The Annals of Statistics
1994, Vol. 22, No. 3, 1256–1274

ESTIMATING A MONOTONE DENSITY FROM CENSORED OBSERVATIONS

BY YOUPING HUANG AND CUN-HUI ZHANG[1]

Columbia University and Rutgers University

We study the nonparametric maximum likelihood estimator (NPMLE) for a concave distribution function F and its decreasing density f based on right-censored data. Without the concavity constraint, the NPMLE of F is the product-limit estimator proposed by Kaplan and Meier. If there is no censoring, the NPMLE of f, derived by Grenander, is the left derivative of the least concave majorant of the empirical distribution function, and its local and global behavior was investigated, respectively, by Prakasa Rao and Groeneboom. In this paper, we present a necessary and sufficient condition, a self-consistency equation and an analytic solution for the NPMLE, and we extend Prakasa Rao's result to the censored model.

1. Introduction. Suppose $(X, Y), (X_1, Y_1), \ldots, (X_n, Y_n)$ are independent identically distributed (iid) random vectors, where X is a lifetime with a concave cumulative distribution function (cdf) F on $(0, \infty)$, and Y is a censoring variable independent of X with a possibly discontinuous cdf G. The observed data are $T_i = \min\{X_i, Y_i\}$ and $\delta_i = I\{X_i \leq Y_i\}$, $i = 1, \ldots, n$. Let f be the left derivative of F. We are interested in the nonparametric maximum likelihood estimator (NPMLE) of f, and its asymptotic distribution.

This problem may arise in nonparametric estimation in renewal processes. We imagine that a renewal process began indefinitely far in the past. The item currently in service is inspected for a period of time Y or until it fails, whichever occurs first. Then, in the absence of censoring, the observed lifetime X of the item under inspection has the limiting distribution of residual lifetime in an ordinary renewal process, which has a monotone density $f(x) = \mu_0^{-1}(1 - F_0(x))$, where F_0 is the cdf of interarrival times and μ_0 its mean. Therefore, finding the NPMLE of f, and eventually F_0, based on iid observations under random censorship is a case of the problem under study.

If the concavity constraint is removed, the NPMLE of F is the well-known product-limit estimator proposed by Kaplan and Meier (1958). If there is no censoring, the NPMLE of F is the least concave majorant (LCM) of the empirical distribution function, derived by Grenander (1956). For various asymptotic properties of the product-limit estimator, see Breslow and Crowley (1974) and Gill (1983); for local and global asymptotics of the Grenander estimator, see Prakasa Rao (1969) and Groeneboom (1985), respectively. We shall extend Prakasa Rao's result to the censored model.

Received June 1991; revised November 1993.

[1]Research partially supported by the National Science Foundation and Army Research Office.

AMS 1991 *subject classifications.* 62G05, 62G30, 62E20.

Key words and phrases. Nonparametric maximum likelihood estimation, censored data, monotone density, self-consistency, least concave majorant, product-limit estimator.

The problem has been studied numerically; see, for example, Laslett (1982), Vardi (1989) and Groeneboom and Wellner (1992) for recursive algorithms. We shall present a necessary and sufficient condition, a self-consistency equation and an analytic solution for the NPMLE. Our computer simulations show that the NPMLE of F performs better than the LCM of the product-limit estimator for a moderate sample size ($n = 50$), but their derivatives have no significant difference, when both X and Y are unit exponential random variables.

Throughout this paper, we use $\int_a^b$ to represent $\int_{(a,b]}$, and "$-$" over any cdf to denote the corresponding survival function.

2. Main results. Let us write observed values of T_i, $1 \le i \le n$, in strictly ascending order:

$$(0 = t_0 <)\, t_1 < t_2 < \cdots < t_n.$$

Define H_{0n}, H_{1n} and H_n to be the respective empirical versions of

$$H_j(t) = P\{T \le t,\ \delta = j\}, \qquad j = 0, 1, \qquad H(t) = H_0(t) + H_1(t).$$

Although ties may occur to censored observations due to discontinuities of G, we assume that the observations have no ties to simplify our notation. Only slight modifications are needed to accomodate the general situation. The NPMLE of f is the left-continuous decreasing nonnegative function maximizing the log-likelihood

$$l = \sum_{i=1}^{n} \delta_i \log f(t_i) + \sum_{i=1}^{n} (1 - \delta_i) \log\bigl(1 - F(t_i)\bigr) \tag{2.1}$$

and satisfying $\int_0^\infty f(t)\,dt \le 1$, where $F(t) = \int_0^t f(s)\,ds$.

Notice that any concave function for which a section is replaced by a linear function is still concave. If F is replaced by a linear function in $[t_{i-1}, t_i]$, then by the left-continuity of f, $f(t_i)$ increases while other items on the right-hand side of (2.1) are unchanged, so that l increases. Furthermore, if $\delta_i = 0$ and F is replaced by a linear function in $[t_{i-1}, t_{i+1}]$, then l also increases, since only three items may change: $\overline{F}(t_i)$ and $f(t_{i+1})$, both increasing, and $f(t_i)$, which decreases but has no effect on l. Therefore, we only need to maximize l over all f of the form of decreasing step function with jumps at uncensored t_i.

Let $\nabla l(\mathbf{f})$ denote the differentiation operator acting on l with respect to $\mathbf{f} = (f(t_1), \dots, f(t_n))$ as a vector in n-dimensional space, and let $\mathbf{1}_k = (1, \dots, 1, 0, \dots, 0)$ be an n-dimensional vector having 1's as its first k components and 0's as its last $n - k$ components, for $k = 1, \dots, n$. The following theorem gives a characterization of the NPMLE.

THEOREM 1. *A left-continuous nonnegative decreasing function f with $\int_0^\infty f(t)\,dt \le 1$ is the NPMLE if and only if both of the following inequalities* *are satisfied*,

$$\langle \nabla l(\mathbf{f}), \mathbf{f} \rangle \geq 0, \tag{2.2}$$

$$\left\langle \nabla l(\mathbf{f}), \mathbf{f} - \frac{1}{t_k}\mathbf{1}_k \right\rangle \geq 0 \quad \textit{for } k = 1, \dots, n. \tag{2.3}$$

In addition, the equality in (2.3) *holds if* f *is the NPMLE and* $df(t_k) < 0$.

REMARK. We can use the following formulas to calculate (2.2) and (2.3),

$$\frac{1}{n}\langle \nabla l(\mathbf{f}), \mathbf{f} \rangle = 1 - \int_0^\infty \frac{dH_{0n}(t)}{1 - F(t)}, \tag{2.4}$$

$$\begin{aligned} &\frac{1}{n}\left\langle \nabla l(\mathbf{f}), \mathbf{f} - \frac{1}{t_k}\mathbf{1}_k \right\rangle \\ &\quad = 1 - \frac{1}{t_k}\left(\int_0^{t_k} \frac{dH_{1n}(t)}{f(t)} - \int_0^{t_k} \frac{t_k - t}{1 - F(t)} dH_{0n}(t) \right) \quad \text{for } k = 1, \dots, n. \end{aligned} \tag{2.5}$$

For any function v on an interval J, define the following: $Mv = M_J v$ is the LCM of v on J; $Dv = D_J v$ is the left derivative of Mv.

Efron (1967) proved that the product-limit estimator F_n^* satisfies $F_n^* = Q_n F_n^*$, where

$$Q_n \Phi(t) = H_n(t) - \int_0^t \frac{1 - \Phi(t)}{1 - \Phi(s)} dH_{0n}(s),$$

and he named the property self-consistency. We shall show an analogy in the following corollary.

COROLLARY 1. *Let* F_n *and* f_n *be the NPMLE's of* F *and its density* f, *respectively. Set*

$$K_n(t) = Q_n F_n(t) = H_n(t) - \int_0^t \frac{1 - F_n(t)}{1 - F_n(s)} dH_{0n}(s). \tag{2.6}$$

Then $F_n(t) \geq K_n(t)$ *for all* $t \geq 0$, *where equality holds if* $df_n(t_i) < 0$, *or, equivalently*, $F_n = MQ_nF_n$ *and* $f_n = DQ_nF_n$.

