Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

Asymptotics for a weighted least squares estimator of the disease onset distribution function for a survival-sacrifice model

Antonio Eduardo Gomes

Relatório Técnico
RTP-03/2002

Relatório Técnico
Série Pesquisa

Asymptotics for a weighted least squares estimator of the disease onset distribution function for a survival-sacrifice model

Antonio Eduardo Gomes
Departamento de Estatística, Universidade Federal de Minas Gerais
aegomes@est.ufmg.br

Abstract

In carcinogenicity experiments with animals where the tumour is not palpable it is common to observe only the time of death of the animal, the cause of death (the tumour or another independent cause, as sacrifice) and whether the tumour was present at the time of death. These last two indicator variables are evaluated after an autopsy. A weighted least squares estimator for the distribution function of the disease onset was proposed by van der Laan et al. (1997). Some asymptotic properties of their estimator are established. A minimax lower bound for the estimation of the disease onset distribution is obtained, as well as the local asymptotic distribution for their estimator.

1 Introduction

Suppose that in an experiment for the study of onset and mortality from undetectable moderately lethal incurable diseases (occult tumours, e.g.) we observe the time of death, whether the disease of interest was present at death, and if present, whether the disease was a probable cause of death. Defining the nonnegative variables T_1 (time of disease onset), T_2 (time of death from the disease) and C (time of death from an unrelated cause), we observe, for the ith individual, $(Y_i, \Delta_{1,i}, \Delta_{2,i})$, where $Y_i = C_i \wedge T_{2,i} = \min\{C_i, T_{2,i}\}$, $\Delta_{1,i} = I(T_{1,i} \leq C_i)$, $\Delta_{2,i} = I(T_{2,i} \leq C_i)$, and $I(.)$ is the indicator function. $T_{1,i}$ and $T_{2,i}$ have an unidentifiable joint distribution function F such that $P(T_{1,i} \leq T_{2,i}) = 1$, C_i has distribution function G and is independent of $(T_{1,i}, T_{2,i})$. Current status data can be seen as a particular case of the survival-sacrifice model above when the disease is nonlethal, i.e., $\Delta_{2,i} = 0$ $(i = 1, \dots, n)$. In this case, $Y_i = C_i$, and $\hat{F}_2 \equiv 0$ for any estimator $\hat{F}_2$ of F_2 (the marginal distribution function of T_2). Right-censored data are a special case of the survival-sacrifice model above when a lethal disease is always present at the moment of death, i.e., $\Delta_{1,i} = 1$ $(i = 1, \ldots n)$. In this case, $\hat{F}_1 \equiv 1$ for any estimator $\hat{F}_1$ of F_1 (the marginal distribution function of T_1).

An example of a real data set studied by Dinse & Lagakos (1982) and Turnbull & Mitchell (1984) is presented in Table 1 and represents the ages at death (in days) of 109 female RFM mice. The disease of interest is reticulum cell sarcoma (RCS). These mice formed the control group in a survival experiment to study the effects of prepubertal ovariectomy in mice given 300 R of X-rays.

Table 1: *Ages at death (in days) in unexposed female RFM mice.*

$\Delta_1 = 1, \Delta_2 = 1$	406,461,482,508,553,555,562,564,570,574,585,588,593,624,626, 629,647,658,666,675,679,688,690,691,692,698,699,701,702,703, 707,717,724,736,748,754,759,770,772,776,776,785,793,800,809, 811,823,829,849,853,866,883,884,888,889
$\Delta_1 = 1, \Delta_2 = 0$	356,381,545,615,708,750,789,838,841,875
$\Delta_1 = 0, \Delta_2 = 0$	192,234,243,300,303,330,339,345,351,361,368,419,430,430,464, 488,494,496,517,552,554,555,563,583,629,638,642,656,668,669, 671,694,714,730,731,732,756,756,782,793,805,821,828,853

The parameter space for the survival-sacrifice model can be taken to be

$$\Theta = \{(F_1, F_2) : F_1 \text{ and } F_2 \text{ are distribution functions with } F_1 <_s F_2\} ,$$

where $F_1 <_s F_2$ means that $F_1(x) \geq F_2(x)$ for every $x \in \mathbb{R}$ and $F_1(x) > F_2(x)$ for some $x \in \mathbb{R}$. The loglikelihood function for this model is

$$\begin{aligned} \mathcal{L}(F_1, F_2) &= \sum_{i=1}^{n} \{(1 - \Delta_{1,i})(1 - \Delta_{2,i}) \log (1 - F_1(Y_i)) \\ &\quad + \Delta_{1,i}(1 - \Delta_{2,i}) \log (F_1(Y_i) - F_2(Y_i)) \\ &\quad + (\Delta_{1,i}\Delta_{2,i}) \log f_2(Y_i)\} + K(g, G) \end{aligned}$$

where $f_2(y) = F_2(y) - F_2(y-)$ and $K(g, G)$ is a term involving only the distribution function G and the probability density function g of C. We will assume without loss of generality that $Y_1 \leq Y_2 \leq \cdots \leq Y_n$.

Kodell et al. (1982) also studied the nonparametric estimation of $S_1 = 1 - F_1$ and $S_2 = 1 - F_2$, but their work is restricted to the case where $R(t) = S_1(t)/S_2(t)$ is nonincreasing, an assumption that may not be reasonable, for example, for progressive diseases whose incidence is concentrated in the early or middle part of the life span.

Turnbull & Mitchell (1984) proposed an EM algorithm for the joint estimation of F_1 and F_2 which converges to the nonparametric maximum likelihood estimator of (F_1, F_2) provided the support of the initial estimator contains the support of the maximum likelihood estimator.

Another possible way of estimating F_1 is by plugging in the Kaplan-Meier estimator $\hat{F}_{2,n}$ of F_2 and calculating the nonparametric maximum pseudolikelihood estimator of F_1.

A weighted least squares estimator for F_1 making $F_2 = \hat{F}_{2,n}$ was proposed by van der Laan et al. (1997) . Their estimator is described in Section 2. The proof of its consistency is presented in Gomes (2001). Results about the rate of convergence and the local limit distribution of their estimator are established in Sections 3 and 4, respectively.

