frac{1}{(1-X_2)^2}\right)\Bigg\} \\
&\qquad - \frac{1_{\{X_1 \le t_1, \Delta_1 = 1, X_2 \le t_2, \Delta_2 = 1\}}}{(1-X_1)(1-X_2)(1-X_1 \vee X_2)} \\
&\qquad - \frac{1}{2}\left(\frac{1}{(1-X_1 \wedge t_1)^2} + \frac{1}{(1-X_2 \wedge t_2)^2}\right)\left(\frac{1}{1-X_1 \wedge t_1 \wedge X_2 \wedge t_2} - 1\right) \\
&\qquad\qquad + \frac{1}{3}\left(\frac{1}{(X_1 \wedge t_1 \wedge X_2 \wedge t_2)^3} - 1\right)\Bigg\} \qquad (66)
\end{aligned}
$$

For the Volterra estimator the first few terms are nearly equal to the corresponding terms in case of complete independence. We were not able to solve the last term. From table 7 we see that the approximated asymptotic R.M.S.E. of the Dabrowska estimator is in all points smaller than the asymptotic R.M.S.E. of the pathwise estimator. Comparing the Dabrowska and the Volterra estimator, only in (0.1,0.1) does the last one seem better, but the difference is small and in the other points Dabrowska is the better one.

Table 15 shows that in this case for $n = 200$ observations the conclusions depend on taking the sample or the asymptotic R.M.S.E.. Asymptotically the Dabrowska estimator is the best one. The differences between the sample R.M.S.E.'s are not large enough to be in conflict with this.

5.1.8. Pareto $a = \infty$, $b = 1$

In this case the same conclusions as for $a = 1$ and $b = 0$ holds for the influence curves of the Dabrowska and the Volterra estimator. Table 8 shows that $n = 100$ is too small to select a structurally better estimator in this case. Even $n = 200$ observations give no unambiguous picture of the best estimator as is shown in table 16, but in most points the Dabrowska estimator is the best one.

5.1.9. Pareto $a = \infty$ and $b = \infty$

Both the failure and the censoring time have independent *uniform (0,1)* distributed margins. The variance of the influence curve of the Dabrowska estimator is known from paragraph 3.1.:

$$\begin{aligned} VAR_{dabrowska}(t_1, t_2) &= \frac{1}{4}\left(1-(1-t_1)^2\right)\left(1-(1-t_2)^2\right) \\ &\quad + \frac{1}{2}(1-t_2)^2\left(1-(1-t_1)^2\right) + \frac{1}{2}(1-t_1)^2\left(1-(1-t_2)^2\right) \quad (67) \end{aligned}$$

This gives the opportunity to see if a Monte-Carlo approximation of the asymptotic variance based on 10,000 replicates of the influence curve gives a useful result. Apart from the upper/right border of the support (for instance the point (0.95,0.95)) the values correspond quite well. This is because at this border only a fraction of the observations contribute to the behaviour of the estimator and hence we need a lot more observation points to approach the asymptotic values here. We know that in this special case the Dabrowska estimator is efficient. The asymptotic R.M.S.E. of the pathwise estimator compared to that of the Dabrowska estimator is given by

$$VAR_{pathwise}(t_1, t_2) = VAR_{dabrowska}(t_1, t_2) + \frac{1}{4}\, t_1(2-t_1)t_2(2-t_2) \quad (68)$$

For the Volterra estimator we only were able to determine the influence curve.

$$\begin{aligned} &IC_{volterra}(t_1, t_2) = \\ &\quad -(1-t_1)\left\{\frac{1_{\{X_1 \le t_1, \Delta_1 = 1\}}}{(1-X_1)^2} - \frac{1}{2}\left(\frac{1}{(1-X_1 \wedge t_1)^2} - 1\right)\right\} \\ &\quad -(1-t_2)\left\{\frac{1_{\{X_2 \le t_2, \Delta_2 = 1\}}}{(1-X_2)^2} - \frac{1}{2}\left(\frac{1}{(1-X_2 \wedge t_2)^2} - 1\right)\right\} \\ &\quad - \iint_{(0,t]} \left\{\frac{1}{1-s_2}\left(\frac{1_{\{X_1 \le s_1, \Delta_1 = 1\}}}{(1-X_1)^2} - \frac{1}{2}\left(\frac{1}{(1-X_1 \wedge s_1)^2} - 1\right)\right)\right. \\ &\qquad\qquad \left. + \frac{1}{1-s_1}\left(\frac{1_{\{X_2 \le s_2, \Delta_2 = 1\}}}{(1-X_2)^2} - \frac{1}{2}\left(\frac{1}{(1-X_2 \wedge s_2)^2} - 1\right)\right)\right\} \\ &\qquad\qquad \sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-s_1}{1-t_1}\right)^k\left(\log\frac{1-s_2}{1-t_2}\right)^k ds_1 ds_2 \end{aligned}$$

$$+\frac{1_{\{X_1\le t_1,\,\Delta_1=1\}}\,1_{\{\,X_2\le t_2,\,\Delta_2=1\}}}{(1-X_1)(1-X_2)}\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-X_1}{1-t_1}\right)^k\left(\log\frac{1-X_2}{1-t_2}\right)^k$$

$$-\int\int_{(0,X\wedge t]}\frac{\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-s_1}{1-t_1}\right)^k\left(\log\frac{1-s_2}{1-t_2}\right)^k}{(1-s_1)^2(1-s_2)^2}\,ds_1ds_2 \tag{69}$$

Note that

$$(\log(1-x))^k dx = -\sum_{i=0}^{k}\frac{k!}{i!}\,(-1)^{k-i}(1-x)\,(\log(1-x))^i \tag{70}$$

and

$$\frac{(\log(1-x))^k}{(1-x)^2}\,dx = \sum_{i=0}^{k}\frac{k!}{i!}\,\frac{(\log(1-x))^i}{(1-x)} \tag{71}$$

So the influence curve is

$$-(1-t_1)\left\{\frac{1_{\{X_1\le t_1,\Delta_1=1\}}}{(1-X_1)^2}-\frac{1}{2}\left(\frac{1}{(1-X_1\wedge t_1)^2}-1\right)\right\}$$

$$-(1-t_2)\left\{\frac{1_{\{X_2\le t_2,\Delta_2=1\}}}{(1-X_2)^2}-\frac{1}{2}\left(\frac{1}{(1-X_2\wedge t_2)^2}-1\right)\right\}$$

$$-\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\log\frac{1}{1-t_2}\right)^{k+1}\Bigg\{$$

$$\frac{1}{2}\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\left(\log\frac{1}{1-t_1}\right)^i-(-1)^k(1-t_1)$$

$$-\frac{1}{2}\sum_{i=0}^{k}\frac{1}{i!}\left(\frac{\left(\log\frac{1-X_1\wedge t_1}{1-t_1}\right)^i}{1-X_1\wedge t_1}-\left(\log\frac{1}{1-t_1}\right)^i\right)$$

$$+1_{\{X_1\le t_1\}}\left(1_{\{\Delta_1=1\}}-\frac{1}{2}\right)\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\frac{\left(\log\frac{1-X_1}{1-t_1}\right)^i}{1-X_1}-(-1)^k\frac{1-t_1}{(1-X_1)^2}\Bigg\}$$

$$-\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\log\frac{1}{1-t_1}\right)^{k+1}\Bigg\{$$

$$\frac{1}{2}\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\left(\log\frac{1}{1-t_2}\right)^i-(-1)^k(1-t_2)$$

$$-\frac{1}{2}\sum_{i=0}^{k}\frac{1}{i!}\left(\frac{\left(\log\frac{1-X_2\wedge t_2}{1-t_2}\right)^i}{1-X_2\wedge t_2}-\left(\log\frac{1}{1-t_2}\right)^i\right)$$

$$+1_{\{X_2\le t_2\}}\left(1_{\{\Delta_2=1\}}-\frac{1}{2}\right)\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\frac{\left(\log\frac{1-X_2}{1-t_2}\right)^i}{1-X_2}-(-1)^k\frac{1-t_2}{(1-X_2)^2}\Bigg\}$$

$$+\frac{1_{\{X_1\le t_1,\,\Delta_1=1\}}\,1_{\{X_2\le t_2,\,\Delta_2=1\}}}{(1-X_1)(1-X_2)}\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-X_1}{1-t_1}\right)^k\left(\log\frac{1-X_2}{1-t_2}\right)^k$$
$$-\sum_{k=0}^{\infty}\sum_{i=0}^{k}\frac{1}{i!}\left(\frac{\left(\log\frac{1-X_1\wedge t_1}{1-t_1}\right)^i}{1-X_1\wedge t_1}-\left(\log\frac{1}{1-t_1}\right)^i\right)\sum_{j=0}^{k}\frac{1}{j!}\left(\frac{\left(\log\frac{1-X_2\wedge t_2}{1-t_2}\right)^j}{1-X_2\wedge t_2}-\left(\log\frac{1}{1-t_2}\right)^j\right) \tag{72}$$

We used this expression for a Monte-Carlo approximation of the asymptotic R.M.S.E. with the influence curve. Table 9 of course shows a smallest R.M.S.E. for the Dabrowska estimator. The difference between the pathwise and the Volterra estimator is not clear in the points left below of the support, but the pathwise seems the better one of these two.