In particular, when there is no censoring, the self-consistency equation $F_n = MQ_nF_n$ reduces to $F_n = MH_n$, which is the case of the Grenander estimator.

EXAMPLE 1. Suppose only one censored time $t_1 < 1$ and one uncensored time $t_2 = 1$ are observed. Then $F_n(t) = at$ in $[0, t_2]$ for some $0 < a \leq 1$. The likelihood equals $(1 - at_1)a$, which reaches its maximum at $a = \min\{(2t_1)^{-1}, 1\}$. Therefore, we have $F_n(t_2) < 1$ if $t_1 > 0.5$. In this case, it is also clear that the NPMLE F_n is smaller than $\min\{t, 1\}$, the LCM of the product-limit estimator.

Like the product-limit estimator, $F_n(t_n) < 1$ if the maximum observation t_n is censored. However, this may occur in our case even for uncensored t_n, as indicated in Example 1. It is obvious that F_n is determined up to the maximum observation t_n, censored or uncensored. When $F_n(t_n) < 1$ and $\delta_n = 0$, $F_n(t)$ has to be equal to $F_n(t_n)$ for $t > t_n$; otherwise it would have increased the likelihood to replace F_n by a straight line in $[t_{n-1}, t_n + c]$ for some large c. However, when $F_n(t_n) < 1$ and $\delta_n = 1$, $F_n(t)$ may be defined arbitrarily for $t > t_n$, as long as it is kept concave. When $F_n(\infty) < 1$ happens, F_n is not a proper cdf. Such an F_n may be looked at as if it has a positive mass at ∞.

We write jumps of f_n in $[0, t_n]$ in strictly ascending order,

$$(0 = s_0 <)\, s_1 < s_2 < \cdots < s_m.$$

They are among uncensored observations. In order to calculate F_n analytically, we shall introduce a function $\varphi(\cdot\,; k, t_i)$ for $k = 1, \dots, m$ and uncensored $t_i > s_{k-1}$ as

$$\varphi(u; k, t_i) = (t_i - s_{k-1})\left(1 - \int_0^{s_{k-1}} \frac{dH_{0n}(t)}{\overline{F_n}(t)}\right) - \frac{1}{u}\int_{s_{k-1}}^{t_i} dH_{1n}(t) - \int_{s_{k-1}}^{t_i} \frac{(t_i - t)\, dH_{0n}(t)}{\overline{F_n}(s_{k-1}) - (t - s_{k-1})u},$$

which depends on F_n only through its values on $(0, s_{k-1}]$. It can be found that φ is a concave function of u with at least one root in $(0, \overline{F_n}(s_{k-1})(t_i - s_{k-1})^{-1}]$. Let $u(k, t_i)$ be the smallest root. Then we have the following theorem on calculating f_n and F_n.

THEOREM 2. *The NPMLE f_n is a left-continuous step function in $[0, t_n]$ which jumps at uncensored s_k, $k = 1, \dots, m$. The values of s_k and $f_n(s_k)$ can be calculated in the order of ascending s_k by*

$$s_k = \max\{t_i\colon\ u(k, t_i) = u^*(k),\ t_i > s_{k-1},\ \delta_i = 1\}, \qquad f_n(s_k) = u(k, s_k),$$

where $u^*(k) = \max\{u(k, t_i)\colon t_i > s_{k-1}, \delta_i = 1\}$ and s_m is the maximum uncensored observation. If $F_n(s_m) < 1$ and $s_m = t_n$, $f_n(t)$ is unspecified for $t > s_m$; otherwise, $f_n(t) = 0$ for $t > s_m$.

For any function $\psi(t)$ and vector $(t_1, \dots, t_k)$, we shall use the following notation to denote the k-dimensional vector:

$$\psi(t)|_{t = t_1, \dots, t_k} = (\psi(t_1), \dots, \psi(t_k)).$$

Prakasa Rao (1969) and Groeneboom (1985) derived the asymptotic distribution of the density of the Grenander estimator. We present an analogous result in the following theorem.

THEOREM 3. *Suppose $f'(x) < 0$, $G(x) < 1$ and $G(x+t) - G(x) = o(t^{1/2})$ as $t \to 0$, at some $x > 0$. Define*

$$\gamma_n(t) = \left(\frac{2n\overline{G}(x)}{f(x)|f'(x)|}\right)^{1/3}\left(f_n\left(x+tn^{-1/3}\right) - f(x)\right) \quad \text{and} \quad a(x) = \left(\frac{|f'(x)|^2\overline{G}(x)}{4f(x)}\right)^{1/3}$$

Then, for every vector $(t_1,\dots,t_k)$, $\gamma_n(t)|_{t=t_1,\dots,t_k}$ converges to $D\eta(t)|_{t=a(x)t_1,\dots,a(x)t_k}$ in distribution as $n \to \infty$, where $\eta(t) = W(t) - t^2$ and $\{W(t), -\infty < t < \infty\}$ is a two-sided standard Brownian motion with $W(0) = 0$.

Since $\psi_n(t)|_{t=t_1,\dots,t_k} \to \psi(t)|_{t=t_1,\dots,t_k}$ for all vectors $(t_1,\dots,t_k)$ implies $\psi_n \to \psi$ in

$$L^p[-c,c] = \left\{\psi\colon \int_{-c}^{c} |\psi(t)|^p\,dt < \infty\right\}$$

for monotone functions ψ_n, Theorem 3 immediately implies the following corollary.

COROLLARY 2. *Under the conditions of Theorem 3, $\gamma_n(t) \to_{\mathcal{D}} D\eta(a(x)t)$ in $L^p[-c,c]$ as $n \to \infty$, for all $c > 0$.*

Theorem 1 and Corollary 1 will be proved in Section 3. The proof of Theorem 2 is quite technical, and it can be found in Huang (1994). Some brief discussion of the function φ is given at the end of Section 3. In Section 4, we show that the LCM of the product-limit estimator is always greater than or equal to the NPMLE of F, and we discuss some preliminary results concerning uniform consistency of F_n and f_n. Section 5 is devoted to the proof of Theorem 3. Finally, in Section 6, we compare the NPMLE of F with the LCM of the product-limit estimator by computer simulations.

3. A necessary and sufficient condition. We replace f_n and F_n by $\widehat{f}$ and $\widehat{F}$, respectively, in this section since n is fixed.

The following proof employs a method similar to that in Proposition 1.1 of Groeneboom and Wellner [(1992), page 39].

PROOF OF THEOREM 1. We shall use f_i and F_i for $f(t_i)$ and $F(t_i)$ respectively. Then $\mathbf{f} = (f_1,\dots,f_n)$.