2 The Weighted Least Squares Estimator of F_1

One possibility for the estimation of F_1 is to calculate a weighted least squares estimator as suggested by van der Laan et al. (1997). Making $S_1 = 1 - F_1$ and $S_2 = 1 - F_2$, in terms of populations, $R(c) = S_1(c)/S_2(c)$ is the proportion of subjects alive at time c who are disease-free (i.e., $1 - R(c)$ is the prevalence function at time c). It can be written as

$$\begin{aligned} R(c) &= \frac{S_1(c)}{S_2(c)} = \frac{1 - F_1(c)}{1 - F_2(c)} = \frac{\mathrm{pr}\,(T_1 > c)}{\mathrm{pr}\,(T_2 > c)} \\ &= \frac{\mathrm{pr}\,(T_1 > c, T_2 > c)}{\mathrm{pr}\,(T_2 > c)} = \mathrm{pr}\,(T_1 > C \mid C = c, T_2 > C) \\ &= E\,\{I\,(T_1 > C) \mid C = c, T_2 > C\} = E(1 - \Delta_1 \mid C = c, T_2 > C). \end{aligned}$$

So, it is possible to rewrite

$$\begin{aligned} S_1(c) &= R(c)S_2(c) = S_2(c)E(1 - \Delta_1 \mid C = c, T_2 > C) \\ &= E\{S_2(C)\,(1 - \Delta_1) \mid C = c, T_2 > C\}. \end{aligned}$$

Estimating S_1 can be viewed, then, as a regression of $S_2(C)\,(1 - \Delta_1)$ on the observed C_i's under the constraint of monotonicity. If we substitute S_2 by its Kaplan-Meier estimator $\hat{S}_{2,n}$, we automatically have an estimator for S_1 minimising

$$\frac{1}{n}\sum_{i=1}^{n}\{(1 - \Delta_{1,i})\hat{S}_{2,n}(Y_i) - S_1(Y_i)\}^2(1 - \Delta_{2,i})$$

under the constraint that S_1 is nonincreasing. This minimisation problem can be solved by using results from the theory of isotonic regression (see Barlow et al., 1972) and its solution is given by

$$\hat{S}_1(Y_m) = \min_{l \le m}\max_{k \ge m}\frac{\sum_{j=l}^{k}\hat{S}_{2,n}(Y_j)(1 - \Delta_{1,j})(1 - \Delta_{2,j})}{\sum_{j=l}^{k}(1 - \Delta_{2,j})},$$

$(m = 1, \ldots, n)$.
However, $\mathrm{var}\{S_2(C)\,(1-\Delta_1) \mid C = c, T_2 > C\}$ is not constant. In fact,

$$
\begin{aligned}
& \mathrm{var}\{S_2(C)\,(1-\Delta_1) \mid C = c, T_2 > C\} \\
= \; & S_2^2(c)\mathrm{var}\{(1-\Delta_1) \mid C = c, T_2 > C\} \\
= \; & S_2^2(c)\mathrm{pr}\,(T_1 > C \mid C = c, T_2 > C)\,\{1 - \mathrm{pr}\,(T_1 > C \mid C = c, T_2 > C)\} \\
= \; & S_2^2(c)E(1-\Delta_1 \mid C = c, T_2 > C)\{1 - E(1-\Delta_1 \mid C = c, T_2 > C)\} \\
= \; & S_2^2(c)R(c)\{1 - R(c)\}.
\end{aligned}
$$

We may, then, use a weighted least squares estimator with weights w_i inversely proportional to the variance $S_2^2(C_i)R(C_i)\{1-R(C_i)\}$ $(i = 1, \ldots, n)$. This expression for the variance involves the unknown value $S_1(C_i)$ that we want to estimate, suggesting the use of an iterative procedure. In each step, the estimate would be given by

$$
\hat{S}_1(Y_m) = \min_{l \le m} \max_{k \ge m} \frac{\sum_{j=l}^{k} \hat{S}_{2,n}(Y_j)(1-\Delta_{1,j})(1-\Delta_{2,j})/[\hat{S}_{2,n}^2(Y_j)R(Y_j)\{1-R(Y_j)\}]}{\sum_{j=l}^{k}(1-\Delta_{2,j})/[\hat{S}_{2,n}^2(Y_j)R(Y_j)\{1-R(Y_j)\}]}
$$

$(m = 1, \ldots, n)$. If we use $w_j = (1-\Delta_{2,j})/\hat{S}_{2,n}^2(Y_j)$ instead, we have an estimator with a closed form, as suggested by van der Laan et al. (1997).

The estimators expressed as the solution for an isotonic regression problem have a geometric interpretation. Consider the least concave majorant determined by the points $(0,0), (W_1, V_1), \ldots, (W_n, V_n)$, where $W_j = \sum_{i=1}^{j} w_i$ and

$$
V_j = \sum_{i=1}^{j} w_i(1-\Delta_{1,i})\hat{S}_{2,n}(Y_i) = \sum_{i=1}^{j} \frac{(1-\Delta_{1,i})(1-\Delta_{2,i})}{\hat{S}_{2,n}(Y_i)} = \sum_{i=1}^{j} \frac{(1-\Delta_{1,i})}{\hat{S}_{2,n}(Y_i)}.
$$

$\hat{S}_1(t)$ is the slope of the least concave majorant at W_j if $t \in (Y_{j-1}, Y_j]$. Barlow et al. (1972) and Robertson et al. (1988) give a detailed description of these equivalent representations.

So, we can write

$$
\hat{S}_{1,n}(Y_m) = \min_{l \le m} \max_{k \ge m} \frac{\sum_{j=l}^{k}(1-\Delta_{1,j})/\hat{S}_{2,n}(Y_j)}{\sum_{j=l}^{k}(1-\Delta_{2,j})/\hat{S}_{2,n}^2(Y_j)}, \tag{2.1}
$$

$(m = 1, \ldots, n)$. It is easy to see that (2.1) reduces to the expression of the nonparametric maximum likelihood estimator of $F \equiv F_1$ for current status data (see Groeneboom & Wellner, 1992) when $\Delta_{2,i} = 0$, $(i = 1, \ldots, n)$, since in this case $\hat{S}_{2,n} \equiv 1$.

The weighted least squares estimator of F_1 proposed by van der Laan et al. (1997) and the Kaplan-Meier estimator of F_2 do not coincide with the nonparametric

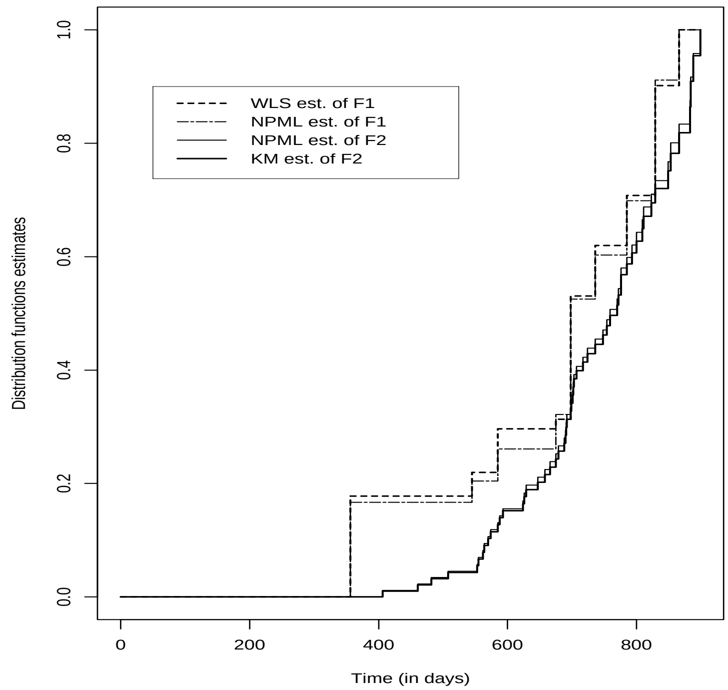

Figure 1: **Weighted least squares estimate of F_1, Kaplan-Meier estimate of F_2, and nonparametric maximum likelihood estimates of F_1 and F_2.**

maximum likelihood estimators of F_1 and F_2, respectively. Figure 1 shows those estimators for the data in Table 1. The smoother picture for the estimates of F_2 is a consequence of the $n^{-1/3}$ rate of convergence for the estimation of F_1 compared to the $n^{-1/2}$ rate for the estimation of F_2.

