

ASYMPTOTIC THEORY FOR NONPARAMETRIC ESTIMATION OF SURVIVAL CURVES UNDER ORDER RESTRICTIONS

BY JENS THOMAS PRÆSTGAARD AND JIAN HUANG

College of American Pathologists and University of Iowa

We consider two problems in nonparametric survival analysis under the restriction of stochastic ordering. The first problem is that of estimating a survival function $\bar{F}(t)$ under the restriction $\bar{F}(t) \geq \bar{F}_0(t)$, all t , where $\bar{F}_0(t)$ is known. The second problem consists of estimating two unknown survival functions $\bar{F}^{(1)}(t)$ and $\bar{F}^{(2)}(t)$ when it is known that $\bar{F}^{(1)}(t) \geq \bar{F}^{(2)}(t)$, all t . The nonparametric maximum likelihood estimators in these problems were derived by Brunk, Franck, Hansen and Hogg and Dykstra. In the present paper we derive their large-sample distributions. We present two sets of proofs depending on whether or not the data are right-censored. When centered and scaled by $n^{1/2}$, the estimators converge in distribution to limiting processes related to the concave majorant of Brownian motion. The limiting distributions are not known in closed form, but can be simulated for the purpose of forming asymptotic pointwise confidence limits.

1. Introduction. In the present article we consider two estimation problems which arise in survival analysis when order restrictions are present. For the first problem suppose that the unobserved survival times X_1, \dots, X_n are iid with survival function $\bar{F} = 1 - F$, and that, independent hereof, C_1, \dots, C_n are iid censoring times with distribution G . What we observe is, in the usual right-censoring setup, the smallest of survival and censoring time and whether the observation was due to death or censoring; that is, the observations are iid pairs $(Z_i, \Delta_i) \equiv (X_i \wedge C_i, I(X_i \leq C_i))$. Let \bar{F}_0 be some fixed survival function, for example, the population survival curve available from demographical studies. The first estimation problem we consider is estimating \bar{F} subject to the restriction

$$(1.1) \quad \bar{F}(t) \geq \bar{F}_0(t) \quad \text{for all } 0 < t < 1.$$

$\bar{F}_0(t)$, may, for instance, be a survival function for a known reference population.

For the second problem we assume that we have two different sets of right-censored survival data from two different populations; that is, with obvious notation,

$$(Z_i^{(j)}, \Delta_i^{(j)}) \equiv (X_i^{(j)} \wedge C_i^{(j)}, I(X_i^{(j)} \leq C_i^{(j)})), \quad i = 1, \dots, n_j, j = 1, 2,$$

Received March 1994; revised July 1995.

AMS 1991 subject classifications. Primary 60E20, 62J02; secondary 60E12.

Key words and phrases. Order restrictions, nonparametric survival analysis, concave majorant.

where $(X_1^{(j)}, \dots, X_{n_j}^{(j)})$ and $(C_1^{(j)}, \dots, C_{n_j}^{(j)})$, $j = 1, 2$, are independent and $X_i^{(j)} \sim_{\text{iid}} F^{(j)}$, $C_i^{(j)} \sim_{\text{iid}} G^{(j)}$. Denoting the underlying survival curves by $\bar{F}^{(1)}$ and $\bar{F}^{(2)}$, respectively, problem 2 consists of estimating $\bar{F}^{(1)}$ and $\bar{F}^{(2)}$ subject to the order restriction

$$(1.2) \quad \bar{F}^{(1)}(t) \geq \bar{F}^{(2)}(t) \quad \text{for all } 0 < t < 1.$$

[This may, for instance, be relevant in studies such as the one reported by Dykstra (1982) where $\bar{F}^{(1)}$ and $\bar{F}^{(2)}$ denote male and female survival functions, and it is necessary to incorporate knowledge that females live longer than males.]

In the case of no censoring, the NLMLEs in the second problem were found by Brunk, Franck, Hanson and Hogg (1966); see also Barlow and Brunk (1972). Dykstra (1982) considered both problems under right censoring and found the nonparametric maximum likelihood estimators of \bar{F} , $\bar{F}^{(1)}$ and $\bar{F}^{(2)}$. As described later, these can be viewed as Kaplan–Meier estimators with a risk set so adjusted that the estimators meet the inequalities imposed in (1.1) and (1.2).

In this paper we derive the asymptotic distribution of the estimators in cases (1.1) and (1.2). We use two different sets of methods, depending on whether or not censoring is present. In the case of no censoring, it is possible to work with the original construction of Brunk, Franck, Hanson, and Hogg (1966) and minimal assumptions to get a limit theorem. Censoring complicates the problem somewhat and forces us to impose (slightly) more restrictive assumptions, such as the standard one of estimating $\bar{F}(t)$ only on $[0, 1]$, where $\bar{F}(1) > 0$. The tools used in the censored case are chiefly the product integral and empirical processes indexed by monotone functions.

The asymptotic distributions in cases (1.1) and (1.2) depend on whether or not the constraints are *active* or *inactive*. The constraints are said to be active for a t such that equality holds and inactive if the inequality is strict. It is folklore in order-restricted inference that the order-restricted estimators are $n^{1/2}$ equivalent to ordinary maximum likelihood estimators if the constraint is inactive, whereas an active constraint produces a usually nonnormal limiting law. This is true in our problem, too; if there is strict inequality for all t in (1.1) or (1.2), then the limiting laws are those of the respective Kaplan–Meier estimators. We mention this fact without proof; it is not surprising and easily shown. A more interesting situation occurs when the constraints are active for every t . In this case the limiting distributions are no longer Gaussian processes, but related to the concave majorant of Brownian motion. We focus our effort on this special case, since an everywhere active constraint illustrates our techniques and is the most important underlying model. The general situation in which the constraint is part active, part inactive on $[0, 1]$ is treated in the technical report preceding this paper; see Præstgaard and Huang (1995). The limiting law is, roughly, “in between” the laws from the everywhere active and everywhere inactive cases; in particular, the limiting distribution for any fixed t is nonnormal if only the constraint is active somewhere in $[0, 1]$.

2. Main results. We now state our first main result: the asymptotic distribution of the estimator from problem (1.1) when $\bar{F}(t) = \bar{F}_0(t)$ for all $0 \leq t \leq 1$. Assume that the distribution function $F(t)$ is differentiable on $]0, 1[$ with bounded and continuous intensity function, $\lambda(t) \equiv \dot{F}(t)/\bar{F}(t)$. Let $G(t)$ denote the censoring distribution and assume $\bar{G}(1) > 0$. Assume that $\lambda(t)$ is strictly positive on $[0, 1]$. Let $p(t) = \bar{F}(t)\bar{G}(t)$ denote the probability of remaining under study at time t and define the "time scale" $\phi(t) \equiv \int_0^t \lambda(u)/p(u) du$, $0 < t < 1$. Note that $p(1) > 0$ and $\phi(1) \equiv \tau < \infty$ since $\bar{F}(1) > 0$, $\bar{G}(1) > 0$. This is a standard assumption in survival analysis: that the life and censoring variables have positive probability of falling outside of the interval $[0, 1]$ at which we do the estimation; since the choice of finite interval $[0, 1]$ is arbitrary, this is not a real restriction.

With these definitions the following result holds.

THEOREM 2.1. *Let \bar{F}_n denote Dykstra's nonparametric estimator of the survival function under the restriction (1.1). Let W be a standard Brownian motion on $[0, \infty[$ and let $K(t)$ be the smallest nondecreasing concave function which majorizes W on $[0, \tau] \equiv [0, \phi(1)]$. Then*

$$\sqrt{n}(\bar{F}_n(t) - \bar{F}(t)) \Rightarrow -\bar{F}(t)(W(\phi(t)) - K(\phi(t)))$$

weakly in $l^\infty[0, 1]$.

The process $K(t)$ is formally defined as

$$K(t) \equiv \inf\{K^*(t): K^* \text{ concave and nondecreasing,} \\ K^*(u) \geq W(u), \text{ all } u \in [0, \tau]\}.$$

This process has been studied extensively in Groeneboom (1983), following Groeneboom and Pyke (1983). From these references it follows that the concave majorant of Brownian motion on all of \mathbf{R}_+ is a.s. increasing. We, however, consider a concave majorant only on $[0, \tau]$, and it might well decrease on the last part of the interval; hence the requirement that it be nondecreasing.

In the special case of no censoring, we have a theorem which avoids the technicality of assuming positive probability outside of $[0, 1]$ and, in turn, uses the concave majorant on the whole real line. To state this result, we recall that by the Doob transformation $B(t) \equiv (1-t)W(t/(1-t))$ is a Brownian bridge on $[0, 1]$ when $W(t)$ is Brownian motion on \mathbf{R}_+ . If $K(t)$ denotes the concave majorant of $W(t)$ on all of \mathbf{R}_+ and $M(t)$ denotes the concave majorant of $B(t)$ on $[0, 1]$, then Groeneboom (1983) showed that

$$(2.3) \quad M(t) = (1-t)K\left(\frac{t}{1-t}\right) \quad \text{for all } 0 < t < 1.$$

In the case of no censoring, it holds that $\phi(t) = F(t)/\bar{F}(t)$, and hence, by (2.3),

$$(2.4) \quad \bar{F}(t)(W(\phi(t)) - K(\phi(t))) = M(F(t)) - B(F(t)).$$

Theorem 2.1 can now be strengthened.

THEOREM 2.2. *Suppose that F is continuous and strictly increasing. Then there exists a Brownian bridge $B(t)$ so that*

$$\sup_{-\infty \leq x \leq \infty} \left| \sqrt{n} (\bar{F}_n(x) - \bar{F}(x)) + (M(F(x)) - B(F(x))) \right| \rightarrow_p 0 \quad \text{as } n \rightarrow \infty.$$

For estimation under the restriction (1.2), we get similar results. Again, we consider a constraint which is everywhere active, that is, $\bar{F}^{(1)}(t) = \bar{F}^{(2)}(t)$ for all t . Let the common force of mortality and not necessarily common censoring intensities be, respectively, $\lambda(t)$ and $\mu_i(t)$. Make the same assumptions about these as in Theorem 2.1. Let $n \equiv n_1 + n_2$ be the combined sample size and assume that $n_i/n \rightarrow c_i \in]0, 1[$. Define $\phi_i(t) \equiv \int_0^t \lambda_i(t)/(c_i p_i(t)) dt$ and $\phi(t) \equiv \phi_1(t) + \phi_2(t)$. Let $\dot{\phi}_i(t)$ denote the derivative of $\phi_i(t)$ and so on. We then have the following result.

THEOREM 2.3. *Let $\bar{F}_n^{(i)}$, $i = 1, 2$, denote Dykstra's nonparametric maximum likelihood estimator of $\bar{F}^{(i)}$ under the restriction (1.2). Let W_1 and W_2 be independent Brownian motions and define*

$$(2.5) \quad V(t) \equiv W_1(\phi_1(\phi^{-1}(t))) - W_2(\phi_2(\phi^{-1}(t))).$$

Let $K(t)$ denote the smallest nondecreasing concave majorant of $V(t)$ on $[0, \tau] \equiv [0, \phi(1)]$. Then

$$(2.6) \quad \begin{aligned} & \sqrt{n} \begin{pmatrix} \bar{F}_n^{(1)}(t) - \bar{F}(t) \\ \bar{F}_n^{(2)}(t) - \bar{F}(t) \end{pmatrix} \\ & \Rightarrow -\bar{F}(t) \begin{pmatrix} W_1(\phi_1(t)) - \int_0^t \dot{K}(\phi(u)) \dot{\phi}_1(u) du \\ W_2(\phi_2(t)) + \int_0^t \dot{K}(\phi(u)) \dot{\phi}_2(u) du \end{pmatrix}. \end{aligned}$$

Here we take $\dot{K}(u)$ to denote the right derivative at u , although this is not important since K is differentiable except in countably many points. The convergence in distribution in Theorems 2.1 and 2.3 is according to Hoffmann-Jørgensen (1984); see Dudley (1985) or Van der Vaart and Wellner (1995) for a description.

The limiting distribution in Theorem 2.3 becomes more tractable if it can be assumed that the censoring distributions agree for all t , that is, $\mu_1(t) = \mu_2(t)$, all t . In this case it holds that $p_1(u) = p_2(u) \equiv p(u)$ and $\phi_i(t) = \int_0^t \lambda(u)/(c_i p(u)) du = \phi(t)/c_i$. It follows that $\phi(t) = \phi(t)/(c_1 c_2)$, and $\phi_i \circ \phi(t) = (1 - c_i)t$. The expression (2.5) hence reduces to

$$(2.7) \quad V(u) = W_1(c_2 u) - W_2(c_1 u).$$

Note that

$$\begin{aligned} \int_0^t \dot{K}(\phi(u)) \dot{\phi}_i(u) du &= (1 - c_i) \int_0^t \dot{K}(\phi(u)/(c_1 c_2)) \dot{\phi}(u)/(c_1 c_2) du \\ &= (1 - c_i) K(\phi(t)/(c_1 c_2)), \end{aligned}$$

which is by linearity equal to $(1 - c_i)L(\phi(t))$ with $L(t)$ the least nondecreasing concave majorant on $[0, \phi(1)]$ of $V(t/(c_1c_2)) = W_1(t/c_1) - W_2(t/c_2)$. By Brownian scaling, the latter difference equals

$$V(t/(c_1c_2)) = c_1^{-1/2}W_1^*(t) - c_2^{-1/2}W_2^*(t)$$

for $W_i^*(t) \equiv c_i^{1/2}W_i(t/c_i)$ iid Brownian motions. The limit (2.6) now reduces to

$$(2.8) \quad -\bar{F}(t) \left(\begin{matrix} c_1^{-1/2}W_1^*(\phi(t)) - c_2L(\phi(t)) \\ c_2^{-1/2}W_2^*(\phi(t)) + c_1L(\phi(t)) \end{matrix} \right).$$

If there is no censoring present, $\phi(t) = F(t)/\bar{F}(t)$, and (2.8) becomes

$$(2.9) \quad -\bar{F}(t) \left(\begin{matrix} c_1^{-1/2}W_1^*(F(t)/\bar{F}(t)) - c_2L(F(t)/\bar{F}(t)) \\ c_2^{-1/2}W_2^*(F(t)/\bar{F}(t)) + c_1L(F(t)/\bar{F}(t)) \end{matrix} \right).$$

Define now, for $0 < t/(c_1c_2(1 - t)) < \tau$, $M(t) \equiv (1 - t)L(t/(1 - t))$. Then $M(t)$ is concave, and by the Doob transformation it can be recognized as the least concave majorant (on $[0, \tau^*]$ where $\tau^*/(c_1c_2(1 - \tau^*)) = \tau$) of

$$(1 - t)V\left(\frac{t}{(1 - t)c_1c_2}\right).$$

This quantity equals

$$(2.10) \quad \begin{aligned} &(1 - t)V\left(\frac{t}{c_1c_2(1 - t)}\right) \\ &= \left(c_1^{-1/2}W_1^*\left(\frac{t}{1 - t}\right) + c_2^{-1/2}W_2^*\left(\frac{t}{1 - t}\right) \right) \\ &= (1 - t)(c_1^{-1/2}B_1(t) - c_2^{-1/2}B_2(t)) \end{aligned}$$

for independent Brownian bridges $B_i(t) = (1 - t)W_i^*(t/(1 - t))$. Equation (2.9) can now be rewritten further as

$$(2.11) \quad -\left(\begin{matrix} c_1^{-1/2}B_1(F(t)) - c_2M(F(t)) \\ c_2^{-1/2}B_2(F(t)) + c_1M(F(t)) \end{matrix} \right).$$

Analogously to Theorem 2.2, we can in the case of no censoring prove a version of Theorem 2.3 which strengthens the result (2.11). For this, we rely on the explicit description of the NPMLs of Brunk, Franck, Hanson and Hogg (1966). This version extends the weak convergence to the whole real line in that the assumption $\bar{F}(1) > 0$ is no longer needed. For the same reason we can work with a concave majorant of Brownian bridge on all of $[0, 1]$, that is, not restricted to $[0, \tau^*]$. The smoothness assumptions on the underlying distributions, which were needed in Theorem 2.3, can also be reduced.