(*Necessity*.) Suppose $\widehat{f}$ is the NPMLE. Notice that, for $k = 1,\dots,n$, $(1-\varepsilon)\widehat{\mathbf{f}} + \varepsilon\cdot t_k^{-1}\mathbf{1}_k$ corresponds to the density of a concave distribution, possibly subdistribution, if $\varepsilon > 0$ is small, or if $|\varepsilon|$ is small and $df(t_k) < 0$ ($f_k > f_{k+1}$). Therefore, we have

$$\begin{aligned} 0 &\geq \lim_{\varepsilon \to 0+} \frac{l\left((1-\varepsilon)\widehat{\mathbf{f}} + \varepsilon\cdot t_k^{-1}\mathbf{1}_k\right) - l(\widehat{\mathbf{f}})}{\varepsilon} \\ &= \sum_{i=1}^{n} \frac{\partial l(\widehat{\mathbf{f}})}{\partial \widehat{f}_i}\left(-\widehat{f}_i + \frac{I\{i \leq k\}}{t_k}\right) = -\left\langle \nabla l(\widehat{\mathbf{f}}), \widehat{\mathbf{f}} - \frac{1}{t_k}\mathbf{1}_k\right\rangle, \end{aligned}$$

where equality holds if $df(t_k) < 0$. In addition, since $(1-\varepsilon)\widehat{\mathbf{f}}$ corresponds to the density of a concave subdistribution for small $\varepsilon > 0$,

$$0 \geq \lim_{\varepsilon \to 0+} \frac{l((1-\varepsilon)\widehat{\mathbf{f}}) - l(\widehat{\mathbf{f}})}{\varepsilon} = \sum_{i=1}^{n} \frac{\partial l(\widehat{\mathbf{f}})}{\partial \widehat{f}_i}(-\widehat{f}_i) = -\langle \nabla l(\widehat{\mathbf{f}}), \widehat{\mathbf{f}} \rangle.$$

(*Sufficiency*.) Suppose (2.2) and (2.3) hold for some $\widehat{\mathbf{f}}$. Define $f_{n+1} = 0$. Then $\mathbf{f} = \sum_{k=1}^{n}(f_k - f_{k+1})\mathbf{1}_k$ and

$$\sum_{k=1}^{n} t_k(f_k - f_{k+1}) = \int_0^{t_n} f(t)\,dt = 1 - a(\mathbf{f}),$$

where $a(\mathbf{f}) \geq 0$. It follows from the concavity of $l(\mathbf{f})$ that

$$\begin{aligned} l(\mathbf{f}) - l(\widehat{\mathbf{f}}) &\leq \langle \nabla l(\widehat{\mathbf{f}}), \mathbf{f} - \widehat{\mathbf{f}} \rangle \\ &= \sum_{k=1}^{n} (f_k - f_{k+1}) \langle \nabla l(\widehat{\mathbf{f}}), \mathbf{1}_k - t_k \widehat{\mathbf{f}} \rangle - a(\mathbf{f}) \langle \nabla l(\widehat{\mathbf{f}}), \widehat{\mathbf{f}} \rangle \leq 0. \end{aligned}$$ □

Equation (2.4) can be derived by

$$\begin{aligned} \frac{1}{n}\langle \nabla l(\mathbf{f}), \mathbf{f} \rangle &= \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\delta_i}{f_i} - \sum_{j=i}^{n} (1-\delta_j) \frac{t_i - t_{i-1}}{1 - F_j} \right) f_i \\ &= \frac{1}{n} \sum_{i=1}^{n} \delta_i - \frac{1}{n} \sum_{j=1}^{n} \frac{1-\delta_j}{1-F_j} \sum_{i=1}^{j} f_i(t_i - t_{i-1}) \\ &= \int_0^{\infty} dH_{1n}(t) - \int_0^{\infty} \frac{F(t)}{1-F(t)}\, dH_{0n}(t) \\ &= 1 - \int_0^{\infty} \frac{dH_{0n}(t)}{1-F(t)}, \end{aligned}$$

and (2.5) can be obtained from (2.4) and the following:

$$\begin{aligned} \frac{1}{n}\langle \nabla l(\mathbf{f}), \mathbf{1}_k \rangle &= \frac{1}{n} \sum_{i=1}^{k} \left(\frac{\delta_i}{f_i} - \sum_{j=i}^{n} (1-\delta_j) \frac{t_i - t_{i-1}}{1 - F_j} \right) \\ &= \frac{1}{n} \sum_{i=1}^{k} \frac{\delta_i}{f_i} - \frac{1}{n} \sum_{j=1}^{n} \frac{1-\delta_j}{1-F_j} \sum_{i=1}^{j \wedge k} (t_i - t_{i-1}) \\ &= \int_0^{t_k} \frac{dH_{1n}(t)}{f(t)} - \int_0^{\infty} \frac{t_k \wedge t}{1-F(t)}\, dH_{0n}(t). \end{aligned}$$

PROOF OF COROLLARY 1. We shall show by induction the following equivalent inequality:

$$\left(1 - \int_0^{t_i} \frac{dH_{0n}(s)}{1-\widehat{F}(s)}\right)(1 - \widehat{F}(t_i)) \leq \int_{t_i}^{\infty} dH_n(s) \quad \text{for } i = 0, \dots, n. \tag{3.1}$$

If $i = 0$, then the equality in (3.1) holds trivially. Let $s_{k-1} = t_j < t_i \le s_k$. Then it follows from (2.5) that

$$
\begin{aligned}
0 &\le \frac{1}{n}\langle \nabla l(\widehat{\mathbf{f}}), t_i\widehat{\mathbf{f}} - \mathbf{1}_i\rangle - \frac{1}{n}\langle \nabla l(\widehat{\mathbf{f}}), t_j\widehat{\mathbf{f}} - \mathbf{1}_j\rangle \\
&= (t_i - s_{k-1})\left(1 - \int_0^{s_{k-1}} \frac{dH_{0n}(s)}{1-\widehat{F}(s)}\right) - \int_{s_{k-1}}^{t_i} \frac{dH_{1n}(t)}{\widehat{f}(t)} \\
&\quad - \int_{s_{k-1}}^{t_i} \frac{t_i - t}{1-\widehat{F}(t)}\, dH_{0n}(t).
\end{aligned}
\tag{3.2}
$$

Multiplying both sides of (3.2) by $\widehat{f}(t_i)$, which is unchanged for $s_{k-1} < t_i \le s_k$, and then moving the last two terms to the left-hand side, we obtain

$$
\begin{aligned}
&\int_{s_{k-1}}^{t_i} dH_{1n}(t) + \int_{s_{k-1}}^{t_i} \frac{\widehat{F}(t_i) - \widehat{F}(t)}{1-\widehat{F}(t)}\, dH_{0n}(t) \\
&\qquad \le \left(\widehat{F}(t_i) - \widehat{F}(s_{k-1})\right)\left(1 - \int_0^{s_{k-1}} \frac{dH_{0n}(s)}{1-\widehat{F}(s)}\right),
\end{aligned}
\tag{3.3}
$$

where equality holds if $t_i = s_k$. Now,

$$
\begin{aligned}
&(1-\widehat{F}(t_i))\left(1 - \int_0^{t_i} \frac{dH_{0n}(s)}{1-\widehat{F}(s)}\right) \\
&= \left((1-\widehat{F}(s_{k-1})) - (\widehat{F}(t_i) - \widehat{F}(s_{k-1}))\right)\left(1 - \int_0^{s_{k-1}} \frac{dH_{0n}(s)}{1-\widehat{F}(s)}\right) \\
&\quad - \int_{s_{k-1}}^{t_i} \frac{1-\widehat{F}(t_i)}{1-\widehat{F}(t)}\, dH_{0n}(t) \\
&\le \int_{s_{k-1}}^{\infty} dH_n(t) - \int_{s_{k-1}}^{t_i} dH_{1n}(t) - \int_{s_{k-1}}^{t_i} dH_{0n}(t) \quad \text{[by induction and (3.3)]} \\
&= \int_{t_i}^{\infty} dH_n(t).
\end{aligned}
$$

This implies $\widehat{F} \ge Q_n\widehat{F}$, and the proof is complete. □

It is easy to see that (3.2) depends on $\widehat{F}$ only through its values on $(0, t_i]$. The definition of $\varphi(u; k, t_i)$ is nothing but replacing $\widehat{f}(t)$ by u for $s_{k-1} < t \le t_i$ in (3.2). We can obtain immediately that $\varphi(\widehat{f}(t_i); k, t_i) \ge 0$ for $s_{k-1} < t_i \le s_k$, where equality holds if $t_i = s_k$. The motivation of the algorithm in Theorem 2 is that after $\widehat{f}$ is determined on $(0, s_{k-1}]$ we pretend that the next jump of $\widehat{f}$ occurs at t_i, and we calculate root $u(k, t_i)$ of $\varphi(\cdot\,; k, t_i)$ for all $t_i > s_{k-1}$ with $\delta_i = 1$, so that $\widehat{f}(s_k)$ is among those roots. Hence the question is to locate the position of s_k. It turns out that $u(k, t_i) \le \widehat{f}(s_k)$ for $t_i \le s_k$, and $u(k, t_i) < \widehat{f}(s_k)$ for $t_i > s_k$. See Huang (1994) for details of the proof.