3 Minimax Lower Bound

We determine here a minimax lower bound for the estimation of $F_1(t_0)$. When n grows, the minimax risk should decrease to zero. The rate δ_n of this convergence is the best rate of convergence an estimator can have for the estimation problem posed.

Let T be a functional and q a probability density in a class $\mathcal{G}$ with respect to a σ-finite measure μ on the measurable space $(\Omega, \mathcal{A})$. Let Tq denote a real-valued parameter and $\{T_n\}, n \geq 1$, be a sequence of estimators of Tq based on samples of size $n, X_1, \ldots, X_n$, generated by q.

$E_{n,q}\{\ell(\mid T_n - Tq \mid)\}$ is the *risk* of the estimator T_n in estimating Tq when the loss function is $\ell : [0, \infty) \longrightarrow \mathbb{R}$ (ℓ is increasing and convex with $\ell(0) = 0$). $E_{n,q}$

denotes the expectation with respect to the product measure $q^{\otimes n}$ associated with the sample $X_1, \dots, X_n$. For fixed n, the minimax risk

$$\inf_{T_n} \sup_{q \in \mathcal{G}} E_{n,q} \{\ell(| T_n - Tq |)\}$$

is a way to measure how hard the estimation problem is.

Lemma 3.1 below is quite helpful in deriving asymptotic lower bounds for minimax risks (see Groeneboom, 1996, chapter 4, for the proof). It will be used to prove Theorem 3.1.

LEMMA 3.1. *Let $\mathcal{G}$ be a set of probability densities on a measurable space $(\Omega, \mathcal{A})$ with respect to a σ-finite measure μ, and let T be a real-valued functional on $\mathcal{G}$. Moreover, let $\ell : [0, \infty) \longrightarrow \mathbb{R}$ be an increasing convex loss function, with $\ell(0) = 0$. Then, for any $q_1, q_2 \in \mathcal{G}$ such that the Hellinger distance $H(q_1, q_2) < 1$,*

$$\begin{aligned} &\inf_{T_n} \max \left[E_{n,q_1} \{\ell(|T_n - Tq_1|)\}, E_{n,q_2} \{\ell(|T_n - Tq_2|)\}\right] \\ \geq \; &\ell[\tfrac{1}{4} |Tq_1 - Tq_2| \left\{1 - H^2(q_1, q_2)\right\}^{2n}]. \end{aligned}$$

In our case, let $X_i = (Y_i, \Delta_{1,i}, \Delta_{2,i})$, and

$$\begin{aligned} q_0(y, \Delta_1, \Delta_2) = \; & [g(y)\{1 - F_1(y)\}]^{(1-\Delta_1)(1-\Delta_2)} \\ & \times [g(y)\{F_1(y) - F_2(y)\}]^{\Delta_1(1-\Delta_2)} \\ & \times [f_2(y)\{1 - G(y)\}]^{\Delta_1\Delta_2}, \end{aligned}$$

$\mu = \lambda \times m$, where λ is the Lebesgue measure and m is the counting measure on the set $\{(0,0), (0,1), (1,1)\}$, $Tq_0 = F_1(t_0)$, and q_n is equal to the density corresponding to the perturbation

$$F_{1,n}(x) = \begin{cases} F_1(x) & \text{if} \quad x < t_0 - n^{-1/3}t \\ F_1(t_0 - n^{-1/3}t) & \text{if} \quad x \in [t_0 - n^{-1/3}t, \; t_0) \\ F_1(t_0 + n^{-1/3}t) & \text{if} \quad x \in [t_0, \; t_0 + n^{-1/3}t) \\ F_1(x) & \text{if} \quad x \geq t_0 + n^{-1/3}t \end{cases} \tag{3.2}$$

for a suitably chosen $t > 0$.

Using the perturbation (3.2) it is seen in the proof of Theorem 3.1 that

$$H^2(q_n, q_0) \sim n^{-1} g(t_0) f_1^2(t_0) t^3 S_2(t_0) / [S_1(t_0)\{S_2(t_0) - S_1(t_0)\}].$$

So, as pointed out by Groeneboom (1996) for current status data, we could say that the Hellinger distance of order $n^{-1/2}$ between q_n and q_0 corresponds to a distance of order $n^{-1/3}$ between $Tq_n = F_{1,n}(t_0)$ and $Tq_0 = F_1(t_0)$.

The perturbation (3.2) is the worst possible. When maximising in t, we are taking the worst possible constant.

THEOREM 3.1.

$$
\begin{aligned}
& n^{1/3} \inf_{\hat{F}_{1,n}} \max\{E_{n,q_0}(\left|\hat{F}_{1,n}(t_0) - F_1(t_0)\right|), E_{n,q_n}(\left|\hat{F}_{1,n}(t_0) - F_{1,n}(t_0)\right|)\} \\
\geq \quad & \frac{1}{4} n^{1/3} \left|F_{1,n}(t_0) - F_1(t_0)\right| 1 - H^2(q_n, q_0)\}^{2n} \\
\longrightarrow \quad & \frac{1}{4} f_1(t_0) t \exp\left[-\frac{2S_2(t_0)g(t_0)f_1^2(t_0)t^3}{S_1(t_0)\{S_2(t_0) - S_1(t_0)\}}\right]
\end{aligned}
$$

and the maximum value of the last expression is

$$k[f_1(t_0)S_1(t_0)\{S_2(t_0) - S_1(t_0)\}/\{S_2(t_0)g(t_0)\}]^{1/3} \tag{3.3}$$

where $k = \frac{1}{4}(6e)^{-1/3}$ *does not depend on* f_1*,* F_1 *or* g*.*

Proof. In the Appendix.

Groeneboom (1987) applied Lemma 3.1 to obtain a minimax lower bound of the form

$$c[f(t_0)F(t_0)\{1 - F(t_0)\}/g(t_0)]^{1/3} \tag{3.4}$$

for the problem of estimating F with current status data. We can easily see that up to the constants k and c (3.3) reduces to (3.4) if we make $F_1 \equiv F$, $f_1 \equiv f$ and $S_2 \equiv 1$, which are the changes that reduce the survival-sacrifice model studied here to current status data.

4 Local Limit Distribution

Theorem 4.1 below gives the local asymptotic behavior of the weighted least squares estimator of F_1 proposed by van der Laan et al. (1997).