5.2. Non-Pareto examples

We now consider two special cases for which we can determine the asymptotic R.M.S.E. with the influence curve for the pathwise and the Dabrowska estimator and approximate it for the Volterra estimator. For the different columns in the tables and the possible differences between the values the same remarks hold as in the previous paragraph.

5.2.1. Only one variable censored

An interesting case is when only one failure time is censored and the other one is not censored at all. We call this type of censoring "half censored". Let $\overline{F}$ be *uniform (0,1)* on the plane and only the second variable *uniform (0,1)* censored. We have independent margins so we can use (32) to detemine the asymptotic variance of the Dabrowska estimator.

$$VAR_{dabrowska}(t_1,t_2)=\frac{1}{2}(1-t_1)t_2(2-t_2)+(1-t_2)^2(1-t_1)t_1 \tag{73}$$

The pathwise estimator that first estimates the first margin is equal to the Dabrowska estimator, because for the second variable we have no censored observations and hence only the empirical distribution is taken. Because of the independent margins we know that this estimator is efficient.

For the Volterra estimator we have (using $\Delta_1\equiv 1$)

$$IC_{volterra}=$$
$$-(1_{\{X_1\le t_1\}}-t_1)$$
$$-(1-t_2)\left\{\frac{1_{\{X_2\le t_2,\Delta_2=1\}}}{(1-X_2)^2}-\frac{1}{2}\left(\frac{1}{(1-X_2\wedge t_2)^2}-1\right)\right\}$$
$$-\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\log\frac{1}{1-t_2}\right)^{k+1}\left\{\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\left(\log\frac{1}{1-t_1}\right)^i-(-1)^k(1-t_1)\right.$$
$$\left.+\frac{1}{(k+1)!}\left(\left(\log\frac{1-X_1\wedge t_1}{1-t_1}\right)^{k+1}-\left(\log\frac{1}{1-t_1}\right)^{k+1}\right)\right\}$$
$$-\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\log\frac{1}{1-t_1}\right)^{k+1}\Bigg\{$$

$$\frac{1}{2}\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\left(\log\frac{1}{1-t_2}\right)^i-(-1)^k(1-t_2)$$

$$-\frac{1}{2}\sum_{i=0}^{k}\frac{1}{i!}\left(\frac{\left(\log\frac{1-X_2\wedge t_2}{1-t_2}\right)^i}{1-X_2\wedge t_2}-\left(\log\frac{1}{1-t_2}\right)^i\right)$$

$$+1_{\{X_2\le t_2\}}\left(1_{\{\Delta_2=1\}}-\frac{1}{2}\right)\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\frac{\left(\log\frac{1-X_2}{1-t_2}\right)^i}{1-X_2}-(-1)^k\frac{1-t_2}{(1-X_2)^2}\Bigg\}$$

$$+\frac{1_{\{X_1\le t_1\}}\,1_{\{X_2\le t_2,\,\Delta_2=1\}}}{1-X_2}\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-X_1}{1-t_1}\right)^k\left(\log\frac{1-X_2}{1-t_2}\right)^k$$

$$-\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\left(\log\frac{1}{1-t_1}\right)^{k+1}-\left(\log\frac{1-X_1\wedge t_1}{1-t_1}\right)^{k+1}\right)$$

$$\sum_{i=0}^{k}\frac{1}{i!}\left(\frac{\left(\log\frac{1-X_2\wedge t_2}{1-t_2}\right)^i}{1-X_2\wedge t_2}-\left(\log\frac{1}{1-t_2}\right)^i\right) \tag{74}$$

We used this result for a Monte-Carlo approximation of the asymptotic R.M.S.E. with the influence curve. Table 10 shows of course that the Volterra estimator is worse than the Dabrowska estimator, but left below of the support the difference with the second pathwise estimator is to the advantage of the Volterra estimator.

5.2.2. Each variable totally censored or not at all

We now observe the situation where the failure time are *uniform (0,1)* distributed on the plane and possibly only one of the variables is available. The censor is defined as

$$(C_1,C_2)=\begin{cases}(1,1) & \text{with probability } p_1\\ (0,1) & \text{with probability } p_2\\ (1,0) & \text{with probability } p_3\end{cases} \tag{75}$$

We took $p_1=p_2=p_3=\frac{1}{3}$. The sample can be split into three sub-samples:

1. Observations not censored at all.
2. First variable not available.$((X_1,X_2,\Delta_1,\Delta_2)=(0,T_2,0,1))$.
3. Second variable not available.$((X_1,X_2,\Delta_1,\Delta_2)=(T_1,0,1,0))$.

Though we have no independence of the censor, we can determine the variance of the influence curve of the Dabrowska estimator exactly. We have

$$\mathrm{E}Y(t_1,t_2)=\begin{cases}\frac{1}{3}\overline{F}(t_1,t_2) & \text{for } t_1>0 \text{ and } t_2>0\\ \frac{2}{3}\overline{F}(t_1,t_2) & \text{for } t_1=0 \text{ and } t_2>0 \text{ or } t_1>0 \text{ and } t_2=0\\ \overline{F}(t_1,t_2) & \text{for } t_1=t_2=0\end{cases} \tag{76}$$

Let

$$IC_k(t_k) = -\overline{F}_k(t_k)\left(\frac{1_{\{X_k\le t_k,\Delta_k=1\}}}{\frac{2}{3}\overline{F}_k(X_k)} - \int\limits_{(0,X_k\wedge t_k]} \frac{\Lambda_k(ds_k)}{\frac{2}{3}\overline{F}_k(s_k)}\right) = -\frac{3}{2}\left(1_{\{X_k\le t_k,\Delta_k=1\}} - t_k\right) \quad (77)$$

be the marginal influence curves as in (30) and define

$$IC'_k(t_k) = -\overline{F}_k(t_k)\left(\frac{1_{\{X_k\le t_k,\Delta_k=1\}}}{\frac{1}{\sqrt{3}}\overline{F}_k(X_k)} - \int\limits_{(0,X_k\wedge t_k]} \frac{\Lambda_k(ds_k)}{\frac{1}{\sqrt{3}}\overline{F}_k(s_k)}\right) = -\sqrt{3}\left(1_{\{X_k\le t_k,\Delta_k=1\}} - t_k\right) \quad (78)$$

then

$$IC_{dabrowska} = \overline{F}_1\, IC_2 + \overline{F}_2\, IC_1 + IC'_1\, IC'_2 \quad (79)$$

Because of independence of IC_1 and IC_2 we get

$$\begin{aligned} VAR_{dabrowska}(t_1,t_2) &= (1-t_1)^2(1-t_2)^2\left\{\frac{3}{2}\left(\frac{1}{1-t_1}-1\right)+\frac{3}{2}\left(\frac{1}{1-t_2}-1\right)\right. \\ &\qquad \left. +3\left(\frac{1}{1-t_1}-1\right)\left(\frac{1}{1-t_2}-1\right)\right\} \\ &= \frac{3}{2}\,(1-t_1)\,(1-t_2)\,(t_1+t_2) \quad (80) \end{aligned}$$

So in this case the Dabrowska estimator is closer to efficiency than both pathwise estimators. One can argue by considering a dummy extra sample of size $\frac{n}{3}$ with censor $(C_1, C_2) = (0, 0)$ that the present problem is essential the same as the problem with $n\ \frac{4}{3}$ larger and all variables independent (the censoring variables equal to 0 or 1 with probability $\frac{1}{2}$), and therefor that the Dabrowska estimator is asymptotically efficient.