4. Uniform consistency. In this section, we compare the NPMLE F_n with the product-limit estimate F_n^* and prove some preliminary results concerning uniform consistency of F_n and f_n.

PROPOSITION 1. *$MF_n^*(t) \geq F_n(t)$ for all $t \geq 0$.*

REMARK. In general, we have $MF_n^*(t) \neq F_n(t)$; see Example 1 in Section 2 or the simulation results in Section 6.

PROOF OF PROPOSITION 1. Efron (1967) proved that

$$F_n^*(t) = H_n(t) - \int_0^t \frac{\overline{F_n^*}(t)}{\overline{F_n^*}(s)}\, dH_{0n}(s). \tag{4.1}$$

Let G_n^* denote the product-limit estimator of G. Noticing $\overline{H_n}(t) = \overline{F_n^*}(t)\overline{G_n^*}(t)$ [Shorack and Wellner (1986), page 295], we have, for $F_n^*(t) < 1$,

$$\overline{F_n^*}(t)\, dG_n^*(t) = dH_{0n}(t), \qquad G_n^*(t) = \int_0^t \frac{dH_{0n}(s)}{\overline{F_n^*}(s)}. \tag{4.2}$$

Note that (4.2) holds for all t since $dH_{0n}(t) = 0$ when $F_n^*(t) = 1$. Subtracting (2.6) from (4.1) yields

$$F_n^*(t) - K_n(t) = \int_0^t \left(\frac{\overline{F_n}(t)}{\overline{F_n}(s)} - \frac{\overline{F_n^*}(t)}{\overline{F_n^*}(s)} \right) dH_{0n}(s).$$

Setting $\lambda_n = F_n^* - F_n$ and $\mu_n = F_n - K_n$, we have, noticing (4.2),

$$\lambda_n(t) + \mu_n(t) = -\int_0^t \frac{\overline{F_n}(t)}{\overline{F_n}(s)} \lambda_n(s)\, dG_n^*(s) + \lambda_n(t)\, G_n^*(t). \tag{4.3}$$

If $F_n(t) < 1$, then

$$\frac{\mu_n(t)}{\overline{F_n}(t)} = -\frac{\lambda_n(t)\overline{G_n^*}(t)}{\overline{F_n}(t)} + \int_0^t \frac{\lambda_n(s)}{\overline{F_n}(s)}\, d\overline{G_n^*}(s) = -\int_0^t \overline{G_n^*}(s-)\, d\frac{\lambda_n(s)}{\overline{F_n}(s)},$$

which implies

$$\frac{\lambda_n(t)}{\overline{F_n}(t)} = -\int_0^t \frac{1}{\overline{G_n^*}(s-)}\, d\frac{\mu_n(s)}{\overline{F_n}(s)} = -\frac{\mu_n(t)}{\overline{F_n}(t)\overline{G_n^*}(t)} + \int_0^t \frac{\mu_n(s)}{\overline{F_n}(s)}\, d\frac{1}{\overline{G_n^*}(s)}.$$

If F_n has a vertex at t, then $\mu_n(t) = 0$, and $\lambda_n(t) \geq 0$ since $\mu_n \geq 0$. On the other hand, if $F_n(t) = 1$, then $F_n^*(t) = 1$. Thus $MF_n^* \geq F_n$. □

LEMMA 1. *Suppose $F(T) < 1$ and $G(T) < 1$. Then $\sup_{0 \leq t \leq T} |F_n(t) - F(t)| = O_p(n^{-1/2})$.*

PROOF. Using the same notation as above and, in addition, $\nu_n = MF_n^* - F_n^*$, we can derive from (4.3) that

$$\begin{aligned}\lambda_n(t) &= \frac{1}{\overline{G_n^*}(t)}\left(-\mu_n(t) - \int_0^t \frac{\overline{F_n}(t)}{\overline{F_n}(s)}\lambda_n(s)\,dG_n^*(s)\right)\\ &\le \frac{1}{\overline{G_n^*}(t)}\int_0^t \frac{\overline{F_n}(t)}{\overline{F_n}(s)}\nu_n(s)\,dG_n^*(s) \le \frac{1}{\overline{G_n^*}(T)}\|\nu_n\|_T,\end{aligned}$$

where $\|\cdot\|_T$ denotes the supremum norm over $[0, T]$. Thus, we obtain

$$\begin{aligned}\overline{G_n^*}(T)\|F_n^* - F_n\|_T &\le \|MF_n^* - F_n^*\|_T\\ &\le \|MF_n^* - F\|_T + \|F_n^* - F\|_T\\ &\le \|F_n^* - F\|_{\tau_n} + \|F_n^* - F\|_T \quad \text{[Marshall (1970)]},\end{aligned}$$

where $\tau_n = \min\{t \ge T\colon DF_n^*(t+) < DF_n^*(t)\}$. Hence, the required result is a consequence of the weak convergence of the product-limit estimators F_n^* and G_n^*; see Breslow and Crowley (1974). □

LEMMA 2. *Suppose $H(T) < 1$. (i) If F is differentiable at $x \in (0, T)$, then $f_n(x) = f(x) + o_p(1)$. (ii) If $f'(t)$ exists for all $t \in (0, T)$ and $0 < a < b < T$, then $\sup_{a \le t \le b}|f_n(t) - f(t)| = O_p(n^{-1/4})$.*

PROOF. Restating Lemma 1, we know that, for any $\varepsilon > 0$, there exist A and n_1 such that, for all $n > n_1$,

$$P\left\{\sup_{0 \le t \le T}|F_n(t) - F(t)| < An^{-1/2}\right\} > 1 - \varepsilon. \tag{4.4}$$

(i) For any $\delta > 0$, there exists an n_2 such that, for all $n > n_2$,

$$\left|\pm n^{1/4}\left[F\left(x \pm n^{-1/4}\right) - F(x)\right] - f(x)\right| < \delta.$$

For $n > \max\{n_1, n_2\}$, we have, from the concavity of F_n and (4.4),

$$\begin{aligned}&P\left\{f_n(x) - f(x) < \delta + 2An^{-1/4}\right\}\\ &\quad\ge P\left\{\frac{F_n\left(x - n^{-1/4}\right) - F_n(x)}{-n^{-1/4}} < \frac{F\left(x - n^{-1/4}\right) - F(x)}{-n^{-1/4}} + 2An^{-1/4}\right\}\\ &\quad= P\left\{F_n\left(x - n^{-1/4}\right) - F_n(x) > F\left(x - n^{-1/4}\right) - F(x) - 2An^{-1/2}\right\} \ge 1 - \varepsilon.\end{aligned}$$

Similarly, we have $P\{f_n(x) - f(x) > -\delta - 2An^{-1/4}\} \ge 1 - \varepsilon$. Together, they imply

$$P\left\{|f_n(x) - f(x)| < \delta + 2An^{-1/4}\right\} \ge 1 - 2\varepsilon.$$

(ii) There exist constants B and n_2 such that

$$\left|\pm n^{1/4}\left[F\left(x \pm n^{-1/4}\right) - F(x)\right] - f(x)\right| < n^{-1/4}B, \quad \forall\, a \le x \le b,\ n > n_2.$$

Letting $\delta = n^{-1/4}B$, we have, as in part (i), $P\{|f_n(x) - f(x)| < (2A + B)n^{-1/4}, \forall\, x \in [a, b]\} \ge 1 - 2\varepsilon$. □

5. Asymptotic distribution of NPMLE. We shall prove Theorem 3 here. Throughout this section, x is fixed. We need two lemmas, whose proofs will be given in the Appendix. An alternative method for the proof of Lemma 4 can be found in Groeneboom (1985) and Groeneboom and Wellner (1992).