THEOREM 4.1. *Suppose* C, T_1 *and* T_2 *have continuous distribution functions* G*,* T_1 *and* T_2*, respectively, such that* $P_{F_1} \ll P_G$*. Additionally, let* t_0 *be such that* $0 < F_1(t_0) < 1$*,* $0 < G(t_0) < 1$*, and let* F_1 *and* G *be differentiable at* t_0*, with strictly positive derivatives* $f_1(t_0)$ *and* $g(t_0)$*, respectively. Suppose also that* t_0 *is such that for some* $\delta > 0$ *and some* $M > 0$*,* $S_2(t_0 + M) > \delta$*. Then*

$$n^{1/3} \frac{\hat{S}_{1,n}(t_0) - S_1(t_0)}{\left[\frac{1}{2} f_1(t_0)S_1(t_0)\{S_2(t_0) - S_1(t_0)\}/\{g(t_0)S_2(t_0)\}\right]^{1/3}} \longrightarrow 2Z \tag{4.5}$$

in distribution, where $Z = \arg\max_h \{\mathbb{B}(h) - h^2\}$*, and* $\mathbb{B}$ *is a two-sided standard Brownian Motion starting from 0.*

Proof. In the Appendix.

Groeneboom (1989) studied the distribution of the random variable Z, and Groeneboom & Wellner (2001) calculated its quantiles. That made the construction of confidence intervals for F_1 possible.

As noticed in the introduction, current status data is a particular version of the present problem when we have $\Delta_{2,i} = 0$, $(i = 1, \ldots, n)$. This is equivalent to have $S_2(t_0) \equiv 1$ in the expression above, which would reduce it to the well known result about the limit distribution of the nonparametric maximum likelihood estimator of F_1 when we have current status data (see Groeneboom & Wellner (1992)).

Notice that the expression in the denominator of (4.5) is proportional to that in the minimax lower bound in Theorem 3.1.

5 Appendix

5.1 Proof of Theorem 3.1.

Take $\ell(x) = | x |$. Since q_n and q_0 coincide when $\Delta_1 = \Delta_2 = 1$, we have

$$\begin{aligned} H^2(q_n, q_0) &= \int_{t_0 - n^{-1/3}}^{t_0} g(c)[\{S_1(t_0 - n^{-1/3})\}^{1/2} - \{S_1(c)\}^{1/2}]^2 dc \\ &+ \int_{t_0}^{t_0 + n^{-1/3}} g(c)[\{S_1(t_0 + n^{-1/3})\}^{1/2} - \{S_1(c)\}^{1/2}]^2 dc \\ &+ \int_{t_0 - n^{-1/3}}^{t_0} g(c)[\{S_2(c) - S_1(t_0 - n^{-1/3}t)\}^{1/2} - \{S_2(c) - S_1(c)\}^{1/2}]^2 dc \\ &+ \int_{t_0}^{t_0 + n^{-1/3}} g(c)[\{S_2(c) - S_1(t_0 + n^{-1/3}t)\}^{1/2} - \{S_2(c) - S_1(c)\}^{1/2}]^2 dc \end{aligned}$$

$$
\begin{aligned}
&\cong\ g(t_0)[\{S_1(t_0-n^{-1/3}t)\}^{1/2}-\{S_1(t_0)\}^{1/2}]^2 n^{-1/3}t/2\\
&\quad +g(t_0)[\{S_1(t_0+n^{-1/3}t)\}^{1/2}-\{S_1(t_0)\}^{1/2}]^2 n^{-1/3}t/2\\
&\quad +\ g(t_0)[\{S_2(t_0)-S_1(t_0-n^{-1/3}t)\}^{1/2}-\{S_2(t_0)-S_1(t_0)\}^{1/2}]^2 n^{-1/3}t/2\\
&\quad +\ g(t_0)[\{S_2(t_0)-S_1(t_0+n^{-1/3}t)\}^{1/2}-\{S_2(t_0)-S_1(t_0)\}^{1/2}]^2 n^{-1/3}t/2\\
&\cong\ 2g(t_0)\left[\frac{f_1(t_0)n^{-1/3}t}{2\{S_1(t_0)\}^{1/2}}\right]^2\frac{n^{-1/3}t}{2}\ +\ 2g(t_0)\left[\frac{f_1(t_0)n^{-1/3}t}{2\{S_2(t_0)-S_1(t_0)\}^{1/2}}\right]^2\frac{n^{-1/3}t}{2}\\
&=\ g(t_0)f_1^2(t_0)n^{-1}t^3/S_1(t_0)\ +\ g(t_0)f_1^2(t_0)n^{-1}t^3/\{(S_2(t_0)-S_1(t_0)\}\\
&=\ g(t_0)f_1^2(t_0)n^{-1}t^3S_2(t_0)/[S_1(t_0)\{S_2(t_0)-S_1(t_0)\}]
\end{aligned}
$$

Then we have

$$
\begin{aligned}
&n^{1/3}\inf_{\hat{F}_{1,n}}\max\{E_{n,q_0}(\left|\hat{F}_{1,n}(t_0)-F_1(t_0)\right|),E_{n,q_n}(\left|\hat{F}_{1,n}(t_0)-F_{1,n}(t_0)\right|)\}\\
\geq\quad &\frac{1}{4}n^{1/3}\left|F_{1,n}(t_0)-F_1(t_0)\right|\left\{1-H^2(q_n,q_0)\right\}^{2n}\\
\cong\quad &\frac{1}{4}n^{1/3}\left|F_{1,n}(t_0)-F_1(t_0)\right|\left[1-\frac{S_2(t_0)g(t_0)f_1^2(t_0)t^3n^{-1}}{S_1(t_0)\{S_2(t_0)-S_1(t_0)\}}\right]^{2n}\\
\cong\quad &\frac{1}{4}n^{1/3}f_1(t_0)tn^{-1/3}\left[1-\frac{S_2(t_0)g(t_0)f_1^2(t_0)t^3n^{-1}}{S_1(t_0)\{S_2(t_0)-S_1(t_0)\}}\right]^{2n}\\
\longrightarrow\quad &\frac{1}{4}f_1(t_0)t\exp\left[-\frac{2S_2(t_0)g(t_0)f_1^2(t_0)t^3}{S_1(t_0)\{S_2(t_0)-S_1(t_0)\}}\right]\\
\equiv\quad &bt\exp(-at^3).
\end{aligned}
$$

The last expression is maximised over t by $t'=(1/3a)^{1/3}$, yielding the minimax lower bound

$$
bt'e^{-1/3}=k\,[f_1(t_0)S_1(t_0)\{S_2(t_0)-S_1(t_0)\}/\{S_2(t_0)g(t_0)\}]^{1/3}
$$

where $k=(1/4)(6e)^{-1/3}$ does not depend on f_1, F_1 or g.