THEOREM 2.4. *Suppose that F is continuous and strictly increasing. Then there exist two independent Brownian bridges so that*

$$(i) \quad \sup_{-\infty \leq x \leq \infty} \left| \sqrt{n} \left(\bar{F}_n^{(1)}(x) - \bar{F}(x) \right) + \left[c_1^{-1/2} B_1(F(x)) - c_2 M(F(x)) \right] \right| \rightarrow_p 0,$$

$$(ii) \quad \sup_{-\infty \leq x \leq \infty} \left| \sqrt{n} \left(\bar{F}_n^{(2)}(x) - \bar{F}(x) \right) + \left[c_2^{-1/2} B_2(F(x)) + c_1 M(F(x)) \right] \right| \rightarrow_p 0,$$

where $M(t)$ is the least concave majorant of $c_1^{-1/2} B_1(t) - c_2^{-1/2} B_2(t)$ on $[0, 1]$.

3. An application to pacemaker survival data. Dykstra (1982) describes a data set consisting of right-censored survival times of patients after pacemaker implant. He estimated the survival functions of males and females separately, under the restriction that the female survival function be at any time larger than the male; that is, $\bar{F}^{(f)}(t) \geq \bar{F}^{(m)}(t)$ for all $t > 0$. We compute the NPMLEs $\bar{F}_n^{(m)}(t)$ and $\bar{F}_n^{(f)}(t)$ under this restriction and graph them with two sets of pointwise 95% confidence limits. Figures 1 and 2 show the estimated survival functions for females and males, respectively. The dashed curves are the confidence limits under an assumption of strict in-

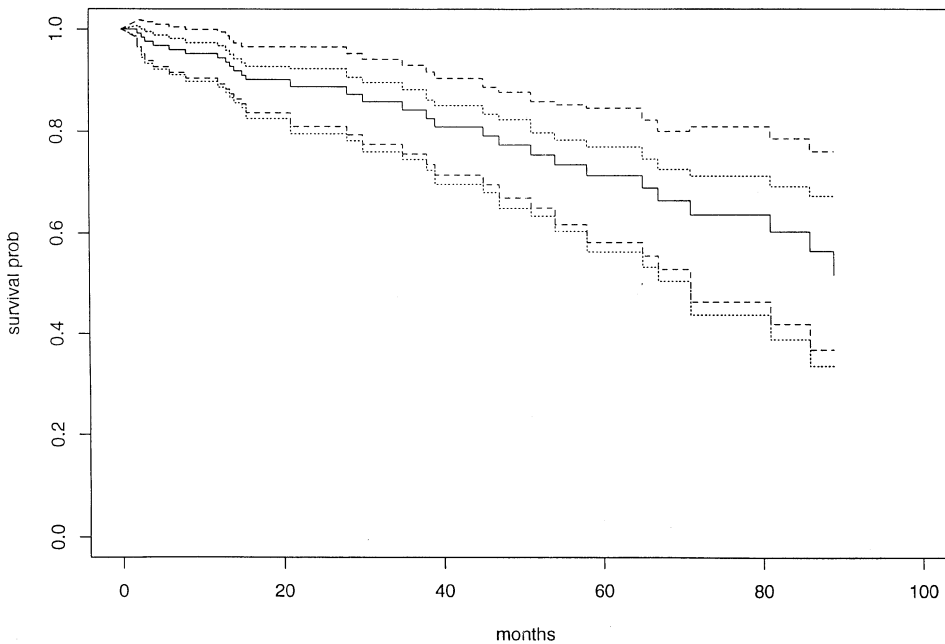


FIG. 1. *Female survival curves.*

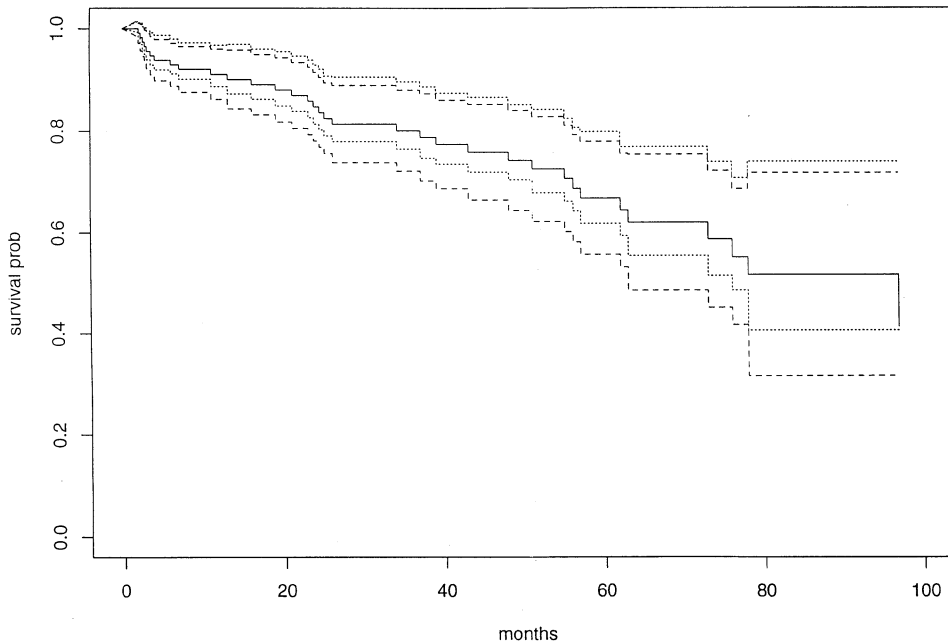


FIG. 2. *Male survival curves.*

equality, that is, an everywhere inactive constraint. The dotted curves are confidence bands calculated under an assumption of strict equality, that is, an everywhere active constraint. We assumed that the censoring distributions were the same for males and females. For the inactive constraint case, the asymptotic distribution is the same as that of the usual Kaplan-Meier estimators, so pointwise confidence bands can be based on a normal approximation, that is, Greenwood's formula; see Breslow and Crowley (1974). For the second case the confidence limits are based on evaluating the distribution in (2.8). Since this distribution is not known to us in a closed form, it must be simulated. 1000 realizations of (2.8) went into each of the dotted curves; on a Sun IPX SPARCstation using C and Splus, it took three minutes to generate these realizations. For both sets of bands we estimated the function $\phi(t) = \int_0^t \lambda(u)/p(u) du$ by $\int_0^t n dN(u)/Y(u)^2$, where $N(u)$ and $Y(u)$ denote the scaled counting process and risk set of the pooled data. This gives a better comparison, since the Greenwood bounds are then based on the Brownian motion parts of (2.8). The proportion c_1 was estimated by the ratio of the number of females to the number of patients, $85/205 = 0.415$.

The pointwise confidence bounds based on an active constraint are seen to be asymmetric. In the case of females, who have the supposedly larger survival probability, the lower pointwise bounds are a little wider than the Greenwood bounds, whereas the upper bounds are substantially narrower than the Greenwood bounds. The opposite situation is true for the males.

That the better bound is the upper bound for the females can be explained as follows. When the constraint is active, there is no difference between male and female survivor, so when using an estimator which has been increased relative to the Kaplan–Meier estimator, the upper bounds need to be decreased relative to those of the Kaplan–Meier estimator. In practice, it is unknown whether the constraint is active or inactive. In view of this, a suggestion for a conservative set of pointwise confidence bands for the order-restricted estimator is to use Greenwood’s formula just as if the estimator were Kaplan–Meier, but more simulations are needed to justify this point.

4. The uncensored case. In this section we prove the results for the uncensored case, that is, Theorems 2.2 and 2.4.

4.1. *Computation of the estimator.* For the one-sample problem, the NPMLE \hat{F}_n is described in Robertson, Wright and Dykstra (1988) as follows. Let F_n be the empirical distribution function of X_1, \dots, X_n . One first forms the *least concave majorant* of the cumulative sum diagram of the points

$$(F_n(X_{(i)}), F_0(X_{(i)})), \quad i = 1, \dots, n,$$

plus the origin $(0, 0)$. Let $\hat{H}(t)$, $0 \leq t \leq 1$, denote this least concave majorant. Then $\hat{F}_n(x) = \hat{H}(F_n(x))$. Based on this expression, the asymptotic properties of \hat{F}_n can be obtained analogously to those of the NPMLE in the two-sample problem. So we will focus on the two-sample problem.

In the two-sample case, the graphic representation of the NPMLE, due to Brunk, Franck, Hanson and Hogg (1966), goes as follows. First, order the combined observations from the smallest one to the largest one. Then take a random walk in the plane starting at the origin, moving one unit up for an $X^{(1)}$ observation and one unit to the right for an $X^{(2)}$ observation. Next, draw the least concave majorant of the points visited during the random walk. The following pattern of data are graphically depicted in Figure 3:

$$X_1^{(1)} < X_1^{(2)} < X_2^{(1)} < X_3^{(1)} < X_2^{(2)} < X_3^{(2)} < X_4^{(1)} < X_5^{(1)} < X_4^{(2)} < X_5^{(2)} < X_6^{(1)}.$$

In the figure an $X^{(1)}$ observation is indicated by X1, and an $X^{(2)}$ observation is indicated by X2.

Observations are said to be in the i th string if they occur after the $(i - 1)$ th vertex up to and including the i th vertex. The origin is taken to be the zeroth vertex. Let j_v denote the number of $X^{(2)}$'s in the v th string and k_v the number of $X^{(1)}$'s in the v th string. Let B be the total number of strings. In the example here, $j_1 = 0$, $k_1 = 1$, $j_2 = 1$, $k_2 = 2$, $j_3 = 2$, $k_3 = 2$, $j_4 = 2$, $k_4 = 1$ and $B = 4$. Theorem 1 of Brunk, Franck, Hanson and Hogg (1966) states that the NPMLEs of F_1 and F_2 , subject to $F_n^{(1)}(x) \leq F_n^{(2)}(x)$ for every x , are given by

$$(4.12) \quad \begin{aligned} F_n^{(1)}(x) &= \sum_{i=1}^{n_1} \hat{f}(X_i^{(1)}) 1_{[X_i^{(1)} \leq x]} + \alpha_1(x), \\ F_n^{(2)}(x) &= \sum_{i=1}^{n_2} \hat{g}(X_i^{(2)}) 1_{[X_i^{(2)} \leq x]} + \alpha_2(x), \end{aligned}$$

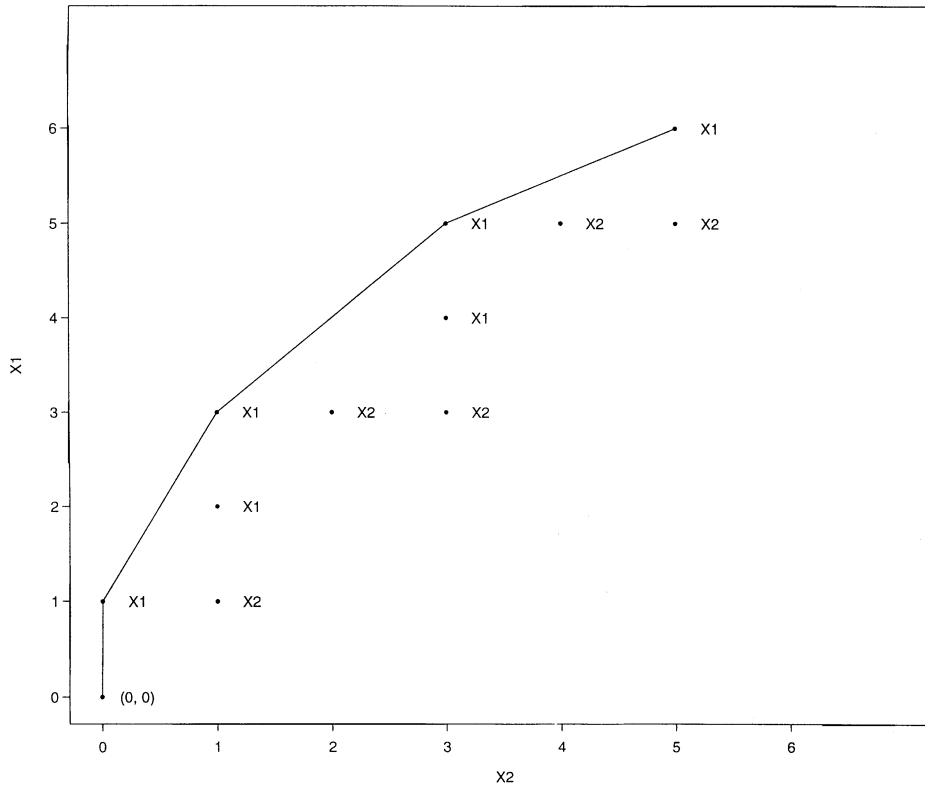


FIG. 3. The random walk and its least concave majorant.

where

$$\hat{f}(X_i^{(1)}) = \frac{j_v + k_v}{k_v n} \quad \text{when } X_i^{(1)} \text{ is in the } v \text{th string,}$$

$$\hat{g}(X_i^{(2)}) = \frac{j_v + k_v}{j_v n} \quad \text{when } X_i^{(2)} \text{ is in the } v \text{th string.}$$

Here $\alpha_1(x) = k_1/n$ if $j_1 = 0$ and $x > X_{(1)}^{(1)}$, otherwise $\alpha_1(x) = 0$; $\alpha_2(x) = j_B/n$ if $k_B = 0$ and $x \geq X_{(n_2)}^{(2)}$, otherwise $\alpha_2(x) = 0$. Here $X_{(1)}^{(1)}$ is the smallest observation in the first sample, and $X_{(n_2)}^{(2)}$ is the largest one in the second sample. Straightforward calculation shows that $\alpha_1(x) = O_p(n^{-1})$ and $\alpha_2(x) = O_p(n^{-1})$. Thus both $\alpha_1(x)$ and $\alpha_2(x)$ are asymptotically negligible and will not be considered in the following. It is useful to write $\hat{f}(x) = (1 + (j_v/k_v))/n$ and $\hat{g}(x) = (1 + (k_v/j_v))/n$. So \hat{f} and \hat{g} depend on j_v and k_v only through k_v/j_v , which can be seen as the slope of the least concave majorant in the v th string. That is, let $s(x)$, $0 < x \leq n_2$, be the left-continuous slope of this least concave majorant, then for $0 < x \leq n_2$, if x is in the v th string, $s(x) = k_v/j_v$ and $s(0) = s(1)$.

Denote empirical distribution functions by $F_{n_1}(x) \equiv n_1^{-1} \sum_{i=1}^{n_1} 1_{[X_i^{(1)} \leq x]}$ and $F_{n_2}(x) \equiv n_2^{-1} \sum_{i=1}^{n_2} 1_{[X_i^{(2)} \leq x]}$. Let $X_{(1)}^{(j)}, X_{(2)}^{(j)}, \dots, X_{(n_j)}^{(j)}$ be the order statistics of $X_1^{(j)}, \dots, X_{n_j}^{(j)}$, $j = 1, 2$. For $0 \leq t \leq 1$, define the left-continuous inverse function $F_{n_1}^{-1}$ of F_{n_1} and the right-continuous inverse function $F_{n_2}^{-1}$ of F_{n_2} by

$$F_{n_1}^{-1}(t) = \inf\{x: F_{n_1}(x) \geq t\} \quad \text{and} \quad F_{n_2}^{-1}(t) = \sup\{y: F_{n_2}(y) \leq t\},$$

respectively, where $F_{n_1}^{-1}(0) = -\infty$ and $F_{n_2}^{-1}(1) = \infty$. Then it can be seen that the least concave majorant of the points of the random walk is the same as the least concave majorant of the points

$$(n_2 F_{n_2}(X_{(i)}^{(2)}), n_1 F_{n_1}(X_{(i+1)}^{(2)})), \quad i = 0, 1, \dots, n_2,$$

plus the origin $(0, 0)$. Here we define $X_{(n_2+1)}^{(2)} = \infty$. Since $X_{(i+1)}^{(2)} = F_{n_2}^{-1}(i/n_2)$, $i = 0, 1, \dots, n_2$, the slope $s(x)$ of this least concave majorant at $x = i$ is the same as n_1/n_2 times $u_n(t)$ evaluated at $t = F_{n_2}(X_i^{(2)})$, $i = 1, \dots, n_2$, where, for $1/n_2 \leq t \leq 1$, $u_n(t)$ is the left-continuous slope of the *least concave majorant* [denoted by $\hat{U}_n(t)$], of $F_{n_1}(F_{n_2}^{-1}(t))$, $t \in [0, 1]$, plus $(0, 0)$, and $u_n(t) = u_n(F_{n_2}(X_{(1)}^{(2)}))$ for $0 \leq t < 1/n_2$. Let $U_n(t) = F_{n_1} F_{n_2}^{-1}(t)$.