LEMMA 3. *If $f'(x) < 0$ and $G(x) < 1$, then there exists an $\varepsilon > 0$ such that $H(x + 2\varepsilon) < 1$ and*

$$\lim_{n \to \infty} P\left\{f_n(t) = D_{[0, x+2\varepsilon]}K_n(t),\ \forall\, t \in [0, x + \varepsilon]\right\} = 1.$$

LEMMA 4. *Suppose $\eta_n(t) = u_n(t) + W_n(v_n(t)) + \alpha_n(t)$, where W_n is a two-sided Brownian motion with $W_n(0) = 0$, u_n is a concave function with $u_n(0) = 0$ and $u_n(t) \le 0$, v_n is a function with $v_n(0) = 0$ and α_n is a process. Let v be a continuous function and let u be a concave function such that $u'(t)$ is continuous at $t_1, \dots, t_k$. If*

$$(5.1) \qquad u_n(t) \to u(t) \quad \text{and} \quad v_n(t) \to v(t) \quad \text{uniformly on compact sets},$$

$$(5.2) \qquad \sup_{|t| \le c} |\alpha_n(t)| \to 0 \quad \text{in probability} \quad \forall\, c > 0,$$

$$(5.3) \qquad \limsup_{t \to \pm\infty} \frac{u(t)}{u(2t)} < \rho \quad \text{for some } \rho < \frac{1}{2},$$

$$(5.4) \qquad \lim_{c \to \infty} \limsup_{n \to \infty} P\left\{ \sup_{|t| \ge c} \left| \frac{\eta_n(t)}{u_n(t)} - 1 \right| > \varepsilon \right\} = 0 \quad \forall\, \varepsilon > 0,$$

then

$$(5.5) \qquad D\eta_n(t)\Big|_{t = t_1, \dots, t_k} \to_{\mathcal{D}} D\Big(u(t) + W\big(v(t)\big)\Big)\Big|_{t = t_1, \dots, t_k},$$

where W is a two-sided Brownian motion with $W(0) = 0$.

PROOF OF THEOREM 3. Lemma 3 enables us to restrict t in $[0, x + 2\varepsilon]$, and we will do so without stating it explicitly.

We divide the proof into three steps: first define $\eta_n(t)$ of the form as in Lemma 4; then verify the conditions; and, finally, calculate both sides of (5.5).

Step 1. [Define $\eta_n(t)$.] Subtracting

$$F(t) = H(t) - \int_0^t \frac{\overline{F}(t)}{\overline{F}(s)}\, dH_0(s)$$

from both sides of (2.6) gives

$$\begin{aligned} K_n(t) - F(t) &= \big(H_n(t) - H(t)\big) - \int_0^t \frac{\overline{F_n}(t)}{\overline{F_n}(s)}\, dH_{0n}(s) + \int_0^t \frac{\overline{F}(t)}{\overline{F}(s)}\, dH_0(s) \\ &= \big(H_n(t) - H(t)\big) - \int_0^t \frac{\overline{F_n}(t)}{\overline{F_n}(s)}\, d\big(H_{0n}(s) - H_0(s)\big) \\ &\quad - \int_0^t \left(\frac{\overline{F_n}(t)}{\overline{F_n}(s)} - \frac{\overline{F}(t)}{\overline{F}(s)}\right)\overline{F}(s)\, dG(s). \end{aligned}$$

Integrating by parts with the first integration on the right-hand side and some algebra with the second yield

$$K_n(t) = F(t) + \big(H_{1n}(t) - H_1(t)\big) + \big(F_n(t) - F(t)\big)G(t) + \zeta_n(t),$$

where $\zeta_n(t) = \zeta_{0n}(t) - \zeta_{1n}(t)$ and

$$\begin{aligned} \zeta_{0n}(t) &= \int_0^t \big(H_{0n}(s-) - H_0(s-)\big)\, d_s \frac{\overline{F_n}(t)}{\overline{F_n}(s)}, \\ \zeta_{1n}(t) &= \int_0^t \frac{\overline{F_n}(t)}{\overline{F_n}(s)}\big(F_n(s) - F(s)\big)\, dG(s). \end{aligned}$$

Define

$$\xi_n(t) = K_n(x+t) - K_n(x) - \big(F_n(x+t) - F_n(x)\big)G(x) - tf(x)\overline{G}(x). \tag{5.6}$$

Then,

$$\begin{aligned} \xi_n(t) = {} & \big[F(x+t) - F(x) - tf(x)\big]\overline{G}(x) \\ & + \big[H_{1n}(x+t) - H_{1n}(x) - H_1(x+t) + H_1(x)\big] \\ & + \big[F_n(x+t) - F(x+t)\big]\big[G(x+t) - G(x)\big] + \big[\zeta_n(x+t) - \zeta_n(x)\big]. \end{aligned}$$

By Komlós, Major and Tusnády (1975),

$$H_{1n}(t) - H_1(t) = n^{-1/2}B_n\big(H_1(t)\big) + O_p\big(n^{-1}\log n\big),$$

where B_n is a Brownian bridge and O_p is uniform in t.

Define

$$\eta_n(t) = n^{2/3}\xi_n\big(tn^{-1/3}\big). \tag{5.7}$$

Then $\eta_n(t) = u_n(t) + W_n(v_n(t)) + \alpha_n(t)$, where, with a standard normal random variable Z independent of B_n,

$$\begin{aligned}
u_n(t) &= n^{2/3}\Big[F\big(x+tn^{-1/3}\big) - F(x) - tn^{-1/3}f(x)\Big]\overline{G}(x),\\
W_n(s) &= n^{1/6}\Big[B_n\big(sn^{-1/3} + H_1(x)\big) - B_n\big(H_1(x)\big) + sn^{-1/3}Z\Big],\\
v_n(t) &= n^{1/3}\Big[H_1\big(x+tn^{-1/3}\big) - H_1(x)\Big],\\
\alpha_n(t) &= n^{2/3}\Big[F_n\big(x+tn^{-1/3}\big) - F\big(x+tn^{-1/3}\big)\Big]\Big[G\big(x+tn^{-1/3}\big) - G(x)\Big]\\
&\quad + O_p\big(n^{-1/3}\log n\big) - n^{1/6}\Big[H_1\big(x+tn^{-1/3}\big) - H_1(x)\Big]Z\\
&\quad + n^{2/3}\Big[\zeta_n\big(x+tn^{-1/3}\big) - \zeta_n(x)\Big].
\end{aligned}$$

Clearly, W_n is a two-sided Brownian motion with $W_n(0) = 0$ and η_n has the required form.

Step 2. Verify the conditions of Lemma 4. It is easy to see that (5.1) and (5.3) hold with $u(t) = \frac{1}{2}t^2 f'(x)\overline{G}(x)$ and $v(t) = tf(x)\overline{G}(x)$. Before verifying other conditions, we shall show

$$\alpha_n(t) = o_p(1) + o_p\big(t^{1/2}\big) + O_p\big(tn^{-1/6}\big), \tag{5.8}$$

where o_p and O_p are independent of t. By lemma 1, the differentiability of F and H_1 at x, and the Hölder continuity of G at x, we have

$$\alpha_n(t) = o_p(1) + o_p\big(t^{1/2}\big) + O_p\big(tn^{-1/6}\big) + n^{2/3}\Big[\zeta_n\big(x+tn^{-1/3}\big) - \zeta_n(x)\Big].$$