5.2 Proof of Theorem 4.1

In Section 2 we saw that $\hat{S}_{1,n}(t)$ is given by the slope of the least concave majorant at W_j if $t \in (Y_{j-1}, Y_j]$. With $D = \{(t_1,t_2) \in \mathbb{R}^2 \ : \ 0 \le t_1 \le t_2\}$, $I\left(D \times (u,t]\right)$ will denote the indicator function of the set

$$\left\{(t_1,t_2,c) \in \mathbb{R}^3 : 0 \le t_1 \le t_2 < \infty,\ u < c \le t\right\} \ .$$

Let $A = \{(t_1,t_2,c) \in \mathbb{R}^3 : 0 < c < t_1\}$ and $B = \{(t_1,t_2,c) \in \mathbb{R}^3 : 0 < c < t_2\}$, and $Pf = \int f dP$ for any probability measure P. Define the processes

$$\begin{aligned} \mathbb{W}_n(t) &= \mathbb{P}_n\left(\frac{I(B)\, I(D \times (0,t])}{\{1-F_2(c)\}^2}\right) = \int_0^t \int \int_{0<t_1<t_2} \frac{I(t_2 > c)}{\{1-F_2(c)\}^2} d\mathbb{P}_n(t_1,t_2,c) \\ &= \frac{1}{n}\sum_{i=1}^n \frac{I(T_{2,i} > C_i)\, I(C_i \le t)}{\{1-F_2(C_i)\}^2} = W_j \quad \text{for} \quad t \in [Y_j, Y_{j+1}),\ j = 1,\dots,n \end{aligned}$$

and

$$\begin{aligned} \mathbb{V}_n(t) &= \mathbb{P}_n\left(\frac{I(A)\, I(D \times (0,t])}{1-F_2(c)}\right) = \int_0^t \int \int_{0<t_1<t_2} \frac{I(t_1 > c)}{1-F_2(c)} d\mathbb{P}_n(t_1,t_2,c) \\ &= \frac{1}{n}\sum_{i=1}^n \frac{I(T_{1,i} > C_i)\, I(C_i \le t)}{1-F_2(C_i)} = V_j \quad \text{for} \quad t \in [Y_j, Y_{j+1}),\ j = 1,\dots,n. \end{aligned}$$

Then the function $s \mapsto \mathbb{V}_n \circ \mathbb{W}_n^{-1}(s)$ equals the cumulative sum diagram in Section 2. Since $\hat{S}_{1,n}(t)$ is given by the slope of the least concave majorant of the cumulative sum diagram defined by $(\mathbb{W}_n, \mathbb{V}_n)$, we have that if $\hat{S}_{1,n}(t) \le a$ then a line of slope a moved down vertically from $+\infty$ first hits the cumulative sum diagram to the left of t (see Figure 2). The point where the line hits the diagram is the point where $\mathbb{V}_n$ is farthest above the line of slope a through the origin. Thus, with $\hat{s}_n$ defined below,

$$\hat{S}_{1,n}(t) \le a \quad \Leftrightarrow \quad \hat{s}_n(a) = \arg\max_s \{\mathbb{V}_n(s) - a\mathbb{W}_n(s)\} \le t$$

and we can derive the limit distribution of $\hat{S}_{1,n}(t)$ by studying the locations of the maxima of the sequence of processes $s \mapsto \mathbb{V}_n(s) - a\mathbb{W}_n(s)$ since

$$\text{pr}[n^{1/3}\{\hat{S}_{1,n}(t_0) - S_1(t_0)\} \le x] = \text{pr}[\hat{s}_n\{S_1(t_0) + xn^{-1/3}\} \le t_0].$$

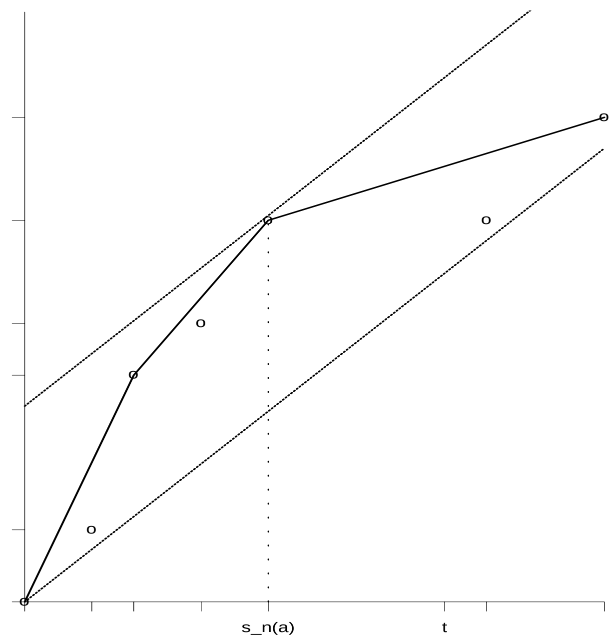

Figure 2: **A cumulative sum diagram and the corresponding Least Concave Majorant.**

Making the change of variables $s \rightarrow t_0 + n^{-1/3}t$ we obtain

$$
\begin{aligned}
& \hat{s}_n\{S_1(t_0) + xn^{-1/3}\} - t_0 \\
= \quad & n^{-1/3} \arg\max_t \left\{ \mathbb{V}_n(t_0 + n^{-1/3}t) - (S_1(t_0) + xn^{-1/3})\mathbb{W}_n(t_0 + n^{-1/3}t) \right\} \\
= \quad & n^{-1/3} \arg\max_t \Bigg\{ \int_0^{t_0+n^{-1/3}t} \int\int_{0<t_1<t_2} \frac{I\,(t_1 > c)\, I\,(t_2 > c)}{1 - F_2(c)} d\mathbb{P}_n\,(t_1, t_2, c) \\
& \qquad - (S_1(t_0) + xn^{-1/3}) \int_0^{t_0+n^{-1/3}t} \int\int_{0<t_1<t_2} \frac{I\,(t_2 > c)}{\{1 - F_2(c)\}^2} d\mathbb{P}_n\,(t_1, t_2, c) \Bigg\} \\
= \quad & n^{-1/3} \arg\max_t \big\{ \mathbb{P}_n I(A) I(D \times (0, t_0 + n^{-1/3}t])/\{1 - F_2(c)\} \\
& \qquad - S_1(t_0)\mathbb{P}_n I(B) I(D \times (0, t_0 + n^{-1/3}t])/\{1 - F_2(c)\}^2 \\
& \qquad - xn^{-1/3}\mathbb{P}_n I(B) I(D \times (0, t_0 + n^{-1/3}t])/\{1 - F_2(c)\}^2 \quad \big\} \; .
\end{aligned}
$$

The location of the maximum of a function does not change when the function is multiplied by a positive constant or shifted vertically. Thus, the arg max above is also a point of maximum of the process

$$\begin{aligned} & n^{2/3}\Bigg[(\mathbb{P}_n - P)\left\{\frac{I\,(A)\,S_2(c) - I(B)S_1(t_0)}{S_2^2(c)}\right\} I(D\times(t_0, t_0+n^{-1/3}t]) \\ & \quad + P\left\{\frac{I(A)S_2(c) - I(B)S_1(t_0)}{S_2^2(c)}\right\} I(D\times(t_0, t_0+n^{-1/3}t]) \\ & \quad - xn^{-1/3}\mathbb{P}_n \frac{I(B)I(D\times(t_0,t_0+n^{-1/3}t])}{S_2^2(c)}\Bigg] \\ = \; & n^{2/3}(M_1+M_2+M_3). \end{aligned} \tag{5.6}$$

So the probability of interest is

$$\mathrm{pr}[\hat{s}_n\{S_1(t_0)+xn^{-1/3}\}\le t_0] = \mathrm{pr}[\arg\max_t\{n^{2/3}(M_1+M_2+M_3)\}\le 0].$$