For the Volterra estimator the influence curve is

$$\begin{aligned} IC_{volterra}(t_1,t_2) &= \\ &-\frac{3}{2}\left(1_{\{X_1\le t_1,\Delta_1=1\}}-t_1\right)-\frac{3}{2}\left(1_{\{X_2\le t_2,\Delta_2=1\}}-t_2\right) \\ &-\frac{3}{2}\iint\limits_{s\le t}\left(1_{\{X_1\le s_1,\Delta_1=1\}}-s_1+1_{\{X_2\le s_2,\Delta_2=1\}}-s_2\right) \\ &\qquad \frac{\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-s_1}{1-t_1}\right)^k\left(\log\frac{1-s_2}{1-t_2}\right)^k}{(1-s_1)\,(1-s_2)}\,ds_1ds_2 \\ &+3\;1_{\{X_1\le t_1,\,\Delta_1=1\}}\;1_{\{X_2\le t_2,\,\Delta_2=1\}}\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-X_1}{1-t_1}\right)^k\left(\log\frac{1-X_2}{1-t_2}\right)^k \\ &-3\iint\limits_{(0,X\wedge t]}\frac{\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-s_1}{1-t_1}\right)^k\left(\log\frac{1-s_2}{1-t_2}\right)^k}{(1-s_1)\,(1-s_2)}\,ds_1ds_2 \end{aligned}$$

$$=$$

$$
\begin{aligned}
&-\frac{3}{2}\left(1_{\{X_1\le t_1,\Delta_1=1\}}-t_1\right)-\frac{3}{2}\left(1_{\{X_2\le t_2,\Delta_2=1\}}-t_2\right)\\
&-\frac{3}{2}\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\log\frac{1}{1-t_2}\right)^{k+1}\left\{\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\left(\log\frac{1}{1-t_1}\right)^{i}-(-1)^k(1-t_1)\right.\\
&\qquad\left.+\frac{1}{(k+1)!}\left(\left(\log\frac{1-X_1\wedge t_1}{1-t_1}\right)^{k+1}-\left(\log\frac{1}{1-t_1}\right)^{k+1}\right)\right\}\\
&-\frac{3}{2}\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\log\frac{1}{1-t_1}\right)^{k+1}\left\{\sum_{i=0}^{k}\frac{(-1)^{k-i}}{i!}\left(\log\frac{1}{1-t_2}\right)^{i}-(-1)^k(1-t_2)\right.\\
&\qquad\left.+\frac{1}{(k+1)!}\left(\left(\log\frac{1-X_2\wedge t_2}{1-t_2}\right)^{k+1}-\left(\log\frac{1}{1-t_2}\right)^{k+1}\right)\right\}\\
&+3\,1_{\{X_1\le t_1,\,\Delta_1=1\}}\,1_{\{X_2\le t_2,\,\Delta_2=1\}}\sum_{k=0}^{\infty}\left(\frac{1}{k!}\right)^2\left(\log\frac{1-X_1}{1-t_1}\right)^k\left(\log\frac{1-X_2}{1-t_2}\right)^k\\
&-3\sum_{k=0}^{\infty}\left(\frac{1}{(k+1)!}\right)^2\left\{\left(\log\frac{1}{1-t_1}\right)^{(k+1)}-\left(\log\frac{1-X_1\wedge t_1}{1-t_1}\right)^{(k+1)}\right\}\\
&\qquad\left\{\left(\log\frac{1}{1-t_2}\right)^{(k+1)}-\left(\log\frac{1-X_2\wedge t_2}{1-t_2}\right)^{(k+1)}\right\}ds_1ds_2 \qquad (81)
\end{aligned}
$$

Again we used this result for a Monte-Carlo approximation of the asymptotic R.M.S.E. with the influence curve. Table 11 shows that the Dabrowska estimator has the smallest R.M.S.E. and that the pathwise estimator is globally better than the Volterra estimator.

5.3. Further Pareto examples

In this paragraph we give some simulation results for Pareto distribution that are more complicated and hence only the sample variances for the Dabrowska and the Volterra estimator are given. This is mainly because the influence curve for the Pareto distribution with parameters unequal to 0, 1 and ∞ are difficult to handle for all three estimators. The simulations are done to check the difference between the Dabrowska and the Volterra estimator in some extra situations. We took $n = 100$ observations for each estimator. Simulations were done for 7 different cases with the reflected pareto(2) or the standard form pareto(2) distribution for respectively the failure times or the censors. Table 17 to table 23 show that the Dabrowska estimator has a smaller R.M.S.E. except for one or two points left below of the support, but the difference there can be caused by sample errors. We have no reason to worry about this.

6. Conclusions

All three estimators are not survival functions themselves, they are not monotone. The Volterra estimator only assigns mass to uncensored observations and at the upper/right region of the support the survival values can be negative. Hence this estimator is only reliable in the lower left region of the support. The pathwise estimator is path dependent and the two paths we considered can give quite different estimators. The Dabrowska estimator can have negative mass in certain points, but there are no negative survival values. In case both the failure and censoring times are independent the Dabrowska estimator is efficient.

In the lower left region of the support the difference between the Dabrowska, Volterra and pathwise estimator is rather small, there is no generally best estimator. When we reduce to the observed cases where we know some asymptotic properties, the Dabrowska and Volterra estimator are well matched here.

Further the overall impression is that the Dabrowska estimator is the best one.

Acknowledgements

This research was done in a project within the scope of civil services (TEG). I would like to thank professor Richard Gill for supervising this research.

References

[1] Becker, R.A., Chambers, J.M. and Wilks, A.R. (1988). *The new S Language*, Wadsworth & Brooks/Cole, California.

[2] Bickel, P.J., Gill, R.D. and Wellner, J.A. (1991). *Inefficient estimators for three multivariate models.* (in preparation).

[3] Campell, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika* **68**, 417–422.

[4] Campell and Földes (1982). Large-sample properties of nonparametric bivariate estimators with censored data. *Nonparametric Statistical Inference* (B.V. Gnedenko, M.L. Puri and I. Vincze, eds.)103–122, North-Holland, Amsterdam.

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

[6] Dabrowska, D.M. (1989). Kaplan-Meier estimate on the plane: Weak Convergence, LIL, and the Bootstrap. *J. Multivar. Anal.* **29**, 308–325.

[7] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood estimation for incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B* **39**, 1–38.

[8] Gill, R.D. (1989). Non- and Semi-parametric Maximum Likelihood Estimators and the von Mises Method (Part1). *Scand. J. Statist.* **16**, 97-128.

[9] Gill, R.D. (1990). *Multivariate survival analysis.* Technical Report, Department of Mathematics, University of Utrecht. To appear in *Proceedings 2nd World Congres of the Bernoulli Society* (Uppsala, 1990).

[10] Gill, R.D. and Johansen, S. (1990). A survey of product-integration with a view towards application in survival analysis. To appear in *Ann. Statist.*.

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

[12] Johnson, M.E. (1987). *Multivariate Statistical Simulation*, Wiley, New York.

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

[14] Laan, M. van der (1990). *Dabrowska's multivariate product limit estimator and the delta-method.* Master's thesis, Department of Mathematics, University of Utrecht.

[15] Muñoz, A. (1981). *Nonparametric estimation from censored bivariate observations.* Technical Report No. 60, Division of Biostatistics, Stanford University, Stanford California.

[16] Pruitt, R.C. (1991). On negative mass assigned by the bivariate Kaplan-Meier estimator. To appear in *Ann. Statist.*.

A. Tables

Table 1.

The asymptotic R.M.S.E. for the pathwise, Dabrowska and Volterra estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with both failure time and censor *uniform (0,1)* on the diagonal.

point	R.M.S.E.					
	$\lrcorner$ pathwise	$\ulcorner$	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) exact
(0.1,0.1)	0.0308221	0.0308221	0.0328019	0.0308221	0.0328019	0.0308221
(0.25,0.25)	0.0467707	0.0467707	0.0408280	0.0467707	0.0408280	0.0467707
(0.25,0.5)	0.0612372	0.0612372	0.0552978	0.0612372	0.0552978	0.0612372
(0.25,0.75)	0.0684653	0.0684653	0.0670722	0.0684653	0.0670722	0.0684653
(0.5,0.5)	0.0612372	0.0612372	0.0552978	0.0612372	0.0552978	0.0612372
(0.5,0.75)	0.0684653	0.0684653	0.0670722	0.0684653	0.0670722	0.0684653
(0.75,0.75)	0.0684653	0.0684653	0.0670722	0.0684653	0.0670722	0.0684653
(0.1,0.9)	0.0703562	0.0703562	0.0834965	0.0703562	0.0834965	0.0703562
(0.9,0.9)	0.0703562	0.0703562	0.0834965	0.0703562	0.0834965	0.0703562
(0.95,0.95)	0.0706222	0.0706222	0.0740182	0.0706222	0.0740182	0.0706222

Table 2.