By (4.12), with probability 1 and without considering $\alpha_2(x)$, we can write

$$\begin{aligned} F_n^{(2)}(x) &= \frac{1}{n} \sum_{i=1}^{n_2} 1_{[X_i^{(2)} \geq 1]} + \frac{1}{n} \sum_{i=1}^{n_2} \frac{n_1}{n_2} u_n(F_{n_2}(X_i^{(2)})) 1_{[X_i^{(2)} \leq x]} \\ (4.13) \quad &= \frac{n_2}{n} F_{n_2}(x) + \frac{n_1}{n} \int_{-\infty}^x u_n(F_{n_2}(x)) dF_{n_2}(x) \\ &= \frac{n_2}{n} F_{n_2}(x) + \frac{n_1}{n} \int_0^{F_{n_2}(x)} u_n(t) dt \\ &= \frac{n_2}{n} F_{n_2}(x) + \frac{n_1}{n} \hat{U}_n(F_{n_2}(x)) + \frac{n_1}{n} \hat{U}_n(0). \end{aligned}$$

Notice that $\hat{U}_n(F_{n_2}(x))$ is $\hat{U}_n(t)$ evaluated at $t = F_{n_2}(x)$. Since $\hat{U}_n(0) = F_{n_1} F_{n_2}^{-1}(0) = F_{n_1}(X_{(1)}^{(2)}) = O_p(n^{-1})$ by a straightforward calculation, so the term $(n_1/n)\hat{U}_n(0)$ is asymptotically negligible. Similarly, for $0 \leq t \leq 1$, let $\hat{V}_n(t)$ be the *greatest convex minorant* (GCM) of $V_n(t) \equiv F_{n_2} F_{n_1}^{-1}(t)$ and let $v_n(t)$ be the left-continuous slope of $\hat{V}_n(t)$. Now we always have $\hat{V}_n(0) = 0$. Similarly to (4.13), we have

$$(4.14) \quad F_n^{(1)}(x) = \frac{n_1}{n} F_{n_1}(x) + \frac{n_2}{n} \hat{V}_n(F_{n_1}(x)) - \frac{n_2}{n} \hat{V}_n(0).$$

With these expressions, to obtain the large-sample properties of $F_n^{(2)}$ and $F_n^{(1)}$, we only need to consider the processes $\hat{U}_n(F_{n_2}(x))$ and $\hat{V}_n(F_{n_1}(x))$, because the asymptotic properties of the empirical distribution functions are well known. The treatment of $\hat{V}_n(F_{n_1}(x))$ is very similar to $\hat{U}_n(F_{n_2}(x))$, so we will focus on the latter one.

The heuristics behind the formal treatment of $\hat{U}_n(F_{n_2}(x))$ are quite simple. First, consider $\hat{U}_n(t)$ for $0 \leq t \leq 1$. Since $U_n(t) = F_{n_1} F_{n_2}^{-1}(t)$ converges to t , $\hat{U}_n(t)$ should converge to the least concave majorant of t , which is t itself. So

$\hat{U}_n(F_{n_2}(x))$ should be close to $F_{n_2}(x)$ when n_2 is large. By (4.13), it is seen that $F_n^{(2)}$ is consistent. For a different proof of consistency, see Brunk, Franck, Hanson and Hogg (1966). It is also seen that the right centering for $\hat{U}_n(t)$ should be t . And we naturally expect that the normalized processes $\sqrt{n}(\hat{U}_n(t) - t)$ converges weakly to some limiting process associated with Brownian bridges, since it can be shown that $\sqrt{n}(F_{n_1}F_{n_2}^{-1}(t) - t)$ converges weakly to a weighted sum of two independent Brownian bridges. By concavity, for every x , $\sqrt{n}[\hat{U}_n(F_{n_2}(x)) - \hat{U}_n(F(x))]$ is sandwiched between $\sqrt{n}u_n(F_{n_2}(x))(F_{n_2}(x) - F(x))$ and $\sqrt{n}u_n(F(x))(F_{n_2}(x) - F(x))$. But the slope $u_n(t)$ converges to 1, the slope of t , for all $0 < t \leq 1$. Hence $\sqrt{n}\hat{U}_n(F_{n_2}(x))$ should be close to $\sqrt{n}\hat{U}_n(F(x)) + \sqrt{n}(F_{n_2}(x) - F(x))$. This is proved in Corollary 4.1. So the task of finding the limiting distribution of $\hat{U}_n(F_{n_2}(x))$ reduces to finding the limiting distribution of $\hat{U}_n(F(x))$. This is done in Theorem 4.1.

4.2. *No censoring: auxiliary results and proofs.* Recall that F is assumed to be continuous and strictly increasing. Let F^{-1} be the inverse of F . For $0 \leq t \leq 1$, define

$$(4.15) \quad B_{n_1}(t) = \sqrt{n_1}(F_{n_1}F^{-1}(t) - t) \quad \text{and} \quad B_{n_2}(t) = \sqrt{n_2}(F_{n_2}F^{-1}(t) - t).$$

Let $D[0, 1]$ be the space of real-valued functions on $[0, 1]$ that are right continuous and have left limits. Let ρ be the supremum norm on $D[0, 1]$; that is, for any f and g in $D[0, 1]$,

$$\rho(f, g) = \sup_{0 \leq x \leq 1} |f(x) - g(x)|.$$

Then there exist special constructions of independent Brownian bridges B_1 and B_2 such that

$$(4.16) \quad \rho(B_{n_1}, B_1) \rightarrow_{\text{a.s.}} 0 \quad \text{and} \quad \rho(B_{n_2}, B_2) \rightarrow_{\text{a.s.}} 0.$$

See Pyke and Shorack (1968), (2.2), page 757. For $t \in [0, 1]$, let

$$(4.17) \quad B(t) = c_1^{-1/2}B_1(t) - c_2^{-1/2}B_2(t).$$

Recall $M(t)$ is the *least concave majorant* of $B(t)$ on $[0, 1]$.

LEMMA 4.1. *Suppose that F is continuous and strictly increasing. Then, as $n \rightarrow \infty$,*

$$\rho(\sqrt{n}(F_{n_1}F_{n_2}^{-1} - I), B) \rightarrow_P 0,$$

where I is the identity map, that is, $I(t) = t$.

PROOF. The proof uses a representation similar to that in Lemma 3.1 of Pyke and Shorack (1968), and it then follows from Lemma 2.4 of that paper. \square

Let \mathcal{P} be the class of all finite discrete probability measures on $[0, 1]$. The least concave majorant $\hat{g}(t)$ of a function $g(t)$ on $[0, 1]$ can be written as

$$\hat{g}(t) = \sup \left\{ \int_{[0,1]} g(u) dP : \int_{[0,1]} u dP = t, P \in \mathcal{P} \right\}.$$

See, for example, Rockafellar (1970), page 36, or Groeneboom and Pyke (1983), page 339. It follows that

$$(4.18) \quad \begin{aligned} & \sqrt{n} (\hat{U}_n(t) - t) \\ &= \sup \left\{ \sqrt{n} \int_{[0,1]} (U_n(u) - u) dP : \int_{[0,1]} u dP = t, P \in \mathcal{P} \right\}. \end{aligned}$$

Combining Lemma 4.1 and (4.18), we have the following theorem.

THEOREM 4.1. *Suppose that F is continuous and strictly increasing. Then, as $n \rightarrow \infty$,*

$$(4.19) \quad \sup_{0 \leq t \leq 1} |\sqrt{n} (\hat{U}_n(t) - t) - M(t)| \rightarrow_p 0.$$

PROOF. By (4.18), it suffices to show that

$$(4.20) \quad \sup_{0 \leq t \leq 1} \left| \sup \left\{ \int_{[0,1]} \sqrt{n} (U_n(u) - u) dP : \int_{[0,1]} u dP = t, P \in \mathcal{P} \right\} - M(t) \right| \rightarrow_p 0$$

as $n \rightarrow \infty$. By (4.18), we can write

$$M(t) = \sup \left\{ \int_{[0,1]} B(u) dP : \int_{[0,1]} u dP = t, P \in \mathcal{P} \right\}.$$

By the triangle inequality for the supremum norm, the left-hand side of (4.20) is less than

$$\sup_{0 \leq t \leq 1} |\sqrt{n} (U_n(t) - t) - B(t)|,$$

which converges to 0 in probability by Lemma 4.1. \square

COROLLARY 4.1. *Let $D_n(x) = \sqrt{n} [\hat{U}_n(F_{n_2}(x)) - \hat{U}_n(F(x)) - (F_{n_2}(x) - F(x))]$. Then*

$$\sup_{-\infty \leq x \leq \infty} |D_n(x)| \rightarrow_p 0.$$

PROOF. By the triangle inequality,

$$\begin{aligned} & \sup_{-\infty \leq x \leq \infty} |D_n(x)| \\ & \leq \sup_{-\infty \leq x \leq \infty} \left| \sqrt{n} [\hat{U}_n(F_{n_2}(x)) - F_{n_2}(x)] - M(F_{n_2}(x)) \right| \\ & \quad + \sup_{-\infty \leq x \leq \infty} \left| \sqrt{n} [\hat{U}_n(F(x)) - F(x)] - M(F(x)) \right| \\ & \quad + \sup_{-\infty \leq x \leq \infty} \left| M(F_{n_2}(x)) - M(F(x)) \right| \\ & \leq 2 \sup_{0 \leq t \leq 1} \left| \sqrt{n} [\hat{U}_n(t) - t] - M(t) \right| + \sup_{-\infty \leq x \leq \infty} \left| M(F_{n_2}(x)) - M(F(x)) \right|. \end{aligned}$$

The first term in the last line converges to 0 in probability by Theorem 4.1. The second term converges to 0 in probability, since $M(t)$ is concave and hence has almost surely continuous sample path, and since $F_{n_2}(x)$ converges to $F(x)$ uniformly almost surely by the Glivenco–Cantelli theorem. \square

PROOF OF THEOREM 2.4. We prove part (ii). Part (i) can be proved in a completely analogous way. By Corollary 4.1,

$$U_n(F_2(x)) = \hat{U}_n(F(x)) + F_{n_2}(x) - F(x) + n^{-1/2}o_P(1),$$

where here and in the following $o_P(1)$ converges to 0 in probability uniformly in x . By (4.13), it follows that

$$\begin{aligned} & \sqrt{n} (F_n^{(2)}(x) - F(x)) \\ &= \sqrt{n} \frac{n_2}{n} (F_{n_2}(x) - F(x)) + \sqrt{n} \frac{n_1}{n} (\hat{U}_n(F_{n_2}(x)) - F(x)) + \frac{n_1}{n} \hat{U}_n(0) \\ &= \sqrt{n} \frac{n_2}{n} (F_{n_2}(x) - F(x)) + \sqrt{n} \frac{n_1}{n} (\hat{U}_n(F(x)) - F(x) + F_{n_2}(x) - F(x)) \\ & \quad + n^{-1/2}o_P(1) \\ &= \left(\frac{n}{n_2}\right)^{1/2} \sqrt{n_2} (F_{n_2}(x) - F(x)) + \frac{n_1}{n} \sqrt{n} (\hat{U}_n(F(x)) - F(x)) \\ & \quad + n^{-1/2}o_P(1). \end{aligned}$$

So the theorem follows from (4.16) and Theorem 4.1. \square

5. Censored case.

5.1. *Description of the estimators.* With censoring, the MPMLEs are due to Dykstra (1982). In this case there are no graphic representations or explicit expressions available for the estimators. They are best described by using counting processes and the product integral. Consider first Theorem 2.1. Let $N_n(t)$ denote the (scaled) counting process of deaths at time t and let $Y_n(t)$ denote the (scaled) risk set at time t , that is,

$$N_n(t) \equiv \frac{1}{n} \sum_{j=1}^n I(Z_j \leq t, \Delta_j = 1), \quad Y_n(t) \equiv \frac{1}{n} \sum_{j=1}^n I(Z_j \geq t).$$

If Π denotes the product integration operator, then the usual Kaplan–Meier estimator $S_n(t)$ based on (Z_i, Δ_i) , $i = 1, \dots, n$, may be written as

$$S_n(t) = \prod_{]0, t]} \left(1 - \frac{dN_n(u)}{Y_n(u)} \right).$$

For an explanation of the product integral and its merits in counting processes and survival analysis, see Gill and Johansen (1990), in particular, Section 4.

In this setup, the maximum likelihood estimator in the first problem (1.1), given in Dykstra (1982), Theorem 2.1, may be conveniently described as follows. For each $0 < s < t < 1$, define $\beta_n(s, t)$ as the solution for β to the equation

$$(5.21) \quad \pi_n(\beta, s, t) \equiv \prod_{]s, t]} \left(1 - \frac{dN_n(u)}{Y_n(u) + \beta} \right) = \frac{\bar{F}_0(t)}{\bar{F}_0(s)}$$

(and introduce at the same time a more compact notation). Then define further a stochastic process $\gamma_n(t)$ by

$$(5.22) \quad \gamma_n(t) \equiv \left(\min_{s \leq t} \max_{u \geq t} \beta_n(s, u) \right) \vee 0.$$

Then the nonparametric maximum likelihood estimator is given by

$$(5.23) \quad \bar{F}_n(t) \equiv \prod_{]s, t]} \left(1 - \frac{dN_n(u)}{Y_n(u) + \gamma_n(u)} \right).$$

The situation in Theorem 2.3 can be handled by a similar technique. Define $(N_n^{(j)}, Y_n^{(j)})$, $j = 1, 2$, as the counting process of deaths and the risk set for both samples:

$$N_n^{(j)}(t) \equiv \frac{1}{n} \sum_{i=1}^n I(Z_i^{(j)} \leq t, \Delta_i^{(j)} = 1), \quad Y_n^{(j)}(t) \equiv \frac{1}{n} \sum_{i=1}^n I(Z_i^{(j)} \geq t).$$

(Recall that $n = n_1 + n_2$.) Then the maximum likelihood estimators $\bar{F}_n^{(1)}$ and $\bar{F}_n^{(2)}$ given in Dykstra (1982), Theorem 2.2, emerge as follows. Define $\beta_n^*(s, t)$ as the solution to the equation

$$(5.24) \quad \begin{aligned} \pi_n^{(1)}(\beta, s, t) &\equiv \prod_{]s, t]} \left(1 - \frac{dN_n^{(1)}(u)}{Y_n^{(1)}(u) + \beta} \right) \\ &= \prod_{]s, t]} \left(1 - \frac{dN_n^{(2)}(u)}{Y_n^{(2)}(u) - \beta} \right) \equiv \pi_n^{(2)}(-\beta, s, t). \end{aligned}$$

Then, on setting $\gamma_n^*(t) \equiv (\min_{s \leq t} \max_{u \geq t} \beta_n^*(s, u)) \vee 0$, the NPMLEs are given by, respectively,

$$(5.25) \quad \bar{F}_n^{(1)}(t) \equiv \prod_{]0, t]} \left(1 - \frac{dN_n^{(1)}(u)}{Y_n^{(1)}(u) + \gamma_n^*(u)} \right)$$

and

$$(5.26) \quad \bar{F}_n^{(2)}(t) \equiv \prod_{]0, t]} \left(1 - \frac{dN_n^{(2)}(u)}{Y_n^{(2)}(u) - \gamma_n^*(u)} \right).$$