Notice that F_2 is unknown and should be substituted by $\hat{F}_{2,n}$. We will rewrite the arg max of (5.6) as

$$\begin{aligned} & \arg\max_t n^{2/3}\Bigg\{\big(\mathbb{P}_n - P\big)\Bigg[\Bigg\{\frac{I(A)\hat{S}_{2,n}(c) - I(B)S_1(t_0)}{\hat{S}_{2,n}^2(c)} \\ & \qquad - \frac{I(A)S_2(c) - I(B)S_1(t_0)}{S_2^2(c)}\Bigg\} I(D\times(t_0,t_0+n^{-1/3}t])\Bigg] \\ & \quad + \big(\mathbb{P}_n - P\big)\left[\left\{\frac{I(A)S_2(c) - I(B)S_1(t_0)}{S_2^2(c)}\right\} I(D\times(t_0,t_0+n^{-1/3}t])\right] \\ & \quad + P\Bigg[\Bigg\{\frac{I(A)\hat{S}_{2,n}(c) - I(B)S_1(t_0)}{\hat{S}_{2,n}^2(c)} \\ & \qquad - \frac{I(A)S_2(c) - I(B)S_1(t_0)}{S_2^2(c)}\Bigg] I(D\times(t_0,t_0+n^{-1/3}t])\Bigg] \\ & \quad + P\left[\left\{\frac{I(A)S_2(c) - I(B)S_1(t_0)}{S_2^2(c)}\right\} I(D\times(t_0,t_0+n^{-1/3}t])\right] \\ & \quad - xn^{-1/3}\big(\mathbb{P}_n - P\big)\left[I(B)\left\{\frac{1}{\hat{S}_{2,n}^2(c)} - \frac{1}{S_2^2(c)}\right\} I(D\times(t_0,t_0+n^{-1/3}t])\right] \\ & \quad - xn^{-1/3}P\left(\frac{I(B)I(D\times(t_0,t_0+n^{-1/3}t])}{S_2^2(c)}\right) \quad \Bigg\} \\ = \; & \arg\max_t n^{2/3}\left\{I_1+I_2+I_3+I_4+I_5+I_6\right\} \end{aligned} \tag{5.7}$$

We will now analyze each term in the arg max expression separately. For term I_1 we have

$$n^{2/3}\big(\mathbb{P}_n - P\big)\left[I(t_1 > c)\left\{\frac{1}{\hat{S}_{2,n}(c)} - \frac{1}{S_2(c)}\right\} I(D \times (t_0, t_0 + n^{-1/3}t])\right]$$
$$+\, n^{2/3}\big(\mathbb{P}_n - P\big)\left[I(t_2 > c)S_1(t_0)\left\{\frac{1}{\hat{S}^2_{2,n}(c)} - \frac{1}{S^2_2(c)}\right\} I(D \times (t_0, t_0 + n^{-1/3}t])\right]$$

and each term converges uniformly in probability to 0. In fact,

$$\begin{aligned}
& \sup_{0\le t\le t_0+M}\left|n^{2/3}\big(\mathbb{P}_n - P\big)\left[I(A)\left\{\frac{1}{\hat{S}_{2,n}(c)} - \frac{1}{S_2(c)}\right\} I(D \times (t_0, t_0 + n^{-1/3}t])\right]\right| \\
& + \sup_{0\le t\le t_0+M}\left|n^{2/3}\big(\mathbb{P}_n - P\big)\left[I(B)S_1(t_0)\left\{\frac{1}{\hat{S}^2_{2,n}(c)} - \frac{1}{S^2_2(c)}\right\} \right.\right. \\
& \qquad\qquad \left.\left. \times I(D \times (t_0, t_0 + n^{-1/3}t])\right]\right| \\
= & \sup_{0\le t\le t_0+M}\left|n^{1/2}\big(\mathbb{P}_n - P\big)\left[n^{1/6}I(A)\left\{\frac{S_2(c) - \hat{S}_{2,n}(c)}{S_2(c)\hat{S}_{2,n}(c)}\right\}\right.\right. \\
& \qquad\qquad \left.\left. \times I(c \in (t_0, t_0 + n^{-1/3}t])\right]\right| \\
& + \sup_{0\le t\le t_0+M}\left|n^{1/2}\big(\mathbb{P}_n - P\big)\left[n^{1/6}I(B)S_1(t_0)\left\{\frac{S^2_2(c) - \hat{S}^2_{2,n}(c)}{S^2_2(c)\hat{S}^2_{2,n}(c)}\right\}\right.\right. \\
& \qquad\qquad \left.\left. \times I(c \in (t_0, t_0 + n^{-1/3}t])\right]\right| \\
\le & \frac{n^{1/2}\|\hat{S}_{2,n}(c) - S_2(c)\|_0^T}{\hat{S}_{2,n}(t_0+M)S_2(t_0+M)} n^{1/6} S_1(t_0)\big(\mathbb{P}_n - P\big) I(c \in (t_0, t_0 + n^{-1/3}t) \\
& + \frac{n^{1/2}\|\hat{S}^2_{2,n}(c) - S^2_2(c)\|_0^T}{\hat{S}^2_{2,n}(t_0+M)S^2_2(t_0+M)} n^{1/6}\big(\mathbb{P}_n - P\big) I(c \in (t_0, t_0 + n^{-1/3}t]) \\
= & O_p(1)o_p(1) \;+\; O_p(1)o_p(1)
\end{aligned}$$

since the rate of convergence of $\hat{S}_{2,n}$ is $n^{-1/2}$.

Consider I_2 now. Taking

$$f_{n,t}(t_1, t_2, c) = n^{1/6}[\{I\,(A)\,S_2(c) - I\,(B)\,S_1(t_0)\}/S_2^2(c)]I(D \times (t_0, t_0 + n^{-1/3}t])$$

and $F_n(t_1, t_2, c) = (n^{1/6}/\delta^2)I(D \times (t_0, t_0 + n^{-1/3}t])$ we have, by Theorems 2.11.23 and 2.7.11 in van der Vaart & Wellner (1996),

$$n^{2/3}\big(\mathbb{P}_n - P\big)[\{I(A)S_2(c) - I(B)S_1(t_0)\}/S_2^2(c)]I(D \times (t_0, t_0 + n^{-1/3}t])$$

converging to a mean zero Gaussian process with covariance function (for $0 < s < t$) given by

$$\begin{aligned}
& (n^{2/3}/n^{1/2})^2 E\{E([\{I(A)S_2(C) - I(B)S_1(t_0)\}/S_2^2(C)]^2 \\
& \qquad \times I(D \times (t_0 + n^{-1/3}s, t_0 + n^{-1/3}t]) \;\mid\; C)\} \\
= \; & n^{1/3}E\{([\{S_2(C) - S_1(t_0)\}/S_2^2(C)]^2 S_1(C) \\
& \qquad + \{S_1(t_0)/S_2^2(C)\}^2 \mathrm{pr}(T_1 < C < T_2))I(C \in (t_0 + n^{-1/3}s, t_0 + n^{-1/3}t])\} \\
= \; & n^{1/3}\int_{t_0+n^{-1/3}s}^{t_0+n^{-1/3}t} ([\{S_2(u) - S_1(t_0)\}/S_2^2(u)]^2 S_1(u) \\
& \qquad + \big\{S_1(t_0)/S_2^2(u)\big\}^2 \{S_2(u) - S_1(u)\})g(u)du \\
\cong \; & n^{1/3}([\{S_2(t_0) - S_1(t_0)\}/S_2^2(t_0)]^2 S_1(t_0) \\
& + \{S_1(t_0)/S_2^2(t_0)\}^2\{S_2(t_0) - S_1(t_0)\})g(t_0)n^{-1/3}\,|t-s| \\
= \; & [\{S_2(t_0) - S_1(t_0)\}S_1(t_0)\{S_2(t_0) - S_1(t_0) + S_1(t_0)\}g(t_0)n^{-1/3}\,|t-s|\;]/S_2^4(t_0) \\
= \; & \{S_2(t_0) - S_1(t_0)\}S_1(t_0)g(t_0)\,|t-s|\,/S_2^3(t_0)
\end{aligned}$$