The asymptotic R.M.S.E. for the pathwise estimator and Monte-Carlo approximations based on 10,000 replicates of the influence curves for the asymptotic R.M.S.E. and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and the Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the diagonal and the censor standard form *pareto(1)*.

point	R.M.S.E.					
	$\lrcorner$ pathwise	$\ulcorner$	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo
(0.1,0.1)	0.0308221	0.0308221	0.0305718	0.0298748	0.0305824	0.0298798
(0.25,0.25)	0.0467707	0.0467707	0.0450679	0.0449520	0.0451453	0.0451025
(0.25,0.5)	0.0641138	0.0612372	0.0592392	0.0604427	0.0588210	0.0609660
(0.25,0.75)	0.0715574	0.0684653	0.0681470	0.0684656	0.0687092	0.0695382
(0.5,0.5)	0.0612372	0.0612372	0.0573428	0.0569028	0.0602289	0.0608765
(0.5,0.75)	0.0793733	0.0684653	0.0660694	0.0677901	0.0719360	0.0756474
(0.75,0.75)	0.0684653	0.0684653	0.0915286	0.0693847	0.1405326	0.1202944
(0.1,0.9)	0.0707900	0.0703562	0.0714082	0.0672429	0.0718236	0.0674189
(0.9,0.9)	0.0703562	0.0703562	0.0741746	0.1047518	0.2145584	0.3750294
(0.95,0.95)	0.0706222	0.0706222	0.0535545	0.1676014	0.2484845	0.8308576

Table 3.

The asymptotic R.M.S.E. for the pathwise estimator and Monte-Carlo approximations based on 10,000 replicates of the influence curves for the asymptotic R.M.S.E. and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the diagonal and the censor *uniform (0,1)* on the plane.

point	R.M.S.E.					
	⌟ pathwise	⌜	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo
(0.1,0.1)	0.0308221	0.0308221	0.0295686	0.0294135	0.0295495	0.0294026
(0.25,0.25)	0.0467707	0.0467707	0.0421520	0.0434707	0.0421770	0.0437353
(0.25,0.5)	0.0683807	0.0612372	0.0629758	0.0602888	0.0632275	0.0611861
(0.25,0.75)	0.0785429	0.0684653	0.0727365	0.0676739	0.0750704	0.0694945
(0.5,0.5)	0.0612372	0.0612372	0.0597765	0.0574172	0.0677167	0.0666544
(0.5,0.75)	0.0918559	0.0684653	0.0770177	0.0670307	0.0983611	0.0832186
(0.75,0.75)	0.0684653	0.0684653	0.0886146	0.0854638	0.1730119	0.1750662
(0.1,0.9)	0.0741532	0.0703562	0.0741568	0.0691498	0.0744120	0.0694292
(0.9,0.9)	0.0703562	0.0703562	0.0796267	0.1613423	0.2771672	0.5501516
(0.95,0.95)	0.0706222	0.0706222	0.0454243	0.1885851	0.2899816	1.1616326

Table 4.

The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time reflected *pareto(1)* and the censor *uniform (0,1)* on the diagonal.

point	R.M.S.E.					
	⌟ pathwise	⌜	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.)	($n = 100$) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0401444	0.0401444	0.0410654		0.0411072	
(0.25,0.25)	0.0547723	0.0547723	0.0461243		0.0463370	
(0.25,0.5)	0.0597614	0.0630840	0.0571160		0.0572542	
(0.25,0.75)	0.0650444	0.0685868	0.0625598		0.0668589	
(0.5,0.5)	0.0623610	0.0623610	0.0588895		0.0596498	
(0.5,0.75)	0.0616441	0.0678233	0.0539600		0.0682924	
(0.75,0.75)	0.0630840	0.0630840	0.0557382		0.0831198	
(0.1,0.9)	0.0696698	0.0703597	0.0680734		0.0707936	
(0.9,0.9)	0.0621632	0.0621632	0.0545474		0.1875752	
(0.95,0.95)	0.0617251	0.0617251	0.0439521		0.2383120	

Table 5.
The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at n = 100 with the failure time reflected *pareto(1)* and the censor standard form *pareto(1)*.

point	R.M.S.E.					
	┘ pathwise	┌	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	(n = 100) Monte-Carlo	(asympt.)	(n = 100) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0403687	0.0403687	0.0418284		0.0417873	
(0.25,0.25)	0.0557493	0.0557493	0.0551789		0.0551997	
(0.25,0.5)	0.0623814	0.0634992	0.0676258		0.0694918	
(0.25,0.75)	0.0679387	0.0687496	0.0709856		0.0770077	
(0.5,0.5)	0.0660606	0.0660606	0.0682871		0.0733942	
(0.5,0.75)	0.0707327	0.0697674	0.0685592		0.0905841	
(0.75,0.75)	0.0752139	0.0752139	0.0642377		0.1111526	
(0.1,0.9)	0.0701001	0.0703627	0.0765468		0.0794417	
(0.9,0.9)	0.0961529	0.0961529	0.0561447		0.2104226	
(0.95,0.95)	0.1241625	0.1241625	0.0425570		0.2387745	

Table 6.
The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at n = 100 with the failure time reflected *pareto(1)* and the censor *uniform (0,1)* on the plane.

point	R.M.S.E.					
	┘ pathwise	┌	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	(n = 100) Monte-Carlo	(asympt.)	(n = 100) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0407122	0.0407122	0.0413615		0.0413788	
(0.25,0.25)	0.0571548	0.0571548	0.0531977		0.0527985	
(0.25,0.5)	0.0662401	0.0646813	0.0603929		0.0618579	
(0.25,0.75)	0.0743811	0.0694441	0.0641205		0.0673961	
(0.5,0.5)	0.0707107	0.0707107	0.0687160		0.0710403	
(0.5,0.75)	0.0812404	0.0734847	0.0716969		0.0773240	
(0.75,0.75)	0.0874818	0.0874818	0.0846082		0.1399557	
(0.1,0.9)	0.0734252	0.0704026	0.0723414		0.0759002	
(0.9,0.9)	0.1228133	0.1228133	0.0478119		0.2591158	
(0.95,0.95)	0.1659650	0.1659650	0.0346000		0.2990478	

Table 7.
The asymptotic R.M.S.E. for the pathwise, a Monte-Carlo aproximation based on 10,000 replicates of the influence curve for the asymptotic R.M.S.E. for the Dabrowska and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the plane and the censor *uniform (0,1)* on the diagonal.

point	R.M.S.E.					
	⌋ pathwise ⌈		Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo	($n = 100$) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0408595	0.0408595	0.0391877	0.0405560	0.0390977	
(0.25,0.25)	0.0557267	0.0557267	0.0513476	0.0537047	0.0523692	
(0.25,0.5)	0.0588519	0.0631219	0.0635771	0.0578878	0.0647374	
(0.25,0.75)	0.0606497	0.0671693	0.0485636	0.0590988	0.0568315	
(0.5,0.5)	0.0586302	0.0586302	0.0551444	0.0537497	0.0576988	
(0.5,0.75)	0.0522913	0.0605960	0.0467915	0.0492616	0.0657551	
(0.75,0.75)	0.0465615	0.0465615	0.0389483	0.0384572	0.1111473	
(0.1,0.9)	0.0668211	0.0700678	0.0715906	0.0648160	0.0802311	
(0.9,0.9)	0.0308140	0.0308140	0.0249335	0.0238843	0.3115999	
(0.95,0.95)	0.0220787	0.0220787	0.0194498	0.0169662	0.3607625	

Table 8.
The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and the Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the plane and the censor standard form *pareto(1)*.

point	R.M.S.E.					
	⌋ pathwise ⌈		Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.)	($n = 100$) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0410995	0.0410995	0.0441409		0.0440277	
(0.25,0.25)	0.0568508	0.0568508	0.0580810		0.0571957	
(0.25,0.5)	0.0615647	0.0637562	0.0646488		0.0643087	
(0.25,0.75)	0.0633812	0.0676076	0.0625320		0.0715206	
(0.5,0.5)	0.0630238	0.0630238	0.0641659		0.0677975	
(0.5,0.75)	0.0606161	0.0639518	0.0579246		0.0736529	
(0.75,0.75)	0.0587566	0.0587566	0.0555035		0.1118467	
(0.1,0.9)	0.0672339	0.0700928	0.0773074		0.0853678	
(0.9,0.9)	0.0537781	0.0537781	0.0302946		0.2242872	
(0.95,0.95)	0.0519050	0.0519050	0.0183309		0.2508435	

Table 9.