The proofs of Theorems 2.1 and 2.3 are quite similar. We now briefly explain the ideas in proving Theorem 2.1. The first step is to investigate the

asymptotic behavior of the adjustment to the risk set, $\gamma_n(t)$, which is defined in (5.22). The "population" version of $\pi_n(\beta, s, t)$ is

$$\begin{aligned}
 \pi_0(\beta, s, t) &\equiv \prod_{[s, t]} \left(1 - \frac{dEN_n(u)}{EY_n(u) + \beta} \right) \\
 &= \prod_{[s, t]} \left(1 - \frac{\lambda(u)p(u)}{p(u) + \beta} \right) \\
 (5.27) \quad &\left[\text{since } EN_n(u) = \int_0^u \lambda(u)p(u) du, EY_n(u) = p(u) \right] \\
 &= \exp \left(- \int_s^t \frac{\lambda(u)p(u)}{p(u) + \beta} du \right),
 \end{aligned}$$

where the last equality is a general property of the product integral of a continuous interval function, described in Gill and Johansen (1990). Since the constraint is everywhere active, the solution to $\pi_0(\beta, s, t) = \bar{F}_0(t)/\bar{F}_0(s)$ ($= \bar{F}(t)/\bar{F}(s)$) is $\beta = 0$. Hence we expect that $n^{1/2}\beta_n(s, t)$ converges in law to a centered Gaussian process indexed by s and t . This is the content of Theorem 5.1. Having shown this, we use a continuous mapping theorem (Lemma 5.4) for the min-max functional defined in (5.22) to show that $n^{1/2}\gamma_n(t)$ converges along an a.s. representation to the left derivative of the least nondecreasing concave majorant of a Brownian motion sample path. Finally, we use asymptotic linearity of the product integral in (5.21) to find the limit of $n^{1/2}(\bar{F}_n - \bar{F})$. The last part of the proof (Lemma 5.5) is similar to Gill and Johansen's (1990) δ -method proof for asymptotic normality of the Kaplan-Meier estimator, which corresponds to the situation in which no adjustment to the risk set is made, that is, $\gamma_n(t) \equiv 0$. It is known from Groeneboom (1983) that the derivative of the least concave majorant of Brownian motion on $[0, \infty[$ has right limit ∞ at 0 a.s. Hence some uniform integrability of $n^{1/2}\gamma_n$ is needed; this is established by empirical process methods in Lemma 5.3. To explain briefly why such tools are useful, let $\mathbf{P} \equiv F \times G$ and $\mathbf{P}_n \equiv n^{-1} \sum_{j=1}^n \delta_{(X_j, C_j)}$ denote the true and empirical probability measures for the (partly) unobserved data in problem (1.1). Let $\mathbf{X}_n \equiv n^{1/2}(\mathbf{P}_n - \mathbf{P})$ denote the empirical process. If we use the notation

$$\mathbf{P}(f) = Ef(X_1, C_1), \quad \mathbf{P}_n(f) = \sum_{j=1}^n f(X_j, C_j),$$

then we have that

$$\int_s^t \frac{dN_n(u)}{Y_n(u)} = \mathbf{P}_n(g_n(\cdot, \cdot; s, t))$$

and

$$(5.28) \quad \int_s^t \frac{dEN_n(u)}{Y_n(u)} = \mathbf{P}(g_n(\cdot, \cdot; s, t)),$$

where the function $(x, c) \rightarrow g_n(x, c; s, t)$ is given by

$$g_n(x, c; s, t) = \frac{1}{Y_n(x)} \mathbf{1}_{]s, t]}(x) \mathbf{1}_{\{x < c\}}, \quad 0 < s < t < 1, n = 1, 2, \dots$$

Due to the work of Dudley (1978, 1984, 1985, 1987), Giné and Zinn (1984, 1986) and Pollard (1984, 1990) (to mention a few key references), much is known about empirical processes indexed by functions as above. For our purposes it is, in view of (5.28), useful to consider the class of functions given by

$$(5.29) \quad \mathcal{M}(K) \equiv \left\{ m(x) \mathbf{1}_{]s, t]}(x) \mathbf{1}_{\{x < c\}} : 0 < s < t < 1, \right. \\ \left. m \text{ nondecreasing, } |m| \leq K \right\}.$$

By Dudley (1987), Theorem 2.1, the class of nondecreasing functions $f: [0, 1] \rightarrow [0, K]$, $K > 0$, is a Q -Donsker class for any probability Q on $[0, 1]$. It is a simple corollary that in particular the class $\mathcal{M}(K)$ defined above is a \mathbf{P} -Donsker class (recall $\mathbf{P} = F \times G$). This fact is crucial for our asymptotic results. For related uses of empirical processes in survival analysis, see Pollard (1990), Chapter 13.

5.2. Auxiliary results for the censored case. The following lemmas provide some inequalities which are important for determining the asymptotic distribution of the β_n process.

LEMMA 5.1. Define $J_n \equiv Y_n(1)^{-1}$ and $K_n \equiv \sup_{0 < u < 1} |Y_n(u)^{-1} - p(u)^{-1}| = O_p(n^{-1/2})$. We then have for all $s < t$ that

$$\begin{aligned} & |\pi_n(0, s, t) - \pi_n(\beta_n(s, t), s, t)| \\ & \leq 8J_n \exp(J_n) \left(\sup_{u \in]s, t]} \left| \int_s^u \frac{dN_n(v) - dEN_n(v)}{Y_n(v)} \right| + CK_n(t - s) \right) \\ & \rightarrow 0 \quad a.s., \end{aligned}$$

where C is a constant.

PROOF. We have by the definition of $\beta_n(s, t)$ that

$$(5.30) \quad \begin{aligned} & |\pi_n(0, s, t) - \pi_n(\beta_n(s, t), s, t)| \\ & = \left| \pi_n(0, s, t) - \frac{\bar{F}(t)}{\bar{F}(s)} \right| \\ & = \left| \prod_{]s, t]} \left(1 - \frac{dN_n(u)}{Y_n(u)} \right) - \prod_{]s, t]} \left(1 - \frac{dEN_n(u)}{p(u)} \right) \right| \\ & = \left| \prod_{]s, t]} (1 - \alpha_n(s, t, du)) - \prod_{]s, t]} (1 - \alpha(s, t, du)) \right|, \end{aligned}$$

where we define for fixed s, t the additive interval functions

$$\alpha_n(s, t, u) \equiv \begin{cases} 0, & \text{if } u \leq s, \\ \int_s^u Y_n(v)^{-1} dN_n(v), & \text{if } s \leq u \leq t, \\ \int_s^t Y_n(v)^{-1} dN_n(v), & \text{if } t < u \end{cases}$$

and $\alpha(s, t, u)$ similarly by replacing N_n by EN_n and Y_n by p . Then (5.30) is, for fixed s, t by Theorem 7 of Gill and Johansen (1990) (which is an application of Duhamel's equation), bounded by

$$(5.31) \quad 4 \exp(\|\alpha_n(s, t, \cdot)\|_\infty) \|\alpha_n(s, t, \cdot)\|_v \|\mu\|_v \|\alpha_n(s, t, \cdot) - \alpha_0(s, t, \cdot)\|_\infty,$$

where we use these authors' notation and define $\mu(u) \equiv \prod_{]0, u]} (1 - \alpha(s, t, dv)) = \bar{F}((s \vee u) \wedge t) / \bar{F}(s)$. It follows that $\|\mu\|_v < 1$. Notice also by the monotonicity of Y_n that $\|\alpha_n(s, t, \cdot)\|_\infty \leq (\inf_{u \in]s, t]} Y_n(u))^{-1} \leq J_n$. Furthermore, $\|\alpha_n(s, t, \cdot)\|_v = \|\alpha_n(s, t, \cdot)\|_\infty \leq J_n$ holds since $\alpha_n(s, t, u)$ is monotonically increasing in u . For fixed $s < t$ we can hence bound (5.30) by

$$\begin{aligned} & 4J_n e^{J_n} \|\alpha_n(s, t, \cdot) - \alpha(s, t, \cdot)\|_\infty \\ & \leq 8J_n e^{J_n} \sup_{u \in]s, t]} \left| \int_s^u \frac{dN_n(v)}{Y_n(v)} - \frac{dEN_n(v)}{p(v)} \right| \\ & \leq 8J_n \exp(J_n) \sup_{u \in]s, t]} \left| \int_s^u \frac{dN_n(v) - dEN_n(v)}{Y_n(v)} \right. \\ & \quad \left. - \int_s^u \left(\frac{1}{p(v)} - \frac{1}{Y_n(v)} \right) dEN_n(v) \right| \\ & \leq 8J_n \exp(J_n) \left(\sup_{u \in]s, t]} \left| \int_s^u \frac{dN_n(v) - dEN_n(v)}{Y_n(v)} \right| \right. \\ & \quad \left. + \left(\int_s^t dEN_n(v) \right) \sup_{u \in]s, t]} \left| \frac{1}{p(u)} - \frac{1}{Y_n(u)} \right| \right) \\ & \leq 8J_n \exp(J_n) \left(\sup_{u \in]s, t]} \left| \int_s^u \frac{dN_n(v) - dEN_n(v)}{Y_n(v)} \right| + CK_n(t - s) \right) \end{aligned}$$

for $C \equiv \sup_{0 < t < 1} \lambda(u)p(u) < \infty$. Since $Y_n(t)$ is eventually uniformly bounded away from 0 a.s., it follows by a Glivenco-Cantelli theorem and a central limit theorem for $n^{1/2}(Y_n - p)$ that $K_n = O_p(n^{-1/2})$ and $K_n \rightarrow 0$ a.s. Furthermore, the functions

$$(x, c) \rightarrow \frac{1}{Y_n(x)} \mathbf{1}_{]s, u]}(x) \mathbf{1}_{\{x < c\}},$$

$0 < s < u < 1$, $n = 1, 2, \dots$, will eventually belong to the class $\mathcal{M}(K)$ defined in (5.29) for some suitable $K > 0$. Hence, by a functional Glivenko–Cantelli theorem for the class $\mathcal{M}(K)$, it holds for n large enough that

$$\sup_{u \in]s, t]} \left| \int_s^u \frac{dN_n(v) - dEN_n(v)}{Y_n(v)} \right| \leq \sup_{f \in \mathcal{M}(K)} |\mathbf{P}_n(f) - \mathbf{P}(f)| \rightarrow 0 \quad \text{a.s.} \quad \square$$

LEMMA 5.2. *It holds for any $0 < s < t < 1$ that*

$$|\pi_n(\beta_n(s, t), s, t) - \pi_n(0, s, t)| \geq (t - s) |\beta_n(s, t)| L_n,$$

where $\{L_n\}$ is a sequence of random variables which is uniformly bounded away from 0 a.s.

PROOF. Duhamel's equation shows that

$$\begin{aligned} & |\pi_n(\beta_n(s, t), s, t) - \pi_n(0, s, t)| \\ &= \left| \int_s^t \pi_n(0, s, u -) \pi_n(\beta_n(s, t), u, t) \right. \\ & \quad \left. \times \left(\frac{1}{Y_n(u)} - \frac{1}{Y_n(u) + \beta(s, t)} \right) dN_n(u) \right| \\ (5.32) \quad & \geq \pi_n(0, s, t) \pi_n(\beta_n(s, t), s, t) \int_s^t \frac{dN_n(u)}{Y_n(u)(Y_n(u) + \beta_n(s, t))} |\beta_n(s, t)| \\ & \geq \pi_n(0, s, t) \frac{\bar{F}(t)}{\bar{F}(s)} \frac{1}{Y_n(s)} \int_s^t \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} |\beta_n(s, t)| \\ & \geq \pi_n(0, s, t) \frac{\bar{F}(t)}{\bar{F}(s)} \left(\int_s^t \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} \right) |\beta_n(s, t)| \\ & \geq \pi_n(0, s, t) \bar{F}(1) \left(\int_s^t \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} \right) |\beta_n(s, t)|. \end{aligned}$$

Here we have used the monotonicity of $Y_n(u)$ and π_n , that $Y_n(u) \leq$ and that $\pi_n(\beta_n(s, t), s, t) = \bar{F}(t)/\bar{F}(s)$. Since $\pi_n(0, s, t) \geq \pi_n(0, 0, 1) \rightarrow \bar{F}(1) > 0$ a.s., the lemma will follow if it can be shown that

$$(5.33) \quad \int_s^t \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} \geq \text{const} \int_s^t \lambda(u) du \quad \text{a.s.}$$

for some positive constant, for all $s < t$, $n > 0$. For this we use the inequality $1 - x \geq \exp(-c(\eta)x)$ for all $0 \leq x \leq 1 - \eta$, where $c(\eta) = -(\log \eta)/(1 - \eta)$ [Gill and Johansen (1990), page 1533], that is, $x \geq -c(\eta)^{-1} \log(1 - x)$, $0 \leq x \leq 1 - \eta$. It follows that for an $\eta < 1$ such that

$$(5.34) \quad \frac{1/n}{Y_n(u) + \beta_n(s, t)} \leq 1 - \eta \quad \text{when } s < u \leq t,$$

it holds by the definition of $\beta_n(s, t)$ that

$$\begin{aligned} \int_s^t \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} &\geq -\frac{1}{c(\eta)} \sum_{s < u \leq t} \log \left(1 - \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} \right) \\ &= -\frac{1}{c(\eta)} \log \prod_{]s, t]} \left(1 - \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} \right) \\ &= \frac{1}{c(\eta)} \int_s^t \lambda(u) du. \end{aligned}$$

We can assume that at least one death occurs in $]s, t]$ [if not, the value of $\beta_n(s, t)$ is immaterial for the maximum likelihood estimate and may, e.g., be defined as 0]. Letting τ^* denote the last time of death in $]s, t]$, it must hold by the definition of $\beta_n(s, t)$ that

$$1 - \frac{1/n}{Y_n(\tau^*) + \beta_n(s, t)} \geq \prod_{]s, t]} \left(1 - \frac{dN_n(u)}{Y_n(u) + \beta_n(s, t)} \right) = \frac{\bar{F}(t)}{\bar{F}(s)} \geq \bar{F}(1) > 0,$$

and hence $\eta \equiv \bar{F}(1) > 0$ will satisfy (5.34).

It follows that (5.33) is satisfied, and hence the lemma follows. \square

By combining Lemmas 5.1 and 5.2, we get the following corollaries.

COROLLARY 5.1.

$$\sup_{t-s > \delta} |\beta_n(s, t)| \rightarrow 0 \quad a.s., \text{ all } \delta > 0.$$

PROOF. This follows immediately by Lemmas 5.1 and 5.2. \square

COROLLARY 5.2.

$$\limsup_{n \rightarrow \infty} \sup_{0 < t-s < \delta} (t-s)\sqrt{n} |\beta_n(s, t)| \rightarrow_p 0 \quad \text{as } \delta \rightarrow 0.$$

PROOF. By Lemmas 5.1 and 5.2 it holds for any $s < t$ that

$$\begin{aligned} &\sqrt{n} (t-s) |\beta_n(s, t)| \\ &\leq 8L_n^{-1} J_n \exp(J_n) \left(\sup_{s < u \leq t} \sqrt{n} \left| \int_s^u \frac{dN_n(x) - dEN_n(x)}{Y_n(x)} \right| + C\sqrt{n} (t-s) K_n \right), \end{aligned}$$

and hence

$$\begin{aligned} &\sup_{0 < t-s < \delta} (t-s)\sqrt{n} |\beta_n(s, t)| \\ (5.35) \quad &\leq O_p(1) \left(\sup_{0 < s < u < s+\delta} \sqrt{n} \left| \int_s^u \frac{dN_n(x) - dEN_n(x)}{Y_n(x)} \right| + C\sqrt{n} \delta K_n \right). \end{aligned}$$

Recall the definition of $\mathcal{M}(K)$ and let $\mathcal{M}_\delta(K)$ denote the subclass of functions $m(x)1_{]s, u]}(x)1_{\{x < c\}} \in \mathcal{M}(K)$ for which $u - s < \delta$. Then the right-hand side of (5.35) is, for n large enough, bounded by

$$O_P(1) \sup_{m^* \in \mathcal{M}_\delta(K)} (|\mathbf{X}_n(m^*)| + C\sqrt{n} K_n \delta),$$

and the corollary will, in view of $K_n = O_P(n^{-1/2})$, follow if it can be shown that

$$(5.36) \quad \limsup_{n \rightarrow \infty} \sup\{|\mathbf{X}_n(m^*)| : m^* \in \mathcal{M}_\delta(K)\} \rightarrow_P 0 \quad \text{as } \delta \rightarrow 0.$$

Let (X, C) be an observation from the joint probability $F \times G \equiv \mathbf{P}$ and define $\sigma^2(f) \equiv \text{Var } f(X, C)$ for any function $f(x, c)$. Then (5.36) will follow from, for example, Giné and Zinn (1986), Theorem 1.3, if it can be shown that

$$(5.37) \quad \limsup_{\delta \rightarrow 0} \{\sigma^2(m^*) : m^* \in \mathcal{M}_\delta(K)\} = 0.$$

[The property we are using here is a.s. uniform continuity of the Gaussian limit of \mathbf{X}_n with respect to the natural variance pseudometric on $\mathcal{M}(K)$.] (5.37) finally follows because, for $m^* \in \mathcal{M}_\delta(K)$,

$$\sigma^2(m^*) \leq Em^*(X, C)^2 \leq K^2\delta \rightarrow 0 \quad \text{as } \delta \rightarrow 0,$$

in turn proving (5.36) and the corollary. \square

What the previous corollaries do not state is that $\sup_{0 < t < 1} |\beta_n(0, t)| \rightarrow 0$ a.s., and, in fact, this cannot be expected to hold. We need, however, to know something about the size of $\beta_n(0, t)$ for t close to 0. This is the content of the following lemma and its corollary. They are both of the ‘‘Chibisov–O’Reilly’’ type.