For term I_3 we have

$$\begin{aligned}
& n^{2/3}P\left[\left\{\frac{I(A)\hat{S}_{2,n}(c) - I(B)S_1(t_0)}{\hat{S}_{2,n}^2(c)}\right.\right. \\
& \qquad \left.\left. - \frac{I(A)S_2(c) - I(B)S_1(t_0)}{S_2^2(c)}\right\} I(D \times (t_0, t_0 + n^{-1/3}t])\right] \\
= \; & n^{2/3}P\left[\left\{\frac{I(A)}{\hat{S}_{2,n}(c)} - \frac{I(A)}{S_2(c)}\right\} I(D \times (t_0, t_0 + n^{-1/3}t])\right] \\
& -n^{2/3}P\left[\left\{\frac{I(B)S_1(t_0)}{\hat{S}_{2,n}^2(c)} - \frac{I(B)S_1(t_0)}{S_2^2(c)}\right\} I(D \times (t_0, t_0 + n^{-1/3}t])\right] .
\end{aligned}$$

For the first term in the sum above, assuming S_1 and g continuous, we have

$$
\begin{aligned}
& n^{2/3} P\left[\left\{\frac{1}{\hat{S}_{2,n}(c)} - \frac{1}{S_2(c)}\right\} I(A) I(D \times (t_0, t_0 + n^{-1/3}t])\right] \\
\leq\ & n^{2/3} \int_{t_0}^{t_0+n^{-1/3}t} \int\int_{c<t_1<t_2} \frac{\mid S_2(c) - \hat{S}_{2,n}(c) \mid}{\hat{S}_{2,n}(t_0 + n^{-1/3}t) S_2(t_0 + n^{-1/3}t)} dF(t_1, t_2) dG(c) \\
=\ & n^{2/3} \int_{t_0}^{t_0+n^{-1/3}t} \frac{S_1(c) \mid S_2(c) - \hat{S}_{2,n}(c) \mid}{\hat{S}_{2,n}(t_0 + n^{-1/3}t) S_2(t_0 + n^{-1/3}t)} dG(c) \\
\leq\ & n^{2/3} \frac{\|S_1(t) g(t)\|_{t_0}^{t_0+n^{-1/3}t} n^{-1/3} t \|S_2(t) - \hat{S}_{2,n}(t)\|_{t_0}^{t_0+M}}{\hat{S}_{2,n}(t_0 + n^{-1/3}t) S_2(t_0 + n^{-1/3}t)} \\
=\ & \frac{n^{2/3} O(n^{-1/3}) O_p(n^{-1/2})}{\hat{S}_{2,n}(t_0 + n^{-1/3}t) S_2(t_0 + n^{-1/3}t)} \ = \ O_p(n^{-1/6}) \ = \ o_p(1) \ .
\end{aligned}
$$

Similarly, for the second term, assuming S_2 continuous,

$$
\begin{aligned}
& n^{2/3} P\left(\left[I(B) S_1(t_0) \left\{\frac{1}{\hat{S}^2_{2,n}(c)} - \frac{1}{S^2_2(c)}\right\}\right] I(D \times (t_0, t_0 + n^{-1/3}t])\right) \\
\leq\ & n^{2/3} \int_{t_0}^{t_0+n^{-1/3}t} \int\int_{c<t_2} \frac{\mid S^2_2(c) - \hat{S}^2_{2,n}(c) \mid}{\hat{S}^2_{2,n}(t_0 + n^{-1/3}t) S^2_2(t_0 + n^{-1/3}t)} dF(t_1, t_2) dG(c) \\
=\ & n^{2/3} \int_{t_0}^{t_0+n^{-1/3}t} \frac{S_2(c) \mid S^2_2(c) - \hat{S}^2_{2,n}(c) \mid}{\hat{S}^2_{2,n}(t_0 + n^{-1/3}t) S^2_2(t_0 + n^{-1/3}t)} dG(c) \\
\leq\ & n^{2/3} \frac{\|S_2(t) g(t)\|_{t_0}^{t_0+n^{-1/3}t} n^{-1/3} t \|S^2_2(t) - \hat{S}^2_{2,n}(t)\|_{t_0}^{t_0+M}}{\hat{S}^2_{2,n}(t_0 + n^{-1/3}t) S^2_2(t_0 + n^{-1/3}t)} \\
=\ & \frac{n^{2/3} O(n^{-1/3}) O_p(n^{-1/2})}{\hat{S}_{2,n}(t_0 + n^{-1/3}t) S_2(t_0 + n^{-1/3}t)} \ = \ O_p(n^{-1/6}) \ = \ o_p(1) \ .
\end{aligned}
$$

The limit of term I_4 can be easily calculated as

$$
\begin{aligned}
& n^{2/3}E\{E([\{I(A)S_2(C) - I(B)S_1(t_0)\}/S_2^2(C)]I(D \times (t_0, t_0 + n^{-1/3}t]) \mid C)\} \\
= \; & n^{2/3}E\left[\{S_1(C)S_2(C) - S_2(C)S_1(t_0)\}I(C \in (t_0, t_0 + n^{-1/3}t])/S_2^2(C)\right] \\
= \; & n^{2/3}\int_{t_0}^{t_0+n^{-1/3}t} [\{F_1(t_0) - F_1(u)\}g(u)/S_2(u)]du \\
\cong \; & n^{2/3}\frac{1}{2}\left\{\frac{F_1(t_0) - F_1(t_0 + n^{-1/3}t)}{S_2(t_0 + n^{-1/3}t)}\right\} g(t_0 + n^{-1/3}t)n^{-1/3}t \\
\cong \; & -n^{2/3}f_1(t_0)n^{-1/3}tg(t_0)n^{-1/3}t/\{2S_2(t_0)\} \; = \; -f_1(t_0)g(t_0)t^2/\{2S_2(t_0)\}
\end{aligned}
$$

The uniform convergence in probability to zero of term I_5 in (5.7) has been established (up to the constant $S_1(t_0)$) in the evaluation of the limit of term I_1.