Since $|F_n(x + tn^{-1/3}) - F_n(x)| \le O_p(n^{-1/2}) + |F(x + tn^{-1/3}) - F(x)| = O_p(n^{-1/2}) + O(tn^{-1/3})$, we have

$$\begin{aligned}
&n^{2/3}\Big[\zeta_{1n}\big(x+tn^{-1/3}\big) - \zeta_{1n}(x)\Big]\\
&= n^{2/3}\bigg(\int_0^x \frac{\overline{F_n}\big(x+tn^{-1/3}\big) - \overline{F_n}(x)}{\overline{F_n}(s)}\big(F_n(s) - F(s)\big)\,dG(s)\\
&\qquad + \int_x^{x+tn^{-1/3}} \frac{\overline{F_n}\big(x+tn^{-1/3}\big)}{\overline{F_n}(s)}\big(F_n(s) - F(s)\big)\,dG(s)\bigg)\\
&= n^{2/3}\Big[\Big(O_p\big(n^{-1/2}\big) + O\big(tn^{-1/3}\big)\Big)O_p\big(n^{-1/2}\big) + O_p\big(n^{-1/2}\big)o\big(t^{1/2}n^{-1/6}\big)\Big]\\
&= O_p\big(n^{-1/3}\big) + O_p\big(tn^{-1/6}\big) + o_p\big(t^{1/2}\big).
\end{aligned}$$

Similarly, we can get $n^{2/3}[\zeta_{0n}(x + tn^{-1/3}) - \zeta_{0n}(x)] = O_p(n^{-1/3}) + O_p(tn^{-1/6})$. Hence, (5.8) is verified. Because (5.8) implies (5.2) directly, the only condition which has not been checked is (5.4). It follows from (5.8) that

$$\lim_{c\to\infty}\limsup_{n\to\infty} P\left\{\sup_{|t|\ge c}\left|\frac{\alpha_n(t)}{t}\right| > \varepsilon\right\} = 0,$$

for all $\varepsilon > 0$. Since

$$\sup_{|t|\geq c} \frac{v_n(t)}{t} \leq \sup_{s\neq 0} \frac{H_1(x+s)-H_1(x)}{s} = O(1),$$

using the strong law of large numbers for the Brownian motion, we have

$$\lim_{c\to\infty} \limsup_{n\to\infty} P\left\{ \sup_{|t|\geq c} \left| \frac{W_n(v_n(t))}{t} \right| > \varepsilon \right\} = 0,$$

for all $\varepsilon > 0$. Therefore, (5.4) is satisfied, since

$$\begin{aligned} \sup_{|t|\geq c} \left| \frac{t}{u_n(t)} \right| &\leq \frac{1}{\overline{G}(x)} \sup_{s\neq 0} \frac{\min\left(|s|n^{-1/3}, s^2 c^{-1}\right)}{|F(x+s)-F(x)-sf(x)|} \\ &= O\left(n^{-1/3}+c^{-1}\right). \end{aligned}$$

Step 3. [Apply Lemma 4.] It is easy to see that, for any real numbers $\theta \in [0,1]$, a, b and function ψ,

$$M(\psi - \theta M\psi + a + bt) = (1-\theta)M\psi + a + bt$$

and

$$D(\psi - \theta M\psi + a + bt) = (1-\theta)D\psi + b.$$

So, it follows from (5.6) and (5.7) that

$$\begin{aligned} D\eta_n(t) &= D\left(n^{2/3}\left[K_n\left(x+tn^{-1/3}\right) - F_n\left(x+tn^{-1/3}\right)G(x) - tn^{-1/3}f(x)\overline{G}(x) + \cdots\right]\right) \\ &= n^{2/3}\overline{G}(x)\left[DK_n\left(x+tn^{-1/3}\right) - n^{-1/3}f(x)\right] \\ &= n^{1/3}\overline{G}(x)\left[f_n\left(x+tn^{-1/3}\right) - f(x)\right]. \end{aligned}$$

On the other hand, let $s = at$ and $W'(s) = b^{-1}W(b^2 s)$, where $a = a(x)$ as in the statement of the theorem and $b = b(x) = (2f^2|f'|^{-1}\overline{G})^{1/3}$. Then W' is also a Brownian motion and

$$\begin{aligned} D\left(u(t) + W\left(v(t)\right)\right) &= D\left(t^2 2^{-1} f'\overline{G} + W(tf\overline{G})\right) \\ &= D\left(W\left(b^2 s(t)\right) - bs^2(t)\right) \\ &= bD\left(W'(s) - s^2\right)\cdot\frac{ds}{dt} \\ &= \left(2^{-1}f|f'|\overline{G}^2\right)^{1/3} D\left(W'(s) - s^2\right). \end{aligned}$$

TABLE 1

	Mean of $b_n^{(1)}$	Variance of $b_n^{(1)}$	Mean of $b_n^{(2)}$	Variance of $b_n^{(2)}$	Mean of $b_n^{(\infty)}$	Variance of $b_n^{(\infty)}$
F_n	0.112	0.00437	0.092	0.00216	0.118	0.00259
MF_n^*	0.123	0.00432	0.102	0.00240	0.134	0.00333
f_n	0.240	0.00583	0.341	0.09968		
DMF_n^*	0.242	0.00496	0.346	0.09907		

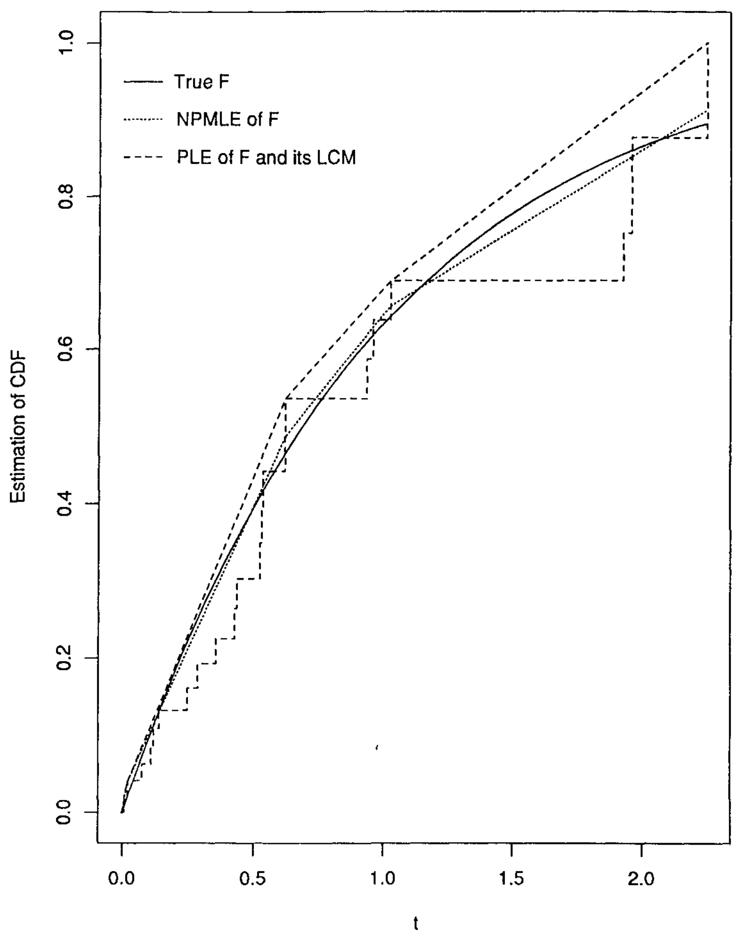

FIG. 1. *NPMLE of F and LCM of PLE.*

Since

$$\gamma_n(t) = \frac{n^{1/3}\overline{G}(x)\left[f_n\left(x + tn^{-1/3}\right) - f(x)\right]}{\left(2^{-1}f|f'|\overline{G}^2\right)^{1/3}} = \frac{D\eta_n(t)}{\left(2^{-1}f|f'|\overline{G}^2\right)^{1/3}},$$

the required result follows from (5.5). □

6. Numerical results. We have performed a computer simulation to compare the NPMLE F_n with MF_n^*, the LCM of the product-limit estimator (PLE), and their derivatives. The simulation was based on 2000 random samples each of size $n = 50$. The lifetime and censoring variables are both exponentially distributed with mean 1. Performance of an estimate $\widetilde{\Phi}_n$ of Φ is measured by $b_n^{(j)} = (\int_0^{s_m} |\widetilde{\Phi}_n(t) - \Phi(t)|^j\, dt)^{1/j}$, for $j = 1, 2$, and $b_n^{(\infty)} = \sup_{0 < t \leq s_m} |\widetilde{\Phi}_n(t) - \Phi(t)|$, where s_m is the largest uncensored observation for each sample. The results are listed in Table 1. We can observe that F_n is indeed better than MF_n^*, while f_n and DMF_n^* show no significant difference.