LEMMA 5.3. *Let $h(t)$ be a nonincreasing, nonnegative function on $]0, 1]$ such that $h(t) = o((t \log_2(1/t))^{-1/2})$ as $t \rightarrow 0$. Then*

$$\sup_{0 < u < t < 1} \sqrt{n} \left| h(u) \int_0^u \frac{dN_n(x) - dEN_n(x)}{Y_n(x)} \right| = O_P(1).$$

PROOF. By Dudley [(1985), Corollary 7.1 and Theorem 6.3], the class of functions

$$\mathcal{F}_1 \equiv \{h(u)1_{]0, u]}(x) : 0 < u < 1\}$$

is a functional Donsker class for the Lebesgue measure on $[0, 1]$, and hence also for the continuous distribution F . Since h is nonincreasing and nonnegative, the collection

$$\mathcal{F}_2 \equiv \{h(v)1_{]0, u]}(x) : 0 < u \leq v < 1\}$$

is contained in the convex hull of \mathcal{F}_1 and is F -Donsker by Dudley [(1985), Theorem 5.3]. The difference of two F -Donsker classes is again F -Donsker; hence, in particular, the collection

$$\mathcal{F}_3 \equiv \{h(v)1_{]u, v]}(x) : 0 < u \leq v < 1\}$$

is F -Donsker. The convex hull of \mathcal{F}_3 , in turn, contains the class

$$\{h(u)m(x)1_{]0, u]}(x) : m \in \mathcal{M}, 0 < u < 1\},$$

where \mathcal{M} denotes the set of all nonnegative, nondecreasing functions on $[0, 1]$ which are bounded above by 1. Finally, by multiplying this collection by the fixed function $1_{\{x < c\}}$, the class

$$(5.38) \quad \mathcal{G} \equiv \left\{g(x; m, u) \equiv h(u)m(x)1_{]0, u]}(x)1_{\{x < c\}} : 0 < u < 1, m \in \mathcal{M}\right\}$$

is a functional $F \times G$ Donsker class.

Let $\mathbf{X}_n = n^{1/2}(\mathbf{P}_n - \mathbf{P})$ denote the empirical process based on life and censoring times (and $\mathbf{P} = F \times G$). Then we may write

$$\sqrt{n}h(u) \int_0^u \frac{dN_n(x) - dEN_n(x)}{Y_n(x)} = \mathbf{X}_n(g(x; Y_n(\cdot)^{-1}, u)).$$

It eventually holds a.s. that $Y_n(t)^{-1}$ is uniformly bounded on $[0, 1]$, and, by a functional central limit theorem for \mathcal{G} , we hence have that

$$\sup_{0 < u < t < 1} \sqrt{n} \left| h(u) \int_0^u \frac{dN_n(x) - dEN_n(x)}{Y_n(x)} \right| \leq \sup_{g \in \mathcal{G}} |\mathbf{X}_n(g)| = O_p(1). \quad \square$$

COROLLARY 5.3. *Let h be a function as in Lemma 5.3. Then it holds that*

$$\sup_{0 < t < 1} |\sqrt{n}th(t)\beta_n(0, t)| = O_p(1).$$

PROOF. By Lemma 5.1 and 5.2 it holds that

$$\begin{aligned} & \sup_{0 < t < 1} |\sqrt{n}th(t)\beta_n(0, t)| \\ & \leq O_p(1)\sqrt{n} \sup_{0 < t < 1} h(t)|\pi_n(0, 0, t) - \pi_n(\beta_n(0, t), 0, t)| \\ & \leq O_p(1) \sup_{0 < u < t < 1} \left(\left| h(t)\sqrt{n} \int_0^u \frac{dN_n(x) - dEN_n(x)}{Y_n(x)} \right| + \sqrt{n}K_n h(t)t \right) \\ & \leq O_p(1) \left(\sup_{0 < u < t < 1} \left| h(u)\sqrt{n} \int_0^u \frac{dN_n(x) - dEN_n(x)}{Y_n(x)} \right| \right. \\ & \qquad \qquad \qquad \left. + \sup_{0 < t < 1} \sqrt{n}K_n h(t)t \right), \end{aligned}$$

where we have used that $h(t)$ is nonincreasing. The last expression is recognized as being of magnitude $O_p(1)$ by Lemma 5.3, noting also that $th(t)$ is bounded on $]0, 1]$. \square

We are now prepared to specify the asymptotic distribution of $\beta_n(s, t)$.

THEOREM 5.1. *Let (Z_N, Z_Y) be defined by*

$$(5.39) \quad \sqrt{n} (N_n - EN_n, Y_n - EY_n) \Rightarrow (Z_N, Z_Y).$$

Then it holds that

$$(5.40) \quad \mathbf{B}_n(s, t) \equiv (\phi(t) - \phi(s))\sqrt{n} \beta_n(s, t) \Rightarrow W(\phi(t)) - W(\phi(s))$$

weakly in $l^\infty[0, 1]^2$. Here $W(t)$ is a Brownian motion given in terms of (Z_N, Z_Y) as

$$W(t) \equiv \int_0^{\phi^{-1}(t)} \frac{dZ_N(u) - Z_Y(u)\lambda(u) du}{p(u)}, \quad 0 \leq t \leq 1.$$

[This identity is explained in Gill and Johansen (1990), page 1538, to whom the notation is also due. Formally, we must define $\mathbf{B}_n(s, t)$ as $\mathbf{B}_n(t, s)$ in case $s > t$.]

PROOF OF THEOREM 5.1. The space $[0, 1]^2$ is totally bounded under the usual Euclidean norm. Hence, by Pollard [(1990), Theorem 10.2, for example], it suffices to show finite-dimensional convergence of $\mathbf{B}_n(s, t)$ to $W(\phi(t)) - W(\phi(s))$ and asymptotic equicontinuity. The latter means showing that

$$(5.41) \quad \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\sup\{|\mathbf{B}_n(s, t) - \mathbf{B}_n(s', t')| : |s - s'|, |t - t'| < \delta\} > \eta) = 0$$

for all $\eta > 0$.

By the assumptions of Theorem 2.1, it holds that $\phi(t) - \phi(s) = \int_s^t \lambda(u)/p(u) du \leq \text{const}(t - s)$ for an appropriate constant, and by Corollary 5.2 we see that (5.41) holds if the supremum is taken over (s, t) and (s', t') such that $|t - s|, |t' - s'| < \delta$. It hence suffices for (5.41) to show that

$$(5.42) \quad \lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\sup\{|\mathbf{B}_n(s, t) - \mathbf{B}_n(s', t')| : |s - s'|, |t - t'| < \delta, |t - s|, |t' - s'| > \rho\} > \eta) = 0$$

for all $\eta, \rho > 0$.

In particular, (5.42) will follow if we can show that $\mathbf{B}_n(s, t)$ converges in distribution to $W(\phi(t)) - W(\phi(s))$ when we restrict $|t - s|$ to be bounded away from 0, that is,

$$(5.43) \quad \mathbf{B}_n(s, t) \Rightarrow \mathbf{B}(s, t)$$

in $l^\infty(\{(s, t) \in [0, 1]^2 : |t - s| \geq \rho\})$ for all $\rho > 0$.

Note now that $\pi_n(\beta, s, t)$ is differentiable w.r.t. β wherever it is defined and that

$$\begin{aligned}
 \frac{\partial}{\partial \beta} [\pi_n(\beta, s, t)] &= \frac{\partial}{\partial \beta} [\exp(\log \pi_n(\beta, s, t))] \\
 &= \pi_n(\beta, s, t) \frac{\partial}{\partial \beta} \left[\sum_{u \in]s, t]} \log \left(1 - \frac{dN_n(u)}{Y_n(u) + \beta} \right) \right] \\
 (5.44) \quad &= \pi_n(\beta, s, t) \sum_{\tau \in]s, t]} \left(\frac{1}{Y_n(\tau) + \beta - n^{-1}} - \frac{1}{Y_n(\tau) + \beta} \right) \\
 &= \pi_n(\beta, s, t) \int_{]s, t]} \frac{dN_n(u)}{(Y_n(u) + \beta)(Y_n(u) + \beta - n^{-1})},
 \end{aligned}$$

where τ denotes an observed time of death. By the mean value theorem we get that

$$\begin{aligned}
 \sqrt{n} (\pi_n(0, s, t) - \pi_n(\beta_n(s, t), s, t)) \\
 (5.45) \quad &= \sqrt{n} (-\beta_n(s, t)) \frac{\partial}{\partial \beta} [\pi_n(\beta_n^*(s, t), s, t)],
 \end{aligned}$$

where $|\beta_n^*(s, t)| < |\beta_n(s, t)|$ for all $0 < s < t < 1$. Since the first line in (5.45) is just the scaled and centered Kaplan–Meier estimator, it holds as demonstrated in Gill and Johansen (1990), page 1537, that

$$\begin{aligned}
 \sqrt{n} (\pi_n(0, s, t) - \pi_n(\beta_n, s, t)) &= \sqrt{n} \left(\prod_{]s, t]} \left(1 - \frac{dN_n(u)}{Y_n(u)} \right) - \frac{\bar{F}(t)}{\bar{F}(s)} \right) \\
 (5.46) \quad &\Rightarrow -\frac{\bar{F}(t)}{\bar{F}(s)} (W(\phi(t)) - W(\phi(s))).
 \end{aligned}$$

[It should be noted that the result quoted from Gill and Johansen (1990) is stated as weak convergence in $D[0, T]$ in the sense of Pollard (1984); since the limiting distribution has uniformly continuous and bounded sample paths, convergence in distribution in $l^\infty[0, 1]^2$ also follows. For more on the δ -method in the setting of Hoffman-Jørgensen weak convergence, see Wellner (1989) or Van der Vaart and Wellner (1995).]

With (5.46) established, (5.43) will follow if it can be shown that

$$(5.47) \quad \sup_{t-s \geq \rho} \sup_{\beta \in]0, \beta_n(s, t)]} \left| \frac{\frac{\bar{F}(t)}{\bar{F}(s)} \int_s^t \frac{\lambda(u)}{p(u)} du}{\frac{\partial}{\partial \beta} [\pi_n(\beta, s, t)]} - 1 \right| \rightarrow_p 0, \quad \text{all } \tau > 0.$$

By the monotonicity of $\beta \rightarrow \pi_n(\beta, s, t)$ it holds for $\beta \in]0, \beta_n(s, t)[$ that

$$(5.48) \quad \left| \pi_n(\beta, s, t) - \frac{\bar{F}(t)}{\bar{F}(s)} \right| = \left| \pi_n(\beta, s, t) - \pi_n(\beta_n(s, t), s, t) \right| \\ \leq \left| \pi_n(0, s, t) - \pi_n(\beta_n(s, t), s, t) \right|$$

and hence the right-hand side of (5.48) is, for fixed $\tau > 0$, bounded by

$$(5.49) \quad \sup_{t-s \geq \rho} \left| \pi_n(0, s, t) - \pi_n(\beta_n(s, t), s, t) \right| = o_p(1),$$

where the equality in (5.49) follows by (5.46). Since $\pi_n(0, s, t)$ and $\pi_n(\beta_n(s, t), s, t) = \bar{F}(t)/\bar{F}(s) \geq \bar{F}(1)$ are both uniformly bounded away from 0 with probability tending to 1, it holds by (5.49) that

$$(5.50) \quad \sup_{t-s \geq \rho} \sup_{\beta \in]0, \beta_n(s, t)[} \left| \frac{\bar{F}(t)/\bar{F}(s)}{\pi_n(\beta, s, t)} - 1 \right| \rightarrow_p 0, \quad \text{all } \rho > 0.$$

In view of (5.50) it remains for (5.47) to show that

$$(5.51) \quad \sup_{t-s \geq \rho} \sup_{\beta \in]0, \beta_n(s, t)[} \left| \frac{\int_s^t \frac{\lambda(u) du}{p(u)}}{\int_s^t \frac{dN_n(u)}{(Y_n(u) + \beta)(Y_n(u) + \beta - n^{-1})}} - 1 \right| \\ \rightarrow_p 0 \quad \text{for all } \tau > 0.$$

By Corollary 5.1 it holds that

$$\liminf_{n \rightarrow \infty} \inf_{t-s \geq \rho} \inf_{\beta \in]0, \beta_n(s, t)[} (Y_n(t) + \beta) \left(Y_n(t) + \beta - \frac{1}{n} \right) > 0 \quad \text{a.s., all } \rho > 0.$$

Hence all functions of the form

$$q(x, y; n, s, t, \beta) \equiv \frac{1}{(Y_n(x) + \beta)(Y_n(x) + \beta - n^{-1})} \mathbf{1}_{]s, t[}(x) \mathbf{1}_{\{x < y\}},$$

where $t-s \geq \rho$, $\beta \in]0, \beta_n(s, t)[$, will, for n sufficiently large, satisfy $q(\cdot, \cdot; n, s, t, \beta) \in \mathcal{M}(K)$ for an appropriate K where the class $\mathcal{M}(K)$ is defined in (5.29). We then see that

$$\int_s^t \frac{dN_n(u)}{(Y_n(u) + \beta)(Y_n(u) + \beta - n^{-1})} = \mathbf{P}_n q(\cdot, \cdot; n, s, t, \beta)$$

and

$$\int_s^t \frac{\lambda(u) p(u) du}{(Y_n(u) + \beta)(Y_n(u) + \beta - n^{-1})} = \mathbf{P} q(\cdot, \cdot; n, s, t, \beta).$$

Hence a Glivenco–Cantelli theorem for $\mathcal{M}(K)$ and uniform convergence of $Y_n(t)$ to $p(t)$ a.s. shows that

$$\sup_{t-s \geq \rho} \sup_{\beta \in]0, \beta_n(s, t)[} \left| \int_s^t \frac{dN_n(u)}{(Y_n(u) + \beta)(Y_n(u) + \beta - n^{-1})} - \int_s^t \frac{\lambda(u)}{p(u)} du \right| \rightarrow 0 \text{ a.s., all } \rho > 0.$$

Since $\int_s^t \lambda(u)/p(u) du = \phi(t) - \phi(s)$ is bounded away from 0 when $t - s \geq \rho$, the theorem follows. \square

For the purpose of applying the min–max functional to the $\beta_n(s, t)$ process, we state explicitly the following corollary in which we make a time transformation by the monotonically increasing function $\phi^{-1}(t)$. Let $\tau \equiv \phi(1)$.