The asymptotic R.M.S.E. for the pathwise and Dabrowska estimator, a Monte-Carlo aproximation based on 10,000 replicates of the influence curve for the asymptotic R.M.S.E. for the Volterra estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with both failure and censoring times *uniform (0,1)* on the plane.

point	R.M.S.E.					
	┘ pathwise	┌	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo
(0.1,0.1)	0.0414669	0.0414669	0.0437168	0.0403640	0.0436509	0.0409183
(0.25,0.25)	0.0584634	0.0584634	0.0516870	0.0542167	0.0525411	0.0546217
(0.25,0.5)	0.0655506	0.0655506	0.0565295	0.0589624	0.0605222	0.0609242
(0.25,0.75)	0.0694566	0.0694566	0.0591066	0.0616346	0.0699101	0.0706484
(0.5,0.5)	0.0684653	0.0684653	0.0633688	0.0572822	0.0730109	0.0661928
(0.5,0.75)	0.0701561	0.0701561	0.0652025	0.05625	0.0914027	0.0899154
(0.75,0.75)	0.0705724	0.0705724	0.0517549	0.0527561	0.1544918	0.1601669
(0.1,0.9)	0.0704237	0.0704237	0.0707438	0.0670019	0.0803102	0.0744061
(0.9,0.9)	0.0707071	0.0707071	0.0262859	0.0504901	0.2868134	0.7446367
(0.95,0.95)	0.0707105	0.0707105	0.0236656	0.0501244	0.3181672	2.5014173

Table 10.

The asymptotic R.M.S.E. for the pathwise and Dabrowska estimator, a Monte-Carlo aproximation based on 10,000 replicates of the influence curve for the asymptotic R.M.S.E. for the Volterra estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the plane and only the second variable *uniform (0,1)* censored.

point	R.M.S.E.					
	┘ pathwise	┌	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo
(0.1,0.1)	0.0397995	0.0408595	0.0383249	0.0397995	0.0382736	0.0388536
(0.25,0.25)	0.0519164	0.0557267	0.0495453	0.0519164	0.0486460	0.0520929
(0.25,0.5)	0.0572822	0.0631219	0.0539914	0.0572822	0.0554076	0.0589948
(0.25,0.75)	0.0602728	0.0671693	0.0625560	0.0602728	0.0699937	0.0680843
(0.5,0.5)	0.05	0.0586302	0.0462747	0.05	0.0487313	0.0542855
(0.5,0.75)	0.05	0.0605960	0.0542853	0.05	0.0747051	0.0726043
(0.75,0.75)	0.0359035	0.0465615	0.0370618	0.0359035	0.1089888	0.0918749
(0.1,0.9)	0.0668132	0.0700678	0.0679091	0.0668132	0.0723438	0.0740657
(0.9,0.9)	0.0224499	0.0308140	0.0177646	0.0224499	0.2703568	0.4066462
(0.95,0.95)	0.0158292	0.0220787	0.0128752	0.0158292	0.3285604	1.3972977

Table 11.

The asymptotic R.M.S.E. for the pathwise and Dabrowska estimator, a Monte-Carlo aproximation based on 10,000 replicates of the influence curve for the asymptotic R.M.S.E. for the Volterra estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the plane and for each observation possibly only one variable is available.

point	R.M.S.E.					
	⌋ pathwise ⌈		Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) exact	($n = 100$) Monte-Carlo	(asympt.) Monte-Carlo
(0.1,0.1)	0.0593591	0.0593591	0.0495824	0.0492950	0.0496576	0.0503493
(0.25,0.25)	0.0761629	0.0761629	0.0630963	0.0649519	0.0694947	0.0721337
(0.25,0.5)	0.0795495	0.0701561	0.0730760	0.0649519	0.0825229	0.0820727
(0.25,0.75)	0.0662913	0.0546652	0.0582809	0.0530330	0.0840098	0.0922902
(0.5,0.5)	0.0684653	0.0684653	0.0737422	0.0612372	0.0977803	0.0945832
(0.5,0.75)	0.0551985	0.0507752	0.0591304	0.0484123	0.1175243	0.1308064
(0.75,0.75)	0.0397748	0.0397748	0.0433315	0.0375000	0.2209726	0.2628665
(0.1,0.9)	0.0494318	0.0369256	0.0373183	0.0367423	0.0574796	0.0891312
(0.9,0.9)	0.0168375	0.0168375	0.0240041	0.0164317	0.7220880	1.4714692
(0.95,0.95)	0.0085458	0.0085458	0.0077202	0.0084410	1.0838144	5.2993375

Table 12.

The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 200$ with the failure time reflected *pareto(1)* and the censor *uniform (0,1)* on the diagonal.

point	R.M.S.E.					
	⌋ pathwise ⌈		Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	(sample) Monte-Carlo	(asympt.)	(sample) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0283864	0.0283864	0.0240472		0.0241057	
(0.25,0.25)	0.0387298	0.0387298	0.0362423		0.0361822	
(0.25,0.5)	0.0422577	0.0446071	0.0420940		0.0420528	
(0.25,0.75)	0.0459933	0.0484982	0.0372613		0.0415213	
(0.5,0.5)	0.0440959	0.0440959	0.0427027		0.0434422	
(0.5,0.75)	0.0435890	0.0479583	0.0393337		0.0484261	
(0.75,0.75)	0.0446071	0.0446071	0.0398002		0.0634898	
(0.1,0.9)	0.0492640	0.0497518	0.0569258		0.0595901	
(0.9,0.9)	0.0439560	0.0439560	0.0394315		0.1341252	
(0.95,0.95)	0.0436463	0.0436463	0.0389522		0.1873497	

Table 13.
The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 200$ with the failure time reflected *pareto(1)* and the censor standard form *pareto(1)*.

point	R.M.S.E.					
	⌟ pathwise ⌜		Dabrowska		Volterra	
	(asympt.)	(asympt.)	(sample)	(asympt.)	(sample)	(asympt.)
	exact	exact	Monte-Carlo		Monte-Carlo	
(0.1,0.1)	0.0285450	0.0285450	0.0264842		0.0264649	
(0.25,0.25)	0.0394207	0.0394207	0.0377049		0.0378795	
(0.25,0.5)	0.0441103	0.0449007	0.0364561		0.0379483	
(0.25,0.75)	0.0480399	0.0486133	0.0498390		0.0539717	
(0.5,0.5)	0.0467119	0.0467119	0.0367931		0.0409497	
(0.5,0.75)	0.0500156	0.0493330	0.0455671		0.0579680	
(0.75,0.75)	0.0531843	0.0531843	0.0439290		0.0837604	
(0.1,0.9)	0.0495683	0.0497540	0.0540859		0.0557963	
(0.9,0.9)	0.0679903	0.0679903	0.0556126		0.1963457	
(0.95,0.95)	0.0877962	0.0877962	0.0382919		0.2528487	

Table 14.
The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 200$ with the failure time reflected *pareto(1)* and the censor *uniform (0,1)* on the plane.

point	R.M.S.E.					
	⌟ pathwise ⌜		Dabrowska		Volterra	
	(asympt.)	(asympt.)	(sample)	(asympt.)	(sample)	(asympt.)
	exact	exact	Monte-Carlo		Monte-Carlo	
(0.1,0.1)	0.0287879	0.0287879	0.0296525		0.0296065	
(0.25,0.25)	0.0404145	0.0404145	0.0376832		0.0383576	
(0.25,0.5	0.0468389	0.0457366	0.0444936		0.0455679	
(0.25,0.75)	0.0525954	0.0491044	0.0430214		0.0472636	
(0.5,0.5)	0.0500000	0.0500000	0.0394243		0.0448534	
(0.5,0.75)	0.0574456	0.0519615	0.0500081		0.0574751	
(0.75,0.75)	0.0618590	0.0618590	0.0602373		0.1138021	
(0.1,0.9)	0.0519194	0.0497822	0.0526800		0.0531375	
(0.9,0.9)	0.0868421	0.0868421	0.0437484		0.2525346	
(0.95,0.95)	0.1173550	0.1173550	0.0282054		0.2893231	

Table 15.
The asymptotic R.M.S.E. for the pathwise, a Monte-Carlo aproximation based on 20,000 replicates of the influence curve for the asymptotic R.M.S.E. for the Dabrowska and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 200$ with the failure time *uniform (0,1)* on the plane and the censor *uniform (0,1)* on the diagonal.

point	R.M.S.E.					
	$\lrcorner$ pathwise	$\ulcorner$	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	(sample) Monte-Carlo	(asympt.) Monte-Carlo	(sample) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0288920	0.0288920	0.0281462	0.0284254	0.0281424	
(0.25,0.25)	0.0394048	0.0394048	0.0392589	0.0380461	0.0397000	
(0.25,0.5)	0.0416146	0.0446339	0.0456363	0.0409732	0.0459778	
(0.25,0.75)	0.0428858	0.0474959	0.0457782	0.0428666	0.0508904	
(0.5,0.5)	0.0414578	0.0414578	0.0423457	0.0380998	0.0433250	
(0.5,0.75)	0.0369755	0.0428478	0.0399970	0.0360820	0.0499239	
(0.75,0.75)	0.0329239	0.0329239	0.0319424	0.0289301	0.0669069	
(0.1,0.9)	0.0472496	0.0495454	0.0487337	0.0470067	0.0538993	
(0.9,0.9)	0.0217888	0.0217888	0.0161159	0.0176462	0.1969951	
(0.95,0.95)	0.0156120	0.0156120	0.0095985	0.0078995	0.3101413	