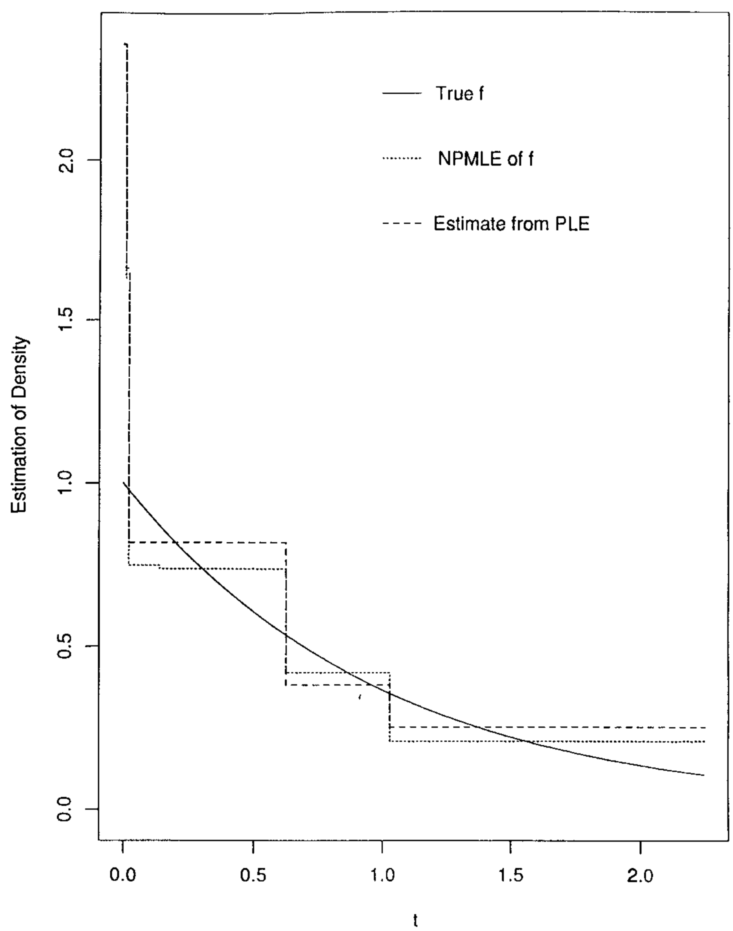

FIG. 2. *NPMLE of f and estimate from PLE.*

The mean and variance of $b_n^{(\infty)}$ for density estimates are left blank in Table 1 because of the instability and inconsistency of both estimators of $f(t)$ at $t = 0+$. Woodroofe and Sun (1993) provided consistent estimators of $f(0+)$ in the absence of censoring, which can also be used to obtain consistent estimates in the right-censoring case as $H_1'(t) = f(t)\{1 - G(t)\}$ is decreasing and $H_1'(0+) = f(0+)$.

Figure 1 is the graph of F_n and MF_n^* along with the true cdf of X, and Figure 2 is the graph of their derivatives along with the true density of X. It is one instance in the simulation. A spiking problem at 0+ with estimates of density can be seen in Figure 2.

APPENDIX

PROOF OF LEMMA 3. There exists an $\varepsilon > 0$ such that $F(t)$ is not a straight line in $[x+\varepsilon, x+2\varepsilon]$ and $H(x+2\varepsilon) < 1$. Since

$$\inf_{a,b} \sup_{x+\varepsilon \le t \le x+2\varepsilon} |at + b - F(t)| = c_0 > 0,$$

we have, from Lemma 1,

$$\begin{aligned} &P\{D_{[0,x+2\varepsilon]}K_n(t) \ne f_n(t),\ \exists\, t \in [0, x+\varepsilon]\} \\ &\quad \le P\{f_n(t) \text{ has no jump in } [x+\varepsilon, x+2\varepsilon]\} \\ &\quad \le P\Big\{\sup_{x+\varepsilon \le t \le x+2\varepsilon} |F_n(t) - F(t)| \ge c_0\Big\} \to 0 \quad \text{as } n \to \infty. \qquad \square \end{aligned}$$

PROOF OF LEMMA 4. It follows from the arguments of Prakasa Rao [(1969), Lemma 4.2] that (5.5) is a consequence of the following three relations:

$$\lim_{c\to\infty} \liminf_{n\to\infty} P\{D_{[-c,c]}\eta_n(t)|_{t=t_1,\dots,t_k} = D\eta_n(t)|_{t=t_1,\dots,t_k}\} = 1; \tag{A.1}$$

$$\lim_{c\to\infty} P\Big\{D_{[-c,c]}\big(u + W(v)\big)\big|_{t=t_1,\dots,t_k} = D\big(u+W(v)\big)\big|_{t=t_1,\dots,t_k}\Big\} = 1; \tag{A.2}$$

$$D_{[-c,c]}\eta_n(t)|_{t=t_1,\dots,t_k} \to_{\mathcal{D}} D_{[-c,c]}\big(u + W(v)\big)\big|_{t=t_1,\dots,t_k} \quad \text{as } n \to \infty \text{ for all } c > t^*, \tag{A.3}$$

where $t^* = \max_{1\le j\le k} |t_j| + 1$.

First we prove (A.1). Define a linear function

$$l_c(t) = \frac{u(c)}{2} - \frac{u(2c)}{4} + t\frac{u(2c)}{2c}.$$

Let $\sigma_1 = \rho/2 + 3/4$. Then $0 < \sigma_1 < 1$. It follows from (5.3) that, for large c,

$$\begin{aligned} l_c(c) &= \frac{u(c)}{2} + \frac{u(2c)}{4} < u(c)\Big(\frac{1}{2} + \frac{1}{4\rho}\Big) < \sigma_1 u(c) < 0 \quad \text{and} \\ l_c(2c) &= \frac{u(c)}{2} + \frac{3u(2c)}{4} > \sigma_1 u(2c), \end{aligned}$$

so that by (5.1) there exists an n_c such that, for $n > n_c$,

$$l_c(c) < \sigma_1 u_n(c) \quad \text{and} \quad l_c(2c) > \sigma_1 u_n(2c),$$

which imply $l_c(t) > \sigma_1 u_n(t)$ for all $t > 2c$ by the concavity of u_n. It follows from (5.4) that

$$\lim_{c\to\infty} \liminf_{n\to\infty} P\{\eta_n(t) < l_c(t),\ \forall t > 2c\} = 1. \tag{A.4}$$

Since

$$\sigma_2 = -\frac{\rho}{2} + \frac{1}{4} - \frac{t^*}{2c} \to -\frac{\rho}{2} + \frac{1}{4} > 0 \quad \text{as } c \to \infty,$$

we can get from (5.3) that

$$\inf_{t<t^*} l_c(t) = l_c(t^*) = \frac{u(c)}{2} - \frac{u(2c)}{4} + \frac{t^* u(2c)}{2c} > \sigma_2 |u(2c)| \quad \text{for large } c,$$

so that $\inf_{t<t^*} l_c(t) \to \infty$ as $c \to \infty$. It follows from (5.4), (5.1) and (5.2) that

$$\begin{aligned} &P\{\eta_n(t) \ge l_c(t),\ \exists t < t^*\} \\ &\quad \le P\Big\{\sup_{t<-c'} \eta_n(t) > 0\Big\} + P\Big\{\sup_{-c'\le t<t^*} \eta_n(t) \ge \inf_{t<t^*} l_c(t)\Big\} \\ &\quad \le P\Big\{\sup_{t<-c'} \Big|\frac{\eta_n(t)}{u_n(t)} - 1\Big| > 1\Big\} + P\Big\{\sup_{-c'\le t<t^*} \eta_n(t) \ge \inf_{t<t^*} l_c(t)\Big\} \\ &\quad \to 0 \quad \text{as } n\to\infty,\ c\to\infty,\ \text{then } c'\to\infty. \end{aligned} \tag{A.5}$$

In addition, we have $\eta_n(c) = u_n(c)(1+o_p(1)) = u(c)(1+o_p(1))$ and $\sigma_1 u(c) > l_c(c)$ for large c, which imply

$$\lim_{c\to\infty} \liminf_{n\to\infty} P\{\eta_n(c) > l_c(c)\} = 1. \tag{A.6}$$

Together (A.4), (A.5) and (A.6) lead to $P\{D\eta_n(t)$ has a vertex in $(t^*, 2c)\} \to 1$ as $n \to \infty$ and then $c \to \infty$. For the same reason, we can derive $P\{D\eta_n(t)$ has a vertex in $(-2c, -t^*)\} \to 1$. Hence, (A.1) is proved.