And finally the limit behavior of term I_6 is calculated below.

$$
\begin{aligned}
& - n^{2/3}xn^{-1/3}\mathbb{P}_n\{I(B)I(D \times (t_0, t_0 + n^{-1/3}t])/S_2^2(c)\} \\
= \; & - n^{2/3}xn^{-1/3}\big(\mathbb{P}_n - P\big)\{I(B)I(D \times (t_0, t_0 + n^{-1/3}t])/S_2^2(c)\} \\
& - n^{2/3}xn^{-1/3}P\{I(B)I(D \times (t_0, t_0 + n^{-1/3}t])/S_2^2(c)\} \\
= \; & - n^{1/2}\frac{n^{1/3}}{n^{1/2}}x\big(\mathbb{P}_n - P\big)\{I(B)I(D \times (t_0, t_0 + n^{-1/3}t])/S_2^2(c)\} \\
& - n^{1/3}xP\{I(B)I(D \times (t_0, t_0 + n^{-1/3}t])/S_2^2(c)\} \; .
\end{aligned}
$$

Taking $F_n(t_1, t_2, c) = xI(D \times (t_0, t_0 + n^{-1/3}t])/(n\delta^2)$, Theorems 2.11.23 and 2.7.11 in van der Vaart & Wellner (1996) imply that the first part of the sum above converges to a mean zero Gaussian process with covariance function given by

$$
\begin{aligned}
& x^2(n^{2/3}/n)E(E[\{I(B)/S_2^2(C)\}^2 I(D \times (t_0 + n^{-1/3}s, t_0 + n^{-1/3}t]) \mid C]) \\
= \; & (x^2/n^{1/3})E\{S_2(C)I(C \in (t_0 + n^{-1/3}s, t_0 + n^{-1/3}t])/S_2^4(C)\} \\
= \; & (x^2/n^{1/3})\int_{t_0+n^{-1/3}s}^{t_0+n^{-1/3}t} \left\{g(u)/S_2^3(u)\right\} du \quad \cong \quad x^2n^{-2/3}\frac{g(t_0)}{S_2^3(t_0)}\,|t - s|
\end{aligned}
$$

which converges to 0 as $n \longrightarrow \infty$. The second part gives

$$\begin{aligned}
& -\, xn^{1/3}P\{I(B)I(D\times(t_0,t_0+n^{-1/3}t])/S_2^2(c)\} \\
= & -\, xn^{1/3}E[E\{I(B)I(D\times(t_0,t_0+n^{-1/3}t])/S_2^2(C) \ \mid\ C\}] \\
= & -\, xn^{1/3}\int_{t_0}^{t_0+n^{-1/3}t}\{S_2(u)g(u)/S_2^2(u)\}du \ \cong\ -\, xn^{1/3}g(t_0)tn^{-1/3}/S_2(t_0) \\
= & -\, xg(t_0)t/S_2(t_0)
\end{aligned}$$

Exercise 3.2.5, page 308, in van der Vaart & Wellner (1996) states that the random variables $\arg\max_t\{a\mathbb{B}(t)-bt^2-ct\}$ and $(a/b)^{2/3}\arg\max_t\{\mathbb{B}(t)-t^2\}-c/(2b)$ are equal in distribution, where $\{\mathbb{B}(t): t\in\mathbb{R}\}$ is a standard two-sided Brownian motion with $\mathbb{B}(0)=0$, and a, b and c are positive constants. Thus, making $a=[S_1(t_0)g(t_0)\{S_2(t_0)-S_1(t_0)\}/S_2^3(t_0)]^{1/2}$, $b=f_1(t_0)g(t_0)/\{2S_2(t_0)\}$ and $c \ = \ g(t_0)x/S_2(t_0)$ we have

$$\begin{aligned}
& \mathrm{pr}[n^{1/3}\{\hat{S}_{1,n}(t_0)-S_1(t_0)\}\le x]=\mathrm{pr}[\hat{s}_n\{S_1(t_0)+xn^{-1/3}\}\le t_0] \\
= \ & \mathrm{pr}[\arg\max_t\{n^{2/3}(M_1+M_2+M_3)\}\le 0]\longrightarrow \mathrm{pr}[\arg\max_t\{a\mathbb{B}(t)-bt^2-ct\}\le 0] \\
= \ & \mathrm{pr}[(a/b)^{2/3}\arg\max_t\{\mathbb{B}(t)-t^2\}-c/(2b)\le 0] \\
= \ & \mathrm{pr}\left(2Z\le\left[\frac{2S_2(t_0)g(t_0)}{f_1(t_0)S_1(t_0)\{S_2(t_0)-S_1(t_0)\}}\right]^{1/3}x\right)
\end{aligned}$$

which implies that

$$\mathrm{pr}\left(n^{1/3}\frac{\hat{S}_{1,n}(t_0)-S_1(t_0)}{[f_1(t_0)S_1(t_0)\{S_2(t_0)-S_1(t_0)\}/\{2S_2(t_0)g(t_0)\}]^{1/3}}\le z\right) \ \longrightarrow \ \mathrm{pr}(2Z\le z).$$

References

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. & Brunk, H.D. (1972). *Statistical Inference under Order Restrictions*, New York: John Wiley & Sons.

Dinse, G.E. & Lagakos, S.W. (1982). Nonparametric estimation of lifetime and disease onset distributions from incomplete observations. *Biometrics* **38**, 921-932.

Gomes, A.E. (2001). *Consistency of a Weighted Least Square Estimator of the Disease Onset Distribution Function for a Survival-Sacrifice Model*, Relatório Técnico RTP-11/2001, Departamento de Estatística, Universidade Federal de Minas Gerais.

Groeneboom, P. (1987). *Asymptotics for the interval censored observations*. Technical Report 87-18, Department of Mathematics, University of Amsterdam.

Groeneboom, P. (1989). Brownian motion with a parabolic drift and Airy functions. *Probab. Th. and Rel. Fields* **81**, 79-109.

Groeneboom, P. (1996). Lectures on Inverse Problems, in *Lecture Notes in Mathematics*, **1648**, Springer-Verlag, 67-164.

GROENEBOOM, P. & WELLNER, J.A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*, Boston: Birkhauser.

GROENEBOOM, P. & WELLNER, J.A. (2001). Computing Chernoff's Distribution. *J. Comput. and Graph. Statist.* **10**, 388-400.

JEWELL, N.P. (1982). Mixtures of exponential distribution. *Ann. Statist.* **10**, 479-484.

KODELL, R., SHAW, G. & JOHNSON, A. (1982). Nonparametric joint estimators for disease resistance and survival functions in survival/sacrifice experiments. *Biometrics* **38**, 43-58.

ROBERTSON, T., WRIGHT, F.T. & DYKSTRA, R.L. (1988), *Order Restricted Statistical Inference*, New York: John Wiley & Sons.

TURNBULL, B.W. & MITCHELL, T.J. (1984). Nonparametric estimation of the distribution of time to onset for specific diseases in survival/sacrifice experiments. *Biometrics* **40**, 41-50.

VAN DER LAAN, M.J., JEWELL, N.P. & PETERSON, D. (1997). Efficient estimation of the lifetime and disease onset distribution. *Biometrika* **84**, 539-554.

VAN DER VAART, A.W. & WELLNER, J.A. (1996). *Weak Convergence and Empirical Processes*, New York: Springer.

WANG, J-G. (1987). A note on the uniform consistency of the Kaplan-Meier estimator. *The Ann. Statist.* **15**, 1313-1316.