Table 16.
The asymptotic R.M.S.E. for the pathwise estimator and a sample R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and the Volterra estimator, at $n = 200$ with the failure time *uniform (0,1)* on the plane and the censor standard form *pareto(1)*.

point	R.M.S.E.					
	$\lrcorner$ pathwise	$\ulcorner$	Dabrowska		Volterra	
	(asympt.) exact	(asympt.) exact	(sample) Monte-Carlo	(asympt.)	(sample) Monte-Carlo	(asympt.)
(0.1,0.1)	0.0290617	0.0290617	0.0295880		0.0296329	
(0.25,0.25)	0.0401996	0.0401996	0.0342606		0.0343319	
(0.25,0.5)	0.0435328	0.0450824	0.0375625		0.0373295	
(0.25,0.75)	0.0448172	0.0478058	0.0397823		0.0437293	
(0.5,0.5)	0.0445646	0.0445646	0.0370021		0.0392714	
(0.5,0.75)	0.0428620	0.0452207	0.0354845		0.0463241	
(0.75,0.75)	0.0415472	0.0415472	0.0349614		0.0778889	
(0.1,0.9)	0.0475416	0.0495631	0.0490007		0.0521846	
(0.9,0.9)	0.0380268	0.0380268	0.0181355		0.1966464	
(0.95,0.95)	0.0367023	0.0367023	0.0110911		0.2722998	

Table 17.

The value of the original survival function and a sample mean and R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the diagonal and the censor standard form *pareto(2)*.

point	value	mean		R.M.S.E.	
		Dabrowska	Volterra	Dabrowska	Volterra
(0.1,0.1)	0.9	0.9011501	0.9011374	0.0321796	0.0321900
(0.25,0.25)	0.75	0.7543970	0.7541872	0.0444008	0.0445962
(0.25,0.5)	0.5	0.5005159	0.5001874	0.0643834	0.0646834
(0.25,0.75)	0.25	0.2459276	0.2453328	0.0686902	0.0706007
(0.5,0.5)	0.5	0.5038301	0.5023479	0.0584540	0.0660271
(0.5,0.75)	0.25	0.2474122	0.2474933	0.0674257	0.0802283
(0.75,0.75)	0.25	0.2448149	0.2374048	0.0757030	0.1575409
(0.1,0.9)	0.1	0.1004546	0.0998269	0.0852298	0.0851620
(0.9,0.9)	0.1	0.0564292	0.0180435	0.0767115	0.2921192
(0.95,0.95)	0.05	0.0247046	-0.0513217	0.0499733	0.3086483

Table 18.

The value of the original survival function and a sample mean and R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time reflected *pareto(1)* and the censor standard form *pareto(2)*.

point	value	mean		R.M.S.E.	
		Dabrowska	Volterra	Dabrowska	Volterra
(0.1,0.1)	0.8181818	0.8256554	0.8256496	0.0433832	0.0433341
(0.25,0.25)	0.6	0.6083120	0.6089496	0.0550482	0.0549779
(0.25,0.5)	0.4285714	0.4398513	0.4402017	0.0593391	0.0603231
(0.25,0.75)	0.2307692	0.2459512	0.2457888	0.0668850	0.0726953
(0.5,0.5)	0.3333333	0.3464694	0.3476341	0.0616784	0.0668986
(0.5,0.75)	0.2	0.2099229	0.2123476	0.0610461	0.0766759
(0.75,0.75)	0.1428571	0.1385896	0.1253656	0.0687254	0.1058538
(0.1,0.9)	0.0989011	0.1203763	0.1204678	0.0753924	0.0816812
(0.9,0.9)	0.0526316	0.0362228	0.0011695	0.0485571	0.2025523
(0.95,0.95)	0.0256410	0.0153462	-0.0596611	0.0310503	0.2425231

Table 19.
The value of the original survival function and a sample mean and R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time reflected *pareto(2)* and the censor *uniform (0,1)* on the diagonal.

point	value	mean		R.M.S.E.	
		Dabrowska	Volterra	Dabrowska	Volterra
(0.1,0.1)	0.8142830	0.8180307	0.8180670	0.0373303	0.0373413
(0.25,0.25)	0.5832498	0.5834260	0.5838993	0.0605424	0.0604360
(0.25,0.5)	0.4062578	0.4006903	0.4019384	0.0576831	0.0578858
(0.25,0.75)	0.2153903	0.2210833	0.2249229	0.0658990	0.0710898
(0.5,0.5)	0.2991195	0.2956831	0.2990542	0.0590553	0.0642720
(0.5,0.75)	0.1715729	0.1781669	0.1842296	0.0571572	0.0714394
(0.75,0.75)	0.1111111	0.1143823	0.1201757	0.0525719	0.0979482
(0.1,0.9)	0.0966647	0.0968424	0.0966074	0.0749391	0.0799835
(0.9,0.9)	0.0352723	0.0289738	0.0486180	0.0449675	0.2344022
(0.95,0.95)	0.0158450	0.0146517	0.0253358	0.0319983	0.2701099

Table 20.
The value of the original survival function and a sample mean and R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time reflected *pareto(2)* and the censor standard form *pareto(1)*.

point	value	mean		R.M.S.E.	
		Dabrowska	Volterra	Dabrowska	Volterra
(0.1,0.1)	0.8142830	0.8162466	0.8162224	0.0362547	0.0361938
(0.25,0.25)	0.5832498	0.5806072	0.5801993	0.0501607	0.0505482
(0.25,0.5)	0.4062578	0.4039866	0.4023286	0.0534613	0.0559891
(0.25,0.75)	0.2153903	0.2137610	0.2129378	0.0628598	0.0711738
(0.5,0.5)	0.2991195	0.2974321	0.2945717	0.0587625	0.0655085
(0.5,0.75)	0.1715729	0.1691549	0.1710580	0.0647637	0.0854180
(0.75,0.75)	0.1111111	0.1045570	0.0915967	0.0595330	0.1381718
(0.1,0.9)	0.0966647	0.0940755	0.0960289	0.0720368	0.0771762
(0.9,0.9)	0.0352723	0.0315438	-0.0680891	0.0442774	0.2556503
(0.95,0.95)	0.0158450	0.0148783	-0.1252081	0.0336343	0.2867076

Table 21.

The value of the original survival function and a sample mean and R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time reflected *pareto(2)* and the censor standard form *pareto(2)*.

point	value	mean		R.M.S.E.	
		Dabrowska	Volterra	Dabrowska	Volterra
(0.1,0.1)	0.8142830	0.8183227	0.8184112	0.0438194	0.0438057
(0.25,0.25)	0.5832498	0.5840472	0.5846069	0.0571790	0.0587351
(0.25,0.5)	0.4062578	0.3966251	0.3960798	0.0640180	0.0646286
(0.25,0.75)	0.2153903	0.2155662	0.2128092	0.0621349	0.0695874
(0.5,0.5)	0.2991195	0.2901228	0.2924964	0.0642540	0.0678191
(0.5,0.75)	0.1715729	0.1766933	0.1799243	0.0603300	0.0854874
(0.75,0.75)	0.1111111	0.1119903	0.1111358	0.0625691	0.1373569
(0.1,0.9)	0.0966647	0.1022695	0.0991087	0.0655302	0.0761984
(0.9,0.9)	0.0352723	0.0324053	-0.0213482	0.0500051	0.2476228
(0.95,0.95)	0.0158450	0.0150031	-0.0752484	0.0365538	0.2625416

Table 22.