COROLLARY 5.4.

$$(5.52) \quad \begin{aligned} &\sqrt{n} (t - s) (\beta_n(\phi^{-1}(s), \phi^{-1}(t))) \\ &\Rightarrow W(t) - W(s) \text{ weakly in } l^\infty([0, \tau]^2), \end{aligned}$$

where W is Brownian motion.

The following lemma is used with Corollary 5.4 to find the limit of the adjustments to the risk set. It is closely related to the “min–max” formulas used in isotonic regression; see Barlow, Bartholomew, Bremner and Brunk [(1972), page 19], or Robertson, Wright and Dykstra [(1988), Theorem 1.4.4]. Theorem 1 in Gill (1995) is similar in flavor; we have, however, not been able to find a result which could be quoted immediately.

LEMMA 5.4 (A min–max continuous mapping theorem). *Suppose that $b_n, n = 1, 2, \dots$, are bounded functions in $0 < s < t < \tau$ and that $w(t)$ is a bounded and continuous function in $0 < t < \tau$ for some $\tau > 0$. Suppose that they all satisfy*

$$\sup_{0 < s < t < \tau} |\sqrt{n} (t - s) b_n(s, t) - (w(s) - w(t))| \rightarrow 0.$$

Let κ denote the smallest concave majorant of w :

$$\kappa(t) \equiv \inf\{k(t) : k \text{ concave, } k(u) \geq w(u), \text{ all } u\}.$$

Let $T \equiv \{t : \kappa(t) = w(t)\}$. Then it holds for any $t \notin T$ that

$$(5.53) \quad \sqrt{n} \gamma_n(t) \vee 0 \equiv \left(\sqrt{n} \min_{s \leq t} \max_{u \geq t} b_n(s, u) \right) \vee 0 \rightarrow \dot{\kappa}(t) \vee 0,$$

where $\dot{\kappa}(t)$ denotes the derivative of κ at t .

PROOF. Let $t \notin T$. Define (s_0, u_0) by

$$(5.54) \quad (s_0, u_0) = \arg \min_{s \leq t} \max_{u \geq t} \frac{w(u) - w(s)}{u - s}.$$

[Since $t \notin T$, (s_0, u_0) is well defined; should more than one arg min–max exist, then we choose the version closest to t .] It then follows that $(w(u_0) - w(s_0))/(u_0 - s_0)$ is the slope of the smallest concave majorant of w over the set $]s_0, u_0[$, that is, $(w(u_0) - w(s_0))/(u_0 - s_0) = \dot{\kappa}(t^*)$ for all $t^* \in]s_0, u_0[$. Furthermore, $s_0 < t - \delta$ for some $\delta > 0$, since $t \notin T$. We now consider for this δ :

$$(5.55) \quad \gamma_n(t, \delta) = \min_{s < t - \delta} \max_{u \geq t} b_n(s, u)$$

and show that

$$(5.56) \quad \sqrt{n} \gamma_n(t, \delta) \rightarrow \dot{\kappa}(t), \quad n \rightarrow \infty.$$

Notice that by assumption there exists a sequence $\varepsilon_n \downarrow 0$ so that, for all $s < u < \tau$,

$$(5.57) \quad \sqrt{n} b_n(s, u) = \frac{(w(u) - w(s)) + \varepsilon_n}{u - s},$$

where the sequence ε_n does not depend on s, u . By the definition of (s_0, u_0) , noting that $u - s > \delta$ in (5.55), this shows that

$$\begin{aligned} \sqrt{n} \gamma_n(t, \delta) &\in \left[\frac{w(u_0) - w(s_0)}{u_0 - s_0} - \frac{\varepsilon_n}{\delta}, \frac{w(u_0) - w(s_0)}{u_0 - s_0} + \frac{\varepsilon_n}{\delta} \right] \\ &= \left[\dot{\kappa}(t) - \frac{\varepsilon_n}{\delta}, \dot{\kappa}(t) + \frac{\varepsilon_n}{\delta} \right]. \end{aligned}$$

Since $\varepsilon_n \downarrow 0$, this proves (5.56).

Note now that, for any $s \in [t - \delta, t[$,

$$\frac{w(u_0) - w(s)}{u_0 - s} \geq \frac{w(u_0) - w(s_0)}{u_0 - s_0} = \dot{\kappa}(t).$$

Hence we have that

$$\begin{aligned} \min_{t - \delta \leq s \leq t} \max_{u \geq t} \sqrt{n} b_n(s, u) &\geq \min_{t - \delta \leq s < t} \sqrt{n} b_n(s, u_0) \\ &= \min_{t - \delta \leq s < t} \frac{w(u_0) - w(s) + \varepsilon_n}{u_0 - s} \\ &\geq \frac{w(u_0) - w(s_0)}{u_0 - s_0} - \frac{\varepsilon_n}{u_0 - t} \\ &\rightarrow \dot{\kappa}(t). \end{aligned}$$

Hence $\liminf_{n \rightarrow \infty} \min_{t - \delta \leq s < t, s \in B} \max_{u > t, u \in B} n^{1/2} b_n(s, u) \geq \dot{\kappa}(t)$, and this, together with (5.56), shows the desired result. \square

The following lemma is a variant on Gill (1989), Lemma 3, suited for the situation where the integrand is the slope of the least convex majorant of a Brownian motion. We let $\|\cdot\|$ denote the sup norm on $l^\infty[0, 1]$.

LEMMA 5.5. Let $x_n, x, y_n, y \in l^\infty[0, 1]$ and set $h_n = t_n(x_n - x)$, $k_n = t_n(y_n - y)$ for a real sequence $t_n \rightarrow \infty$. Suppose $\int_0^1 |dy| < \infty$, $\limsup_{n \rightarrow \infty} \int_0^1 |dy_n| < \infty$ and $\|t_n(y_n - y) - k\| \rightarrow 0$ for some $k \in l^\infty[0, 1]$. Suppose that for all $\varepsilon > 0$ there exists a set $A_\varepsilon \subset [0, 1]$ with Lebesgue measure $|A_\varepsilon| > 1 - \varepsilon$ so that

$$(5.58) \quad \sup_{t \in A_\varepsilon} |h_n(t) - h(t)| \rightarrow 0 \quad \text{for all } \varepsilon > 0,$$

where h is a real function which is bounded on A_ε and has for all $\varepsilon > 0$ only a finite number of discontinuities on A_ε and which satisfies $\int_0^1 |h| |dy| < \infty$. Suppose, however, that

$$(5.59) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_{A_\varepsilon^c} |h_n| |dy_n| = 0,$$

and that (5.59) also holds for the combinations $\int |h_n| |dy|$ and $\int |h| |dy_n|$. Then it is true that

$$\left\| t_n \left(\int_0^{\cdot} x_n dy_n - \int_0^{\cdot} x dy \right) - \int_0^{\cdot} h dy + \int_0^{\cdot} x dk \right\| \rightarrow 0.$$

REMARK. Unlike the situation in Gill's proof, $h \notin l^\infty[0, 1]$ may occur.

PROOF. We modify Gill's proof as follows. First of all

$$\begin{aligned} & \left\| \int_0^{\cdot} h_n dy + \int_0^{\cdot} x dk_n - \left(\int_0^{\cdot} h dy + \int_0^{\cdot} x dk \right) \right\| \\ & \leq \left\| \int_0^{\cdot} (h_n - h) dy \right\| + \left\| \int_0^{\cdot} x (dk - dk_n) \right\|. \end{aligned}$$

Since x, k_n and k are as in Gill's proof, it holds that $\|\int_0^{\cdot} x (dk_n - dk)\| \rightarrow 0$, and furthermore, for any $\varepsilon > 0$,

$$\left\| \int_0^{\cdot} (h_n - h) dy \right\| \leq \sup_{t \in A_\varepsilon} |h_n(t) - h(t)| \left(\int_0^1 |dy| \right) + \int_{A_\varepsilon^c} (|h_n| + |h|) |dy|,$$

which by (5.58) and (5.59), since $\varepsilon > 0$ can be chosen arbitrarily small, shows that

$$\left\| \int_0^{\cdot} (h_n - h) dy \right\| \rightarrow 0.$$

We next need to prove that $\|\int_0^{\cdot} h_n d(y_n - y)\| \rightarrow 0$. We use the bound

$$\begin{aligned} \left\| \int_0^{\cdot} (h_n - h) d(y_n - y) \right\| & \leq \sup_{t \in A_\varepsilon} |h_n(t) - h(t)| \left(\int_0^1 d(|y_n| + d|y|) \right) \\ & \quad + \int_{A_\varepsilon^c} (|h_n| + |h|) d(|y_n| + |y|), \end{aligned}$$

which, again by (5.58) and (5.59), shows that $\|\int_0^{\cdot} (h_n - h) d(y_n - y)\| \rightarrow 0$. It now suffices to show that $\|\int_{A_\varepsilon} h(dy_n - dy)\| \rightarrow 0$ for all $\varepsilon > 0$. But since the restriction of h to A_ε is bounded and has only a finite number of discontinuities, h can be arbitrarily closely approximated in supremum norm by a

step function with a finite number of steps. The remaining part of Gill's proof now carries through. \square

5.3. Proof for the censored case.

PROOF OF THEOREM 2.1. Let $K(t)$ denote the smallest nondecreasing concave majorant over $[0, \phi(1)] \equiv [0, \tau]$ of a standard Brownian motion W . The we shall first show that

$$(5.60) \quad \begin{aligned} \sqrt{n} \left(\int_0^t \frac{dN_n(u)}{Y_n(u) + \gamma_n(u)} - \int_0^t \lambda(u) du \right) \\ \Rightarrow W(\phi(t)) - K(\phi(t)) \quad \text{in } l^\infty[0, 1]. \end{aligned}$$

Let \mathbf{W}_n denote the random element of $l^\infty[0, 1]^2 \times l^\infty[0, \tau]^2$, given by

$$\begin{aligned} \mathbf{W}_n(u, v, s, t) \equiv \sqrt{n} (N_n(u) - EN_n(u), Y_n(v) - EY_n(v), \\ (t - s)\beta_n(\phi^{-1}(s), \phi^{-1}(t))); \end{aligned}$$

then it is by Corollary 5.4 the case that

$$(5.61) \quad \begin{aligned} \mathbf{W}_n(u, v, s, t) \Rightarrow \\ \mathbf{W}(u, v, s, t) \equiv (Z_N(u), Z_Y(v), \mathbf{W}(t) - \mathbf{W}(s)) \\ \text{weakly in } l^\infty[0, 1]^2 \times l^\infty[0, \tau]^2. \end{aligned}$$

Here $W(t)$ is related to (Z_N, Z_Y) by $W(\phi(t)) = \int_0^t (dZ_N(u) - Z_Y(u)\lambda(u) du) / p(u)$. We need to be able to assume that (5.61) holds along almost every sample path; that is, we need an almost sure representation theorem. The almost sure representation theorem in the Hoffmann-Jørgensen theory of weak convergence is due to Dudley (1985); see also Kim and Pollard (1990), Theorem 2.2, and Van der Vaart and Wellner (1995). Assume now that the data and hence also \mathbf{W}_n are defined on some underlying probability space (Ω, \mathcal{G}, P) . Dudley's representation theorem guarantees the existence of a new probability space $(\tilde{\Omega}, \tilde{\mathcal{G}}, \tilde{P})$, a sequence of perfect, measurable maps $\psi_n: \tilde{\Omega} \rightarrow \Omega$ and a Borel measurable random element $\tilde{\mathbf{W}}$ defined on $\tilde{\Omega}$ so that the following conditions all hold:

- (i) $P = \phi_n(\tilde{P})$ for all n ,
- (ii) $\tilde{\mathbf{W}} =_d \mathbf{W}$,
- (iii) $\|\mathbf{W}_n \circ \psi_n(\tilde{\omega}) - \tilde{\mathbf{W}}^\omega\| \rightarrow 0$ a.s.,

where $\|\cdot\|$ denotes the (product) supremum norm on $l^\infty[0, 1]^2 \times l^\infty[0, \tau]^2$. The notion of a perfect mapping is explained in the references above. It is a concept which is necessary because the mappings \mathbf{W}_n are not necessarily measurable, and hence one needs to be careful when discussing almost sure representations.

Write $\tilde{W}_n^{\tilde{\omega}}$ for the composition $\tilde{W}_n \circ \psi_n(\tilde{\omega})$ and $\tilde{N}_n^{\tilde{\omega}}$ and so on for its components. By (iii) we can assume that for all $\tilde{\omega}$ in a set of P probability 1 it holds that

$$(5.62) \quad \sup_{0 < s < t < \tau} \left| \sqrt{n} (t - s) \tilde{\beta}_n^{\tilde{\omega}}(\phi^{-1}(s), \phi^{-1}(t)) - (\tilde{W}^{\tilde{\omega}}(t) - \tilde{W}^{\tilde{\omega}}(s)) \right| \rightarrow 0.$$

Next, we apply Lemma 5.4 for fixed $\tilde{\omega}$ with $\tau = \phi(1)$, $b_n(s, t) = \tilde{\beta}_n^{\tilde{\omega}}(\phi^{-1}(s), \phi^{-1}(t))$ and

$$(5.63) \quad w(t) \equiv \tilde{W}^{\tilde{\omega}}(t) = \int_0^{\phi^{-1}(t)} \frac{d\tilde{Z}_N^{\tilde{\omega}}(u) - \tilde{Z}_Y^{\tilde{\omega}}(u)\lambda(u)}{p(u)} du.$$

Let $\tilde{K}^{\tilde{\omega}}(u)$ denote the smallest nondecreasing concave majorant of $\tilde{W}^{\tilde{\omega}}(t)$ on $[0, \tau]$. Then it holds by Lemma 5.4 for any t so that the derivative $D\tilde{K}^{\tilde{\omega}}(u)$ exists at $u \equiv \phi(t)$ that

$$(5.64) \quad \begin{aligned} \sqrt{n} \tilde{\gamma}_n^{\tilde{\omega}}(t) &\equiv \sqrt{n} \left(\min_{s \leq t} \max_{u \geq t} \tilde{\beta}_n^{\tilde{\omega}}(s, t) \vee 0 \right) \\ &= \sqrt{n} \left(\min_{s \leq \phi(t)} \max_{u \geq \phi(t)} \tilde{\beta}_n^{\tilde{\omega}}(\phi^{-1}(s), \phi^{-1}(t)) \vee 0 \right) \\ &\rightarrow D\tilde{K}^{\tilde{\omega}}(\phi(t)) \vee 0 \equiv \tilde{\gamma}^{\tilde{\omega}}(t). \end{aligned}$$

We can hence describe $\tilde{\gamma}^{\tilde{\omega}}(t)$ as the derivative of the least concave *nondecreasing* majorant of $\tilde{W}^{\tilde{\omega}}(u)$ over $[0, \tau]$ at the point $u = \phi(t)$. Now, since $\tilde{W}^{\tilde{\omega}}(u)$ is a Brownian motion sample path on $[0, \tau]$, it is known from Groeneboom [(1985), page 1019, last three lines], that there are in every interval $[\delta, \tau[$ not containing 0 only finitely many points u^* in which, $\tilde{W}^{\tilde{\omega}}(u^*) = \tilde{K}^{\tilde{\omega}}(u^*)$. Furthermore, $\tilde{K}^{\tilde{\omega}}(u)$ is piecewise linear between these points; that is, $\tilde{\gamma}^{\tilde{\omega}}(t)$ is constant. It follows that there exists for each $\varepsilon > 0$ and $\tilde{\omega}$ a set $A_\varepsilon^{\tilde{\omega}} \subset]0, 1]$ with $|A_\varepsilon^{\tilde{\omega}}| > 1 - \varepsilon$ so that

$$(5.65) \quad \sup_{t \in A_\varepsilon^{\tilde{\omega}}} \left| \sqrt{n} \tilde{\gamma}_n^{\tilde{\omega}}(t) - \tilde{\gamma}^{\tilde{\omega}}(t) \right| \rightarrow 0.$$