Relation (A.2) is a special case of (A.1) since

$$\begin{aligned} &P\Big\{\sup_{|t|\ge c} \Big|\frac{u(t) + W(v(t))}{u(t)} - 1\Big| > \varepsilon\Big\} \\ &\quad = \lim_{c'\to\infty} P\Big\{\sup_{c\le|t|<c'} \Big|\frac{u(t) + W_n(v(t))}{u(t)} - 1\Big| > \varepsilon\Big\} \\ &\quad = \lim_{c'\to\infty}\lim_{n\to\infty} P\Big\{\sup_{c\le|t|<c'} \Big|\frac{\eta_n(t)}{u_n(t)} - 1\Big| > \varepsilon\Big\} \\ &\quad \le \limsup_{n\to\infty} P\Big\{\sup_{|t|\ge c} \Big|\frac{\eta_n(t)}{u_n(t)} - 1\Big| > \varepsilon\Big\}. \end{aligned}$$

We now turn to prove (A.3). In view of (5.1) and (5.2), η_n can be written as $u(t)+W_n(v(t))+\beta_n(t)$ with $\beta_n(t) \to 0$ in probability uniformly on all compact sets, so that it suffices to show that, for any fixed $c > t^*$,

$$\text{(A.7)}\quad D_{[-c,c]}\big(u+W_n(v)+\beta_n\big)\big|_{t=t_1,\dots,t_k} \to_{\mathcal{D}} D_{[-c,c]}\big(u+W(v)\big)\big|_{t=t_1,\dots,t_k} \quad \text{as } n\to\infty.$$

Let $D[-c, c]$ be the metric space of left-continuous functions with right limits on $[-c,c]$ equipped with the uniform metric $\|\cdot\|$, and $C[-c,c]$ the subset of all continuous functions in $D[-c,c]$. Define a subset $\widetilde{C}[-c,c]$ of $C[-c,c]$ by

$$\widetilde{C}[-c,c] = \big\{\psi \in C[-c,c]\colon D_{[-c,c]}\psi(t) \text{ is continuous at } t = t_1,\dots,t_k\big\}.$$

Since $D_{[-c,c]}\psi$ is determined by the values of ψ on the set of all rational numbers, the mapping $\psi \to D_{[-c,c]}\psi|_{t=t_1,\dots,t_k}$ is measurable with respect to the projection σ-field. In adition, $\widetilde{C}[-c,c]$ is measurable with respect to the projection σ-field.

If $\{\psi_n\}$ is any sequence of functions in $D[-c,c]$ converging to a function ψ in $\widetilde{C}[-c,c]$, then $M_{[-c,c]}\psi_n \to M_{[-c,c]}\psi$ uniformly on $[-c,c]$. Let D^- and D^+ denote, respectively, the left and right derivatives of the LCM. Then, $D^-_{[-c,c]}\psi_n(t) \to D^-_{[-c,c]}\psi(t)$ if $D_{[-c,c]}\psi$ is continuous at t, since [see Barlow, Bartholomew, Bremner and Brunk (1972), page 228]

$$\begin{aligned} D^+_{[-c,c]}\psi(t) &\le \liminf_{n\to\infty} D^+_{[-c,c]}\psi_n(t) \le \liminf_{n\to\infty} D^-_{[-c,c]}\psi_n(t)\\ &\le \limsup_{n\to\infty} D^-_{[-c,c]}\psi_n(t) \le D^-_{[-c,c]}\psi(t)\end{aligned}$$

and $D^+_{[-c,c]}\psi(t) = D^-_{[-c,c]}\psi(t)$. Thus, the convergence of ψ_n to ψ in $D[-c,c]$ and the membership of ψ in $\widetilde{C}[-c,c]$ imply

$$D_{[-c,c]}\psi_n(t)|_{t=t_1,\dots,t_k} \to D_{[-c,c]}\psi(t)|_{t=t_1,\dots,t_k} \quad \text{as } n\to\infty.$$

If $D\eta = D(u + W(v))$ is not continuous at t_j, then

$$\inf_{t<t_j} \frac{\eta(t)-\eta(t_j)}{t-t_j} > \sup_{t>t_j} \frac{\eta(t)-\eta(t_j)}{t-t_j}.$$

The continuity of u' at t_j implies

$$\text{(A.8)}\qquad \liminf_{t\to t_j-} \frac{W(v(t)) - W(v(t_j))}{t-t_j} > \limsup_{t\to t_j+} \frac{W(v(t)) - W(v(t_j))}{t-t_j}.$$

It follows from the law of the iterated logarithm for Brownian motion $W(v(t_j)+s)$ as $s \to 0$ that the left-hand side of (A.8) is less than or equal to 0 almost surely, while the right-hand side is larger than or equal to 0 almost surely, so that relation (A.8) holds with probability 0. Hence $P\{u + W(v) \in \widetilde{C}[-c,c]\} = 1$, and (A.7) follows from the continuous mapping theorem [cf., e.g., Pollard (1984)]. □

REFERENCES

BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference under Order Restrictions.* Wiley, New York.

BRESLOW, N. E. and CROWLEY, J. (1974). A large-sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2** 437–453.

EFRON, B. (1967). The two sample problem with censored data. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **4** 831–853. Univ. California Press, Berkeley.

GILL, R. D. (1983). Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.* **11** 49–58.

GRENANDER, U. (1956). On the theory of mortality measurement, Part II. *Scand. Actuar. J.* **39** 125–153.

GROENEBOOM, P. (1985). Estimating a monotone density. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R. A. Olshen, eds.) **2** 539–555. Wadsworth, Belmont, CA.

GROENEBOOM, P. and WELLNER, J. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation.* Birkhäuser, Boston.

HUANG, Y. (1994). Some problems in nonparametric estimation based on incomplete data. Ph.D. dissertation, Columbia Univ.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 437–481.

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent rv's, and the sample df. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131.

LASLETT, G. M. (1982). The survival curve under monotone density constraints with applications to two-dimensional line segment processes. *Biometrika* **69** 153–160.

MARSHALL, A. W. (1970). Discussion on Barlow and van Zwet's paper. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.) 174–176. Cambridge Univ. Press.

POLLARD, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.

PRAKASA RAO, B. L. S. (1969). Estimation of a unimodel density. *Sankhyā Ser. A* **31** 23–36.

SHORACK, G. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

VARDI, Y. (1989). Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. *Biometrika* **76** 751–761.

WOODROOFE, M. and SUN J. (1993). A penalized maximum likelihood estimate of $f(0+)$ when f is nonincreasing. *Statist. Sinica* **3** 501–515.

CENTER FOR RADIOLOGICAL RESEARCH
COLUMBIA UNIVERSITY
630 W. 168TH STREET
NEW YORK, NY 10032

DEPARTMENT OF STATISTICS
HILL CENTER FOR THE
MATHEMATICAL SCIENCES
BUSCH CAMPUS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903