The value of the original survival function and a sample mean and R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time reflected *pareto(2)* and the censor *uniform (0,1)* on the plane.

point	value	mean		R.M.S.E.	
		Dabrowska	Volterra	Dabrowska	Volterra
(0.1,0.1)	0.8142830	0.8194698	0.8195789	0.0440637	0.0438937
(0.25,0.25)	0.5832498	0.5924665	0.5927925	0.0532467	0.0540572
(0.25,0.5)	0.4062578	0.4023690	0.4000212	0.0628660	0.0632445
(0.25,0.75)	0.2153903	0.2220764	0.2162220	0.0722247	0.0773893
(0.5,0.5)	0.2991195	0.2983966	0.2968403	0.0580041	0.0664928
(0.5,0.75)	0.1715729	0.1742516	0.1638655	0.0719212	0.0912879
(0.75,0.75)	0.1111111	0.1133653	0.1041699	0.0747782	0.1605963
(0.1,0.9)	0.0966647	0.0955914	0.0925985	0.0707866	0.0777277
(0.9,0.9)	0.0352723	0.0230414	-0.0886554	0.0400927	0.2706998
(0.95,0.95)	0.0158450	0.0079825	-0.1418894	0.0254986	0.3023783

Table 23.
The value of the original survival function and a sample mean and R.M.S.E. based on 100 replicates of the estimators for the Dabrowska and Volterra estimator, at $n = 100$ with the failure time *uniform (0,1)* on the plane and the censor standard form *pareto(2)*.

point	value	mean		R.M.S.E.	
		Dabrowska	Volterra	Dabrowska	Volterra
(0.1,0.1)	0.81	0.8091810	0.8092784	0.0426959	0.0426176
(0.25,0.25)	0.5625	0.5599269	0.5599475	0.0554576	0.0564913
(0.25,0.5)	0.375	0.3766229	0.3762228	0.0634749	0.0651543
(0.25,0.75)	0.1875	0.1809602	0.1808382	0.0649699	0.0714940
(0.5,0.5)	0.25	0.2511119	0.2517799	0.0557903	0.0625102
(0.5,0.75)	0.125	0.1088948	0.1154940	0.0628718	0.0871055
(0.75,0.75)	0.0625	0.0601313	0.0475202	0.0522410	0.1332206
(0.1,0.9)	0.09	0.0923647	0.0893200	0.0697708	0.0741077
(0.9,0.9)	0.01	0.0111654	-0.0740299	0.0247417	0.2555912
(0.95,0.95)	0.0025	0.0063879	-0.1294346	0.0209569	0.3230628

B. The variance of the influence function of the Pathwise estimator

The variance of the pathwise estimator is given in (45). For the one dimensional case with both the failure and the censor time *uniform (0,1)* distributed we have

$$\int_{(0,t]} \frac{\Lambda(ds)}{\mathrm{E}Y(s)} = \frac{1}{2}\left(\frac{1}{(1-t)^2} - 1\right) \tag{1}$$

Next we give the variance of some special bivariate cases.

B.1. Pareto $a = 0,\ b = 0$

We have $\overline{F} = \overline{G} = \min\{1-t_1, 1-t_2\}$. It is clear that in this case we get the Kaplan-Meier estimate. For $t_1 \le s_2$ we have

$$\overline{F}(t_1, ds_2) = -ds_2 \tag{2}$$

this gives for $t_1 \le t_2$

$$\int_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} = \int_{(t_1,t_2]} \frac{ds_2}{(1-s_2)^3} = \frac{1}{2}\left(\frac{1}{(1-t_2)^2} - \frac{1}{(1-t_1)^2}\right) \tag{3}$$

and

$$VAR_{pathwise} = \frac{1}{2}(1 - (1-t_2)^2) \tag{4}$$

For $t_1 > t_2$

$$VAR_{pathwise} = \frac{1}{2}(1 - (1-t_1)^2) \tag{5}$$

This is the same as the variance of the Dabrowska and the Volterra estimator.

B.2. Pareto $a = 0,\ b = 1$

$\overline{F}(t_1, t_2) = \min\{1-t_1, 1-t_2\}$ and

$$\overline{G}(t_1, t_2) = \frac{(t_1+t_2)(1-t_1)(1-t_2)}{t_1+t_2-t_1t_2} \tag{6}$$

For $t_1 > t_2$

$$VAR_{pathwise} = \frac{1}{2}\left(1 - (1-t_1)^2\right) \tag{7}$$

For $t_1 \le t_2$

$$\begin{aligned}
\int_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} &= \int_{(t_1,t_2]} \frac{t_1+s_2-t_1s_2}{(t_1+s_2)(1-t_1)(1-s_2)^3} ds_2 \\
&= \frac{1}{2\,(1-t_1)(1+t_1)}\left(\frac{1}{(1-t_2)^2} - \frac{1}{(1-t_1)^2}\right) \\
&\quad + \frac{t_1^2}{(1-t_1)(1+t_1)^2}\left(\frac{1}{1-t_2} - \frac{1}{1-t_1}\right) \\
&\quad + \frac{t_1^2}{(1-t_1)(1+t_1)^3}\left(\log\frac{t_1+t_2}{2t_1} - \log\frac{1-t_2}{1-t_1}\right)
\end{aligned} \tag{8}$$

and hence

$$VAR_{pathwise} =$$
$$\frac{1}{2}\left\{\left(\frac{1-t_2}{1-t_1}\right)^2 - (1-t_2)^2\right\} + \frac{1}{2\,(1-t_1)(1+t_1)}\left\{1-\left(\frac{1-t_2}{1-t_1}\right)^2\right\}$$
$$+\frac{t_1^2(1-t_2)}{(1-t_1)(1+t_1)^2}\left(1-\frac{1-t_2}{1-t_1}\right)+\frac{t_1^2(1-t_2)^2}{(1-t_1)(1+t_1)^3}\log\frac{(1-t_1)(t_1+t_2)}{2t_1(1-t_2)} \tag{9}$$

B.3. Pareto $a = 0$, $b = \infty$

$\overline{F}(t_1, t_2) = \min\{1-t_1, 1-t_2\}$ and $\overline{G}$ *uniform (0,1)*. We have to distinguish $t_1 > t_2$ and $t_1 \leq t_2$. For $t_1 > s_2$ we have

$$\overline{F}(t_1, ds_2) = 0 \tag{10}$$

and hence for $t_1 > t_2$

$$VAR_{pathwise} = \frac{1}{2}\left(1-(1-t_1)^2\right) \tag{11}$$

For $t_1 \leq t_2$

$$\int_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} = \int_{(t_1,t_2]} \frac{ds_2}{(1-s_2)^2(1-t_1)(1-s_2)}$$
$$= \frac{1}{2\,(1-t_1)}\left(\frac{1}{(1-t_2)^2}-\frac{1}{(1-t_1)^2}\right) \tag{12}$$

this gives

$$VAR_{pathwise} = (1-t_2)^2\left\{\frac{1}{2}\left(\frac{1}{(1-t_1)^2}-1\right)+\frac{1}{2\,(1-t_1)}\left(\frac{1}{(1-t_2)^2}-\frac{1}{(1-t_1)^2}\right)\right\}$$
$$= \frac{1}{2}\left\{\left(\frac{1-t_2}{1-t_1}\right)^2-(1-t_2)^2+\frac{1}{1-t_1}\left(1-\left(\frac{1-t_2}{1-t_1}\right)^2\right)\right\} \tag{13}$$

B.4. Pareto $a = 1$, $b = 0$

We have

$$\overline{F}(t_1, t_2) = \frac{(1-t_1)(1-t_2)}{(1-t_1t_2)} \tag{14}$$

and $\overline{G} = \min\{1-t_1, 1-t_2\}$. Then

$$\int_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} = 1_{\{t_1>t_2\}}\int_{(0,t_2]} \frac{ds_2}{(1-t_1)(1-s_2)^2}$$
$$+\,1_{\{t_1\leq t_2\}}\left\{\int_{(0,t_1]}\frac{ds_2}{(1-t_1)(1-s_2)^2}+\int_{(t_1,t_2]}\frac{ds_2}{(1-s_2)^3}\right\} \tag{15}$$

this gives

$$VAR_{pathwise} =$$

$$\frac{1}{(1-t_1t_2)^2}\left\{\frac{1}{2}t_1(2-t_1)(1-t_2)^2 + 1_{\{t_1>t_2\}}(1-t_1)t_2(1-t_2)\right.$$

$$\left. + 1_{\{t_1\le t_2\}}\left(t_1(1-t_2)^2 + \frac{1}{2}\left((1-t_1)^2 - (1-t_2)^2\right)\right)\right\} \tag{16}$$

B.5. Pareto $a = 1$, $b = 1$

We have

$$\overline{F}(t_1,t_2) = \frac{(1-t_1)(1-t_2)}{(1-t_1t_2)} \tag{17}$$

$$\overline{G}(t_1,t_2) = \frac{(t_1+t_2)(1-t_1)(1-t_2)}{t_1+t_2-t_1t_2} \tag{18}$$

and

$$-\overline{F}(t_1,ds_2) = \left(\frac{1-t_1}{1-t_1s_2}\right)^2 ds_2 \tag{19}$$