Next, linearize to get

$$(5.66) \quad \begin{aligned} &\sqrt{n} \left(\frac{1}{\tilde{Y}_n^{\tilde{\omega}}(t) + \tilde{\gamma}_n^{\tilde{\omega}}(t)} - \frac{1}{p(t)} \right) \\ &= \left(- \frac{\sqrt{n} (\tilde{Y}_n^{\tilde{\omega}}(t) - p(t) + \tilde{\gamma}_n^{\tilde{\omega}}(t))}{p(t)(\tilde{Y}_n^{\tilde{\omega}}(t) + \tilde{\gamma}_n^{\tilde{\omega}}(t))} \right) \equiv \tilde{H}_n^{\tilde{\omega}}(t), \end{aligned}$$

which by (5.65) and (5.39) converges uniformly on $A_\varepsilon^{\tilde{\omega}}$ to

$$(5.67) \quad - \frac{\tilde{Z}_Y^{\tilde{\omega}}(t) + \tilde{\gamma}^{\tilde{\omega}}(t)}{p(t)^2} \equiv \tilde{H}^{\tilde{\omega}}(t).$$

We now apply Lemma 5.5 with $x_n(t) = (\tilde{Y}_n^{\tilde{\omega}}(t) + \tilde{\gamma}_n^{\tilde{\omega}}(t))^{-1}$, $x(t) = p(t)^{-1}$, $y_n(t) = \tilde{N}_n^{\tilde{\omega}}(t)$, $y(t) = EN_n(t) = \int_0^t p(u)\lambda(u) du$, $t_n = n^{1/2}$ and, by (5.67), $h_n(t) = \tilde{H}_n^{\tilde{\omega}}(t)$, $h(t) = \tilde{H}^{\tilde{\omega}}(t)$, $k_n(t) = {}^{1/2}(\tilde{N}_n^{\tilde{\omega}}(t) - EN_n(t))$ and $k(t) = \tilde{Z}_N^{\tilde{\omega}}$ to show that

$$(5.68) \quad \sup_{0 < t < 1} \left| \sqrt{n} \left(\int_0^t \frac{d\tilde{N}_n^{\tilde{\omega}}(v)}{\tilde{Y}_n^{\tilde{\omega}}(v) + \tilde{\gamma}_n^{\tilde{\omega}}(v)} - \int_0^t \lambda(v) dv \right) - \left(\int_0^t \tilde{H}^{\tilde{\omega}}(v) \lambda(v) p(v) dv + \int_0^t \frac{d\tilde{Z}_N^{\tilde{\omega}}(v)}{p(v)} \right) \right| \rightarrow 0.$$

By (5.67),

$$\int_0^t \tilde{H}^{\tilde{\omega}}(v) \lambda(v) p(v) dv = - \int_0^t \frac{\lambda(v) (\tilde{Z}_Y^{\tilde{\omega}}(v) + \tilde{\gamma}^{\tilde{\omega}}(v))}{p(v)} dv,$$

and, since $\dot{\phi}(v) = \lambda(v)/p(v)$, (5.68) shows that

$$(5.69) \quad \begin{aligned} & \sqrt{n} \left(\int_0^t \frac{d\tilde{N}_n^{\tilde{\omega}}(v)}{\tilde{Y}_n^{\tilde{\omega}}(v) + \tilde{\gamma}_n^{\tilde{\omega}}(v)} - \int_0^t \lambda(v) dv \right) \\ & \rightarrow \int_0^t \frac{d\tilde{Z}_N^{\tilde{\omega}}(v) - \lambda(v) (\tilde{Z}_Y^{\tilde{\omega}}(v) + \tilde{\gamma}^{\tilde{\omega}}(v))}{p(v)} du \\ & = \int_0^t \frac{d\tilde{Z}_N^{\tilde{\omega}}(v) - \lambda(v) \tilde{Z}_Y^{\tilde{\omega}}(v)}{p(v)} dv - \int_0^t \tilde{\gamma}^{\tilde{\omega}}(v) \dot{\phi}(v) dv. \end{aligned}$$

Recall by (5.63) and (5.64) that $\tilde{\gamma}^{\tilde{\omega}}(t)$ is the derivative of $\tilde{K}^{\tilde{\omega}}(u)$ evaluated at $u = \phi(t)$, and hence

$$\int_0^t \tilde{\gamma}^{\tilde{\omega}}(v) \dot{\phi}(v) dv = \int_0^t \frac{d}{du} [\tilde{K}^{\tilde{\omega}}(u)]_{u=\phi(v)} \dot{\phi}(v) dv = \tilde{K}^{\tilde{\omega}}(\phi(t)).$$

By reverting to the original sequence and noting that (5.69) is distributed as $W(\phi(t)) - K(\phi(t))$, this proves (5.62).

For applying Lemma 5.5, notice that our choices of y_n , y , k_n and k are as required in the lemma; indeed, they are the same as assumed by Gill and Johansen (1990) in their application of Gill's (1989) Lemma 5.2. The problem we face is that $\tilde{\gamma}(t)$ is unbounded and has right limit ∞ at 0. By the remark preceding (5.67) it is true that (5.58) is satisfied with $A_\varepsilon \equiv A_\varepsilon^{\tilde{\omega}}$. Furthermore, $|\tilde{H}^{\tilde{\omega}}|$ is integrable since $\tilde{Z}_Y^{\tilde{\omega}}(u)$ is uniformly continuous and $\tilde{\gamma}^{\tilde{\omega}}(t)$ is a derivative. In view of the fact that $\tilde{Z}_Y^{\tilde{\omega}}$ and $n^{1/2}(Y_n - p)$ are as in Gill and Johansen (1990) and that $p(u)$ is bounded uniformly away from 0 and $\lambda(u)p(u)$ from ∞ ,

it suffices for (5.59) to show that for every $\varepsilon > 0$ the following conditions hold for each $\tilde{\omega}$:

$$(5.70) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_0^\varepsilon \tilde{\gamma}^{\tilde{\omega}}(u) d\tilde{N}_n^{\tilde{\omega}}(u) = 0,$$

$$(5.71) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_0^\varepsilon \sqrt{n} \tilde{\gamma}_n^{\tilde{\omega}}(u) du = 0$$

and

$$(5.72) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_0^\varepsilon \sqrt{n} \tilde{\gamma}_n^{\tilde{\omega}}(u) d\tilde{N}_n^{\tilde{\omega}}(u) = 0.$$

By one of the laws of the iterated logarithm for Brownian motion [see Breiman (1968), Theorem 12.29, there exists for each $\tilde{\omega}$ a $0 < \tau^{\tilde{\omega}} < \eta$ so that $\tilde{W}^{\tilde{\omega}}(u) < (4u \log_2(1/u))^{1/2}$ when $0 < u < \tau^{\tilde{\omega}}$. Let $\phi(t) = u$. Since by concavity $\tilde{\gamma}^{\tilde{\omega}}(t) = (DK^{\tilde{\omega}})(\phi(t)) < \tilde{K}^{\tilde{\omega}}(\phi(t))/\phi(t) \leq \tilde{W}_n^{\tilde{\omega}}(\phi(t))/\phi(t)$, we get that

$$\tilde{\gamma}^{\tilde{\omega}}(t) \leq C \sqrt{\frac{4 \log_2(1/\phi(t))}{\phi(t)}}, \quad 0 < t < \phi^{-1}(\tau^{\tilde{\omega}}).$$

The right-hand side is an integrable function in t , and it holds by a strong law of large numbers for $\tilde{N}_n^{\tilde{\omega}}$ that

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_0^\varepsilon \sqrt{\frac{4 \log_2(1/\phi(u))}{\phi(u)}} d\tilde{N}_n^{\tilde{\omega}}(u) \\ &= \lim_{\varepsilon \rightarrow 0} \int_0^\varepsilon \sqrt{\frac{2 \log_2(1/\phi(u))}{\phi(u)}} \lambda(u) p(u) du = 0, \end{aligned}$$

thus proving (5.70). For establishing (5.72), let h be a function as in Lemma 5.3. Then it holds that

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_0^\varepsilon \sqrt{n} \tilde{\gamma}_n^{\tilde{\omega}}(u) d\tilde{N}_n^{\tilde{\omega}}(u) \\ &= \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_0^\varepsilon \frac{\sqrt{n} \tilde{\gamma}_n^{\tilde{\omega}}(u)}{uh(u)} uh(u) d\tilde{N}_n^{\tilde{\omega}}(u) \\ (5.73) \quad & \leq \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \left(\sup_{0 < u < \varepsilon} \sqrt{n} uh(u) \tilde{\gamma}_n^{\tilde{\omega}}(u) \right) \int_0^\varepsilon \frac{1}{uh(u)} d\tilde{N}_n^{\tilde{\omega}}(u) \\ & \leq \sup_{0 < t < \varepsilon} (th(t) \tilde{\beta}_n^{\tilde{\omega}}(0, t)) \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \int_0^\varepsilon \frac{d\tilde{N}_n^{\tilde{\omega}}(u)}{uh(u)}. \end{aligned}$$

Since $1/(uh(u))$ is integrable, it follows by Corollary 5.3 that (5.73) equals 0 a.s. This proves (5.72), and (5.71) now follows in the same way without our having to invoke the law of large numbers.

We have now proved (5.60). Theorem 2.1 entails an application of the δ -method to show that

$$\begin{aligned} & \sqrt{n} \left(\prod_{]0, t]} \left(1 - \frac{dN_n(u)}{Y_n(u) + \gamma_n(u)} \right) - \prod_{]0, t]} (1 - \lambda(u) du) \right) \\ & \Rightarrow - \prod_{]0, t]} (1 - \lambda(u) du) (W(\phi(t)) - K(\phi(t))) \\ & = -\bar{F}(t) (W(\phi(t)) - K(\phi(t))). \quad \square \end{aligned}$$

PROOF OF THEOREM 2.3. We now describe how the proof of Theorem 2.1 may be modified to prove Theorem 2.3. There is a slight complication in that the natural scaling for the counting process and risk set in group $i = 1, 2$ is n_i and not $n = n_1 + n_2$ as used in the nonparametric maximum likelihood estimators (5.25) and (5.26). By defining, however, $N^{(i)}(t) = (n/n_i)N_n^{(i)}(t)$ and $Y^{(i)}(t) = (n/n_i)Y_n^{(i)}(t)$ (and suppressing the dependence on n for convenience), we may write (5.24) as

$$\begin{aligned} (5.74) \quad \pi_n^{(1)}(\beta, s, t) & \equiv \prod_{]s, t]} \left(1 - \frac{dN^{(1)}(u)}{Y^{(1)}(u) + (n/n_1)\beta} \right) \\ & = \prod_{]s, t]} \left(1 - \frac{dN^{(2)}(u)}{Y^{(2)}(u) - (n/n_2)\beta} \right) \\ & \equiv \pi_n^{(2)}(-\beta, s, t). \end{aligned}$$

By the mean value theorem [in the form $f(x) - f(x_0) - (g(-x) - g(-x_0)) = (f'(\xi) + g'(-\xi))(x - x_0)$, $\xi \in]x_0, x[$], we get as in (5.45) that

$$\begin{aligned} (5.75) \quad & \sqrt{n} (\pi_n^{(1)}(0, s, t) - \pi_n^{(1)}(\beta_n(s, t), s, t)) \\ & - \sqrt{n} (\pi_n^{(2)}(0, s, t) - \pi_n^{(2)}(-\beta_n(s, t), s, t)) \\ & = \sqrt{n} (-\beta_n(s, t)) \left(\frac{\partial}{\partial \beta} \pi_n^{(1)}(\beta_n^*(s, t), s, t) \right. \\ & \quad \left. + \frac{\partial}{\partial \beta} \pi_n^{(2)}(-\beta_n^*(s, t), s, t) \right), \end{aligned}$$

where $\beta_n^*(s, t) \in]0, \beta_n(s, t)[$. On the other hand, we have by (5.74) that the terms involving $\beta_n(s, t)$ cancel out, so that it equals

$$\begin{aligned} (5.76) \quad & \sqrt{n} (\pi_n^{(1)}(0, s, t) - \pi_n^{(2)}(0, s, t)) \\ & = \sqrt{\frac{n}{n_1}} \sqrt{n_1} \left(\prod_{]s, t]} \left(1 - \frac{dN^{(1)}(u)}{Y^{(1)}(u)} \right) - \prod_{]s, t]} \left(1 - \frac{dEN^{(1)}(u)}{EY^{(1)}(u)} \right) \right) \\ & \quad - \sqrt{\frac{n}{n_2}} \sqrt{n_2} \left(\prod_{]s, t]} \left(1 - \frac{dN^{(2)}(u)}{Y^{(2)}(u)} \right) - \prod_{]s, t]} \left(1 - \frac{dEN^{(2)}(u)}{EY^{(2)}(u)} \right) \right). \end{aligned}$$

Let $(Z_{N_i}(t), Z_{Y_i}(t))$ denote the two-dimensional Gaussian process which is the limit in distribution of $n_i^{1/2}(N^{(i)}(t) - E^{(i)}(t), Y^{(i)}(t) - EY^{(i)}(t))$. We then ex-

pect by the δ -method as in Theorem 5.1 that (5.76) converges in distribution to

$$(5.77) \quad -c_1^{-1/2} \frac{\bar{F}(t)}{\bar{F}(s)} \int_s^t \left(\frac{dZ_{N_1}(u)}{p_1(u)} - \frac{Z_{Y_1}(u)\lambda(u) du}{p_1(u)} \right) + c_2^{-1/2} \frac{\bar{F}(t)}{\bar{F}(s)} \int_s^t \left(\frac{dZ_{N_2}(u)}{p_2(u)} - \frac{Z_{Y_2}(u)\lambda(u) du}{p_2(u)} \right).$$

Returning to the derivative in (5.75), we get using (5.44) that

$$\begin{aligned} & \frac{\partial}{\partial \beta} \pi_n^{(1)}(\beta, s, t) + \frac{\partial}{\partial \beta} \pi_n^{(2)}(-\beta, s, t) \\ &= \frac{n}{n_1} \pi_n^{(1)}(\beta, s, t) \int_s^t \frac{dN^{(1)}(u)}{(Y^{(1)}(u) + (n/n_1)\beta)(Y^{(1)}(u) + (n/n_1)\beta - n_1^{-1})} \\ & \quad + \frac{n}{n_2} \pi_n^{(2)}(-\beta, s, t) \\ & \quad \times \int_s^t \frac{dN^{(2)}(u)}{(Y^{(2)}(u) - (n/n_2)\beta)(Y^{(2)}(u) - (n/n_2)\beta - n_2^{-1})}. \end{aligned}$$

Hence we expect that if $\beta_n(s, t)$ is, for large n , close to 0, then the right-hand side of (5.75) is asymptotically equivalent to

$$\begin{aligned} & -\sqrt{n} \beta_n(s, t) \left(\frac{1}{c_1} \frac{\bar{F}(t)}{\bar{F}(s)} \int_s^t \frac{dEN^{(1)}(u)}{(EY^{(1)}(u))^2} + \frac{1}{c_2} \frac{\bar{F}(t)}{\bar{F}(s)} \int_s^t \frac{dEN^{(2)}(u)}{(EY^{(2)}(u))^2} \right) \\ & \equiv \frac{\bar{F}(t)}{\bar{F}(s)} (\phi_1(t) - \phi_1(s) + \phi_2(t) - \phi_2(s)), \end{aligned}$$

where we recall $\phi_i(t) = \int_0^t \lambda(u)/(c_i p_i(u))$. By combining (5.75), (5.77) and (5.78), we find that the terms involving $\bar{F}(t)$ cancel out, and we get that

$$(5.78) \quad \left(\sum_{i=1,2} \phi_i(t) - \phi_i(s) \right) \sqrt{n} (\beta_n(s, t)) \Rightarrow \int_s^t \frac{dZ_{N_1}(u) - Z_{Y_1}(u)\lambda(u) du}{p_1(u)} - \int_s^t \frac{dZ_{N_2}(u) - Z_{Y_2}(u)\lambda(u) du}{p_2(u)}.$$

It is natural to write

$$c_i^{-1/2} \int_0^t \frac{dZ_{N_i}(u) - Z_{Y_i}(u)\lambda_i(u) du}{p_i(u)} \equiv c_i^{-1/2} W_i(c_i \phi_i(t))$$

for W_1, W_2 iid Brownian motions on $[0, \tau]$. Let $\phi(t) = \phi_1(t) + \phi_2(t)$. Then we can argue similarly to Corollary 5.4 that (5.78) implies

$$(5.79) \quad \sqrt{n} (t - s) (\beta_n(\phi^{-1}(s), \phi^{-1}(t))) \Rightarrow V(t),$$

where $V(t) \equiv c_1^{-1/2}W_1(c_1\phi_1(\phi^{-1}(t))) - c_2^{-1/2}W_2(c_2\phi_2(\phi^{-1}(t)))$. Next, we invoke the a.s. representation theorem and argue that

$$\begin{aligned}
 \sqrt{n} \gamma_n(t) &\equiv \sqrt{n} \left(\min_{s \leq t} \max_{u \geq t} \beta_n(s, u) \right) \vee 0 \\
 (5.80) \quad &= \sqrt{n} \left(\min_{s \leq \phi.(t)} \max_{u \geq \phi.(t)} \beta_n(\phi^{-1}(s), \phi^{-1}(u)) \right) \vee 0 \\
 &\rightarrow \dot{K}(\phi.(t)),
 \end{aligned}$$

where $\dot{K}(\phi.(t))$ denotes the derivative (left or right) at $u = \phi.(t)$ of the least nondecreasing concave majorant of $V(u)$ on $[0, \tau]$ evaluated at $u = \phi.(t)$. The convergence in (5.80) can, as in Lemma 5.3, be assumed pointwise in every point where $K(\phi.(t))$ is defined.