This gives

$$\begin{aligned}\int_{(0,t_2]} \frac{\Lambda_{01}(t_1,ds_2)}{\mathrm{E}Y(t_1,s_2)} &= \int_{(0,t_2]} \frac{t_1+s_2-t_1s_2}{(t_1+s_2)(1-t_1)(1-s_2)^3} ds_2 \\ &= \frac{1}{2\,(1-t_1)(1+t_1)}\left(\frac{1}{(1-t_2)^2}-1\right) \\ &\quad + \frac{t_1^2}{(1-t_1)(1+t_1)^2}\left(\frac{1}{1-t_2}-1\right) \\ &\quad + \frac{t_1^2}{(1-t_1)(1+t_1)^3}\left(-\log(1-t_2)+\log\frac{t_1+t_2}{t_1}\right)\end{aligned} \tag{20}$$

Hence the variance of the influence function of the pathwise estimator is in this case

$$VAR_{pathwise} =$$

$$\frac{1}{(1-t_1t_2)^2}\left\{\frac{1}{2}\,t_1(2-t_1)(1-t_2)^2\right.$$

$$\left. + \frac{1-t_1}{1+t_1}\left(\frac{1}{2}\,t_2(2-t_2) + \frac{t_1^2t_2(1-t_2)}{1+t_1} + \frac{t_1^2(1-t_2)^2}{(1+t_1)^2}\log\frac{t_1+t_2}{t_1(1-t_2)}\right)\right\} \tag{21}$$

B.6. Pareto $a = 1$, $b = \infty$

We have

$$\overline{F}(t_1,t_2) = \frac{(1-t_1)(1-t_2)}{(1-t_1t_2)} \tag{22}$$

and $\overline{G}$ is *uniform (0,1)*. Then

$$\int\limits_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} = \frac{1}{2\,(1-t_1)} \left(\frac{1}{(1-t_2)^2} - 1 \right) \tag{23}$$

and

$$VAR_{pathwise} = \frac{1}{2\,(1-t_1t_2)^2} \left\{ t_1(2-t_1)(1-t_2)^2 + (1-t_1)t_2(2-t_2) \right\} \tag{24}$$

B.7. Pareto $a = \infty$, $b = 0$

$\overline{F}$ is *uniform (0,1)* and $\overline{G}(t_1, t_2) = \min\{1-t_1, 1-t_2\}$. Then

$$\int\limits_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} = \begin{cases} \displaystyle\int\limits_{(0,t_2]} \frac{(1-t_1)ds_2}{(1-t_1)^2(1-s_2)^2(1-t_1)} & \text{for } t_1 > t_2 \\ \displaystyle\int\limits_{(0,t_1]} \frac{(1-t_1)ds_2}{(1-t_1)^2(1-s_2)^2(1-t_1)} + \int\limits_{(t_1,t_2]} \frac{(1-t_1)ds_2}{(1-t_1)^2(1-s_2)^3} & \text{for } t_1 \le t_2 \end{cases} \tag{25}$$

this gives

$$\begin{aligned} VAR_{pathwise} &= \frac{1}{2}\, t_1(2-t_1)(1-t_2)^2 + 1_{\{t_1 > t_2\}} t_2(1-t_2) \\ &\quad + 1_{\{t_1 \le t_2\}} \left\{ \frac{t_1(1-t_2)^2}{1-t_1} + \frac{1}{2\,(1-t_1)} \left((1-t_1)^2 - (1-t_2)^2 \right) \right\} \end{aligned} \tag{26}$$

B.8. Pareto $a = \infty$, $b = 1$

$\overline{F}$ is *uniform (0,1)* distributed and

$$\overline{G}(t_1, t_2) = \frac{(t_1+t_2)(1-t_1)(1-t_2)}{t_1 + t_2 - t_1t_2} \tag{27}$$

Then

$$\begin{aligned} \int\limits_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} &= \int\limits_{(0,t_2]} \frac{t_1 + s_2 - t_1s_2}{(t_1+s_2)(1-t_1)^2(1-s_2)^3} ds_2 \\ &= \frac{1}{2\,(1-t_1)^2(1+t_1)} \left(\frac{1}{(1-t_2)^2} - 1 \right) \\ &\quad + \frac{t_1^2}{(1-t_1)^2(1+t_1)^2} \left(\frac{1}{1-t_2} - 1 \right) \\ &\quad + \frac{t_1^2}{(1-t_1)^2(1+t_1)^3} \left(-\log(1-t_2) + \log \frac{t_1+t_2}{t_1} \right) \end{aligned} \tag{28}$$

and

$$\begin{aligned} VAR_{pathwise} &= \frac{1}{2}(1-t_2)^2 t_1(2-t_1) \\ &\quad + \frac{1}{1+t_1} \left(\frac{1}{2} t_2(2-t_2) + \frac{t_1^2 t_2(1-t_2)}{1+t_1} + \frac{t_1^2(1-t_2)^2}{(1+t_1)^2} \log \frac{t_1+t_2}{t_1(1-t_2)} \right) \end{aligned} \tag{29}$$

B.9. Pareto $a = \infty$, $b = \infty$

Both $\overline{F}$ and $\overline{G}$ are *uniform (0,1)*. We have

$$\int_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} = \int_{(0,t_2]} \frac{(1-t_1)ds_2}{(1-t_1)^3(1-s_2)^3} = \frac{1}{2\,(1-t_1)^2}\left(\frac{1}{(1-t_2)^2} - 1\right) \tag{30}$$

and this gives

$$\begin{aligned} VAR_{pathwise} &= (1-t_1)^2(1-t_2)^2 \left\{\frac{1}{2}\left(\frac{1}{(1-t_1)^2} - 1\right) + \frac{1}{2\,(1-t_1)^2}\left(\frac{1}{(1-t_2)^2} - 1\right)\right\} \\ &= \frac{1}{2}\left(1 - (1-t_1)^2(1-t_2)^2\right) \end{aligned} \tag{31}$$

B.10. Only one variable censored

$\overline{F}$ is *uniform (0,1)* and $\overline{G}(t_1, t_2) = 1 - t_2$. In this case we have no symmetry, so we have to treat the two observed paths separately. When first the first margin is estimated, we have

$$\int_{(0,t_2]} \frac{\Lambda_{01}(t_1, ds_2)}{\mathrm{E}Y(t_1, s_2)} = \int_{(0,t_2]} \frac{(1-t_1)\,ds_2}{(1-t_1)^2(1-s_2)^3} \tag{32}$$

and so

$$VAR_{pathwise} = (1-t_2)^2(1-t_1)t_1 + \frac{1}{2}(1-t_1)(2-t_2)t_2 \tag{33}$$

When we start with estimating the second margin, we have

$$\int_{(0,t_1]} \frac{\Lambda_{10}(ds_1, t_2)}{\mathrm{E}Y(s_1, t_2)} = \int_{(0,t_1]} \frac{(1-t_2)\,ds_1}{(1-s_1)^2(1-t_2)^3} \tag{34}$$

and so

$$VAR_{pathwise} = \frac{1}{2}(1-t_1)^2(2-t_2)t_2 + (1-t_1)t_1 \tag{35}$$

B.11. Each variable totally censored or not at all

$\overline{F}$ is *uniform (0,1)* and

$$\overline{G}(t_1, t_2) = \begin{cases} \frac{1}{3} & \text{for } t_1 > 0 \text{ and } t_2 > 0 \\ \frac{2}{3} & \text{for } t_1 = 0 \text{ and } t_2 > 0 \text{ or } t_1 > 0 \text{ and } t_2 = 0 \\ 1 & \text{for } t_1 = t_2 = 0 \end{cases} \tag{36}$$

We have

$$\int_{(0,t_1]} \frac{\Lambda_{10}(ds_1,0)}{\mathrm{E}Y(s_1,0)} = \frac{3}{2}\int_{(0,t_1]} \frac{ds_1}{(1-s_1)} = \frac{3}{2}\left(\frac{1}{1-t_1}-1\right) \tag{37}$$

and

$$\int_{(0,t_2]} \frac{\Lambda_{01}(t_1,ds_2)}{\mathrm{E}Y(s_1,t_2)} = \frac{3}{1-t_1}\int_{(0,t_2]} \frac{ds_2}{(1-s_2)^2} = \frac{3}{1-t_1}\left(\frac{1}{1-t_2}-1\right) \text{ for } t_1 \neq 0 \tag{38}$$

This gives

$$VAR_{pathwise} = \begin{cases} \frac{3}{2}(1-t_1)t_1(1-t_2)^2 + 3(1-t_1)(1-t_2)t_2 & \text{for } t_1 \neq 0 \\ \frac{3}{2}(1-t_2)t_2 & \text{for } t_1 = 0 \end{cases} \tag{39}$$