Notice now that the NPMLEs of $\bar{F}^{(i)}(t)$ may be written as

$$\tilde{F}_n^{(i)}(t) = \prod_{]0, t]} \left(1 - \frac{dN^{(i)}(u)}{Y^{(i)}(u) + (n/n_i)(-1)^{i+1}\gamma_n(u)} \right).$$

Since, furthermore,

$$\begin{aligned}
 \prod_{]0, t]} \left(1 - \frac{dEN^{(i)}(u)}{EY^{(i)}(u)} \right) &= \prod_{]0, t]} \left(1 - \frac{p_i(u)\lambda(u) du}{p_i(u)} \right) \\
 &= \exp\left(-\int_0^t \lambda(u) du\right) = \bar{F}^{(i)}(t),
 \end{aligned}$$

an application of the δ -method together with (5.80), still along an a.s. subsequence [noting that $n_i^{1/2}(n/n_i)\gamma_i(t) = n^{1/2}(n/n_i)^{1/2}\gamma_i(t) \rightarrow c_i^{-1/2}\dot{K}(\phi.(t))$], shows that

$$\begin{aligned}
 &\sqrt{n} (\bar{F}_n^{(i)}(t) - \bar{F}^{(i)}(t)) \\
 &= \sqrt{\frac{n}{n_i}} \sqrt{n_i} \left(\prod_{]0, t]} \left(1 - \frac{dN^{(i)}(u)}{Y^{(i)}(u) + (n/n_i)(-1)^{i+1}\gamma_n(u)} \right) \right. \\
 &\quad \left. - \prod_{]0, t]} \left(1 - \frac{dEN^{(i)}(u)}{EY^{(i)}(u)} \right) \right) \\
 &\rightarrow -c_i^{-1/2}\bar{F}^{(i)}(t) \\
 &\quad \times \int_0^t \left(\frac{dZ_{N_i}(u)}{EY^{(i)}(u)} - \frac{Z_{Y_i}(u) + (-1)^{i+1}c_i^{-1/2}\dot{K}(\phi.(u)) dEN^{(i)}(u)}{EY^{(i)}(u)^2} \right) \\
 &= -c_i^{1/2}\bar{F}^{(i)}(t) \left(\int_0^t \frac{dZ_{N_i}(u) - Z_{Y_i}(u)\lambda(u) du}{p_i(u)} \right. \\
 &\quad \left. - (-1)^{i+1}\bar{F}^{(i)}(t)c_i^{-1} \int_0^t \dot{K}(\phi.(u)) \frac{\lambda(u)}{p_i(u)} du \right).
 \end{aligned}$$

Remembering the definition of $W_i(c_i\phi_i(t))$, the limit in (5.81) may now be written as

$$-\bar{F}^{(i)}(t) \left(c_i^{-1/2} W_i(c_i\phi_i(t)) + (-1)^i \int_0^t \dot{K}(\phi(u)) \dot{\phi}_i(u) du \right).$$

By noting that $W_i^*(t) \equiv c_i^{-1/2} W_i(c_i, t)$ are themselves independent Brownian motions and that $V(t) = W_1^*(\phi_1 \circ \phi^{-1}(t)) - W_2^*(\phi_2 \circ \phi^{-1}(t))$, the desired result (Theorem 2.3) follows. What is needed in order to make the above line of reasoning rigorous is first of all that Corollaries 5.1, 5.2 and 5.3 all hold for $\beta_n(s, t)$. That is, it must hold that

$$(5.81) \quad \sup_{t-s > \delta} |\beta_n(s, t)| \rightarrow_p 0 \quad \text{for all } \delta > 0,$$

$$(5.82) \quad \limsup_{n \rightarrow \infty} \sup_{0 < t-s < \delta} |\beta_n(s, t)| \rightarrow_p 0, \quad \delta \rightarrow 0,$$

and

$$(5.83) \quad \sup_{0 < t < 1} \sqrt{n} th(t) |\beta_n(0, t)| = O_p(1).$$

For showing (5.81)–(5.83) we define $\rho_n^{(i)}(s, t)$ as the solutions to the equations

$$(5.84) \quad \prod_{[s, t]} \left(1 - \frac{(n_i/n) dN^{(i)}(u)}{(n_i/n) Y^{(i)}(u) + \rho} \right) = \frac{\bar{F}(t)}{\bar{F}(s)}.$$

Since now $\beta_n(s, t)$ satisfies by definition the equation

$$\prod_{[s, t]} \left(1 - \frac{(n_1/n) dN^{(1)}(u)}{(n_1/n) Y^{(1)}(u) + \beta} \right) = \prod_{[s, t]} \left(1 - \frac{(n_2/n) dN^{(2)}(u)}{(n_2/n) Y^{(2)}(u) - \beta} \right),$$

it follows by symmetry that

$$\min(-\rho_n^{(2)}(s, t), \rho_n^{(1)}(s, t)) \leq \beta_n(s, t) \leq \max(-\rho_n^{(2)}(s, t), \rho_n^{(1)}(s, t))$$

or

$$|\beta_n(s, t)| \leq |\rho_n^{(1)}(s, t)| + |\rho_n^{(2)}(s, t)|.$$

By (5.84), $(n/n_i)\rho_n^{(i)}(s, t)$ are the adjustment used in one-sample problems of the type considered in Theorem 2.1 which consist of estimating $\bar{F}^{(i)}(t)$ subject to the restrictions $\bar{F}^{(i)}(t) \geq \bar{F}(t)$. We can now argue that Corollaries 5.1, 5.2 and 5.3 are all satisfied for $(n/n_i)\rho_n^{(i)}(s, t)$. There now remains one point where the proof of Theorem 2.3 differs from that of Theorem 2.1. This has to do with verifying Lemma 5.5 for the quantity H_n when treating the lower survival curve estimator $\bar{F}_n^{(2)}(t)$. In verifying Lemma 5.5 for H_n , defined in (5.66), in the proof of Theorem 2.1, we used that the nonnegative γ_n was added in the denominator, so that uniform integrability of $n^{1/2}\gamma_n$ sufficed. However, in case of $\bar{F}_n^{(2)}$ the corresponding definition is

$$H_n(t) \equiv - \frac{\sqrt{n} (Y^{(2)}(t) - p_2(t) - \gamma_n(t))}{p(t)(Y^{(2)}(t) - \gamma_n(t))};$$

that is, $\gamma_n(t)$ is subtracted in the denominator which may subsequently blow up the fraction.

It would suffice for carrying through the proof of Theorem 2.1 that $\gamma_n(t) \rightarrow 0$ a.s. uniformly in t . By choosing $h(t) = t^{-1/3}$, it follows from (5.83) that $\gamma_n(t) \leq^{-1/2} t^{-2/3} O_p(1)$. Let a_n denote the last time the curves $\bar{F}^{(1)}(t)$ and $\bar{F}^{(2)}(t)$ touch before $n^{-1/2-\varepsilon}$ for some $0 < \varepsilon < \frac{1}{4}$. Then it holds that $\gamma_n(t) \rightarrow 0$ a.s. uniformly on $[a_n, 1]$. Hence we can show, along the a.s. representation, that

$$(5.85) \quad \sqrt{n} \left(\int_{a_n}^t \frac{dN^{(2)}(u)}{Y^{(2)}(u) - \gamma_n(u)} - \int_{a_n}^t \lambda(u) du \right)$$

converges uniformly in $t > a_n$ to the given limit [which is $c_2^{-1/2} W_2(c_2 \phi_2(t)) + \int_0^t \dot{K}(\phi(u)) \dot{\phi}_2(u) du$]. As to $0 < t < a_n$ the proof will be complete if it can be shown that

$$(5.86) \quad \limsup_{n \rightarrow \infty} \sup_{0 < t < a_n} \sqrt{n} \left| \int_0^t \frac{dN^{(2)}(t)}{Y^{(2)}(t) - \gamma_n(t)} - \int_0^t \lambda(u) du \right| = 0 \quad \text{a.s.}$$

Notice to this end that (5.86) is bounded by

$$\sqrt{n} \int_0^{a_n} \frac{dN^{(2)}(t)}{Y^{(2)}(t) - \gamma_n} + \sqrt{n} \int_0^{n^{-1/2-\varepsilon}} \lambda(u) du$$

and that the right-hand term goes to 0 by the boundedness of $\lambda(u)$. Since $\bar{F}^{(1)}(a_n) = \bar{F}^{(2)}(a_n)$, we now get by the inequality $x \leq -\log(1 - x)$ that

$$(5.87) \quad \begin{aligned} \sqrt{n} \int_0^{a_n} \frac{dN^{(2)}(t)}{Y^{(2)}(t) - \gamma_n(t)} &\leq -\sqrt{n} \log \prod_{]0, a_n]} \left(1 - \frac{dN^{(2)}(t)}{Y^{(2)}(t) - \gamma_n(t)} \right) \\ &= -\sqrt{n} \log \bar{F}_n^{(2)}(a_n) \\ &= -\sqrt{n} \log \bar{F}_n^{(1)}(a_n) \\ &\leq -\log S^{(1)}(a_n). \end{aligned}$$

Here $S^{(1)}(t)$ denotes the Kaplan–Meier estimator from the first sample, and the last inequality follows because $\bar{F}_n^{(1)}$ has received a positive adjustment to the risk set so that $\bar{F}^{(1)}(a_n) > S^{(1)}(a_n)$. One can now invoke an almost sure “nearly linear” bound for the Kaplan–Meier estimator to show that for any $\delta > 0$ there exists a constant $C > 0$ so that

$$\liminf_{n \rightarrow \infty} \inf_{0 < t < 1} (S^{(1)}(t) - (1 - Ct^{1-\delta})) = 0 \quad \text{a.s.}$$

For an explanation of a.s. linear bounds (for the empirical process, but easily gotten also for the Kaplan–Meier estimator), see for example Shorack and Wellner (1986), page 405. It now follows that (5.87) can eventually be bounded by $-\sqrt{n} \log(1 - Ca_n^{1-\delta}) \leq -\sqrt{n} \log(1 - Cn^{-(1/2+\varepsilon)(1-\delta)})$. Now choose δ small enough and positive that $(\frac{1}{2} + \varepsilon)(1 - \delta) > \frac{1}{2}$; then it follows by the inequality $-\log(1 - x) < 2x$ if $0 < x < 0.797$ that

$$\limsup_{n \rightarrow \infty} \sqrt{n} \log(1 - Cn^{(1/2+\varepsilon)(1-\delta)}) \leq \limsup_{n \rightarrow \infty} \sqrt{n} 2Cn^{1/2} n^{-(1/2+\varepsilon)(1-\delta)} = 0.$$

This proves (5.86), and together with (5.85) this completes the proof of Theorem 2.3. \square

Acknowledgments. We thank our colleagues Dick Dykstra, Bob Hogg, Tim Robertson and Jon Wellner for their advice.

REFERENCES

- BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference Under Order Restrictions*. Wiley, New York.
- BARLOW, R. E. and BRUNK, H. D. (1972). The isotonic regression problem and its dual. *J. Amer. Statist. Assoc.* **67** 140–147.
- BREIMAN, L. B. (1968). *Probability*. Addison-Wesley, Reading, MA.
- BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2** 437–453.
- BRUNK, H. D., FRANCK, W. E., HANSON, D. L. and HOGG, R. V. (1966). Maximum likelihood estimation of the distribution function of two stochastically ordered random variables. *J. Amer. Statist. Assoc.* **61** 1067–1080.
- DUDLEY, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929.
- DUDLEY, R. M. (1984). A course on empirical processes. *Lecture Notes in Math.* **1097** 2–142. Springer, New York.
- DUDLEY, R. M. (1985). An extended Wichura theorem, definitions of Donsker classes, and weighted empirical distributions. *Lecture Notes in Math.* **1143** 141–178. Springer, New York.
- DUDLEY, R. M. (1987). Some universal Donsker classes of functions. *Ann. Probab.* **15** 1306–1326.
- DYKSTRA, R. L. (1982). Maximum likelihood estimation of the survival function of stochastically ordered random variables. *J. Amer. Statist. Assoc.* **77** 621–628.
- GILL, R. D. (1989). Non- and semiparametric maximum likelihood estimation and the von Mises method. *Scand. J. Statist.* **16** 97–129.
- GILL, R. D. (1995). Nonparametric estimation under censoring and passive registration. Preprint.
- GILL, R. D. and JOHANSEN, S. (1990). A survey of product-integration with a view towards application in survival analysis. *Ann. Statist.* **18** 1501–1556.
- GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes (with discussion). *Ann. Probab.* **12** 929–998.
- GINÉ, E. and ZINN, J. (1986). Lectures on the central limit theorem for empirical processes. *Lecture Notes in Math.* **1221** 50–113. Springer, New York.
- GROENEBOOM, P. (1983). The concave majorant of Brownian motion. *Ann. Probab.* **11** 1016–1027.
- GROENEBOOM, P. and PYKE, R. (1983). Asymptotic normality of statistics based on the convex minorants of empirical distribution functions. *Ann. Probab.* **11** 328–345.
- HOFFMANN-JØRGENSEN, J. (1984). Stochastic process on Polish spaces. Unpublished manuscript.
- KIM, J. and POLLARD, D. (1990). Cube root asymptotics. *Ann. Statist.* **18** 191–219.
- POLLARD, D. (1984). *Convergence in Distribution of Stochastic Processes*. Springer, New York.
- POLLARD, D. (1990). *Empirical Processes: Theory and Applications*. IMS, Hayward, CA.
- PRÆSTGAARD, J. T. and HUANG, J. (1995). Asymptotic theory for nonparametric estimation of survival curves under order restrictions. Technical report, Dept. Statistics and Actuarial Science, Univ. Iowa.
- PYKE, R. and SHORACK, G. R. (1968). Weak convergence of a two sample empirical process and a new approach to Chernoff–Savage theorems. *Ann. Math. Statist.* **39** 755–771.
- ROBERTSON, T., WRIGHT, F. T. and DYKSTRA, R. L. (1988). *Order Restricted Statistical Inference*. Wiley, New York.
- ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.

- SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.
- VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.
- WELLNER, J. A. (1989). Comment on R. D. Gill. Non- and semiparametric maximum likelihood estimation and the von Mises method. I. *Scand. J. Statist.* **16** 97–129.

COLLEGE OF AMERICAN PATHOLOGISTS
325 WAUKEGAN ROAD
NORTHFIELD, ILLINOIS 60093

DEPARTMENT OF STATISTICS AND
ACTUARIAL SCIENCE
14 MACLEAN HALL
UNIVERSITY OF IOWA
IOWA CITY, IOWA 52242
E-MAIL: jian@stat.uiowa.edu