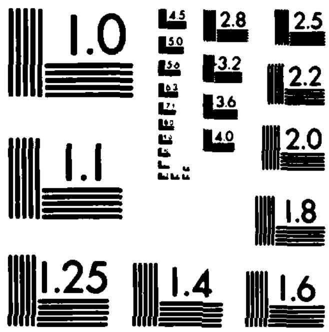
AD-A172 031 HAZARD RATE ESTIMATION FOR CENSORED DATA VIA STRONG 1/1
REPRESENTATION OF THE..(U) CALIFORNIA UNIV DAVIS
INTERCOLLEGE DIV OF STATISTICS S H LO ET AL. AUG 85

UNCLASSIFIED UCD-STATISTICS-64 AFOSR-TR-86-0694 F/G 12/1 NL


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6

(2)

AFOSR-TR- 86-0694

AD-A172 031

HAZARD RATE ESTIMATION FOR CENSORED DATA VIA STRONG REPRESENTATION OF THE KAPLAN-MEIER ESTIMATOR

by

S. H. Lo, Y. P. Mack and J. L. Wang

Technical Report No. 64

August 1985

Division of Statistics

University of California, Davis

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A172 031

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY: NA	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): UCD STATISTICS REPORT No. 64 AFOSR-85-0268	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 86-0694
6a. NAME OF PERFORMING ORGANIZATION: University of California, Davis — 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NM
6c. ADDRESS (City, State and ZIP Code): Division of Statistics, University of California, Davis, CA 95616	7b. ADDRESS (City, State and ZIP Code): Bldg.410, Bolling AFB,DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR — 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-85-0268
8c. ADDRESS (City, State and ZIP Code): BLDG. 410, Bolling AFB,DC	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A5	

11. TITLE (Include Security Classification) (see 16)
Hazard Rate Estimation for Censored Data...

12. PERSONAL AUTHOR(S)
S. H. Lo, Y. P. Mack and J. L. Wang

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical Report	FROM ______ TO ______	August 1985	34

16. SUPPLEMENTARY NOTATION
Hazard Rate Estimation for Censored Data via Strong Representation of the Kaplan-Meier Estimator.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Censored data, hazard rate, kernel method, strong representation, consistency, mean squared error.
XXXXXXXXXXXXXX			

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

~~We study~~ the estimation of a hazard rate function based on censored data by the kernel smoothing method. Our technique is facilitated by a recent result of Lo and Singh (1984) which establishes a strong uniform approximation of the Kaplan-Meier estimator by an average of independent random variables. Pointwise and uniform strong consistency are derived, as well as the mean squared error expression and asymptotic normality, which is obtained using a more traditional method, as compared with the Hajek projection employed by Tanner and Wong (1983).

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Jane-Ling Wang	(916)752-7624	AFOSR/NM

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

86 9 18 031

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

HAZARD RATE ESTIMATION FOR CENSORED DATA VIA STRONG REPRESENTATION OF THE KAPLAN-MEIER ESTIMATOR

Running Title: Hazard rate estimation under censoring

by

S. H. Lo

Rutgers University
New Brunswick, New Jersey

Y. P. Mack[1] and J. L. Wang[2]

University of California
Davis, California

ABSTRACT

We study the estimation of a hazard rate function based on censored data by the kernel smoothing method. Our technique is facilitated by a recent result of Lo and Singh (1984) which establishes a strong uniform approximation of the Kaplan-Meier estimator by an average of independent random variables. Pointwise and uniform strong constistency are derived, as well as the mean squared error expression and asymptotic normality, which is obtained using a more traditional method, as compared with the Hajek projection employed by Tanner and Wong (1983).

[1] Research supported in part by
[2] Research supported in part by U.S. Air Force Grant AFOSR 85-0268.

AMS 1980 subject classifications. Primary 62G05, secondary 62P10, 62E20.

Key words and phrases. Censored data, hazard rate, kernel method, stong representation, consistency, mean squared error.

1. Introduction

Suppose $T_1,\ldots,T_n$ are i.i.d. nonnegative random variables ("lifetimes") with common distribution function (d.f.) $F(\cdot)$ and suppose $C_1,\ldots C_n$ are i.i.d. nonnegative random variables ("censoring sequence") with common d.f. $G(\cdot)$. Assume also that the lifetimes and censoring sequence are independent. In the setting of survival analysis data with random right censorship, one observes the bivariate sample $(X_1,\delta_1),\ldots(X_n,\delta_n)$, where

$$(1) \qquad X_i = T_i \wedge C_i, \quad \delta_i = 1\{T_i < C_i\}$$

with $\wedge$ denoting minimum and $1\{\cdot\}$ denoting the indicator function on a set. One question of interest in survival analysis is the estimation of the hazard rate function $h(\cdot)$, defined as follows when it is further assumed that F has a density $f(\cdot)$:

$$(2) \qquad h(x) = \frac{d}{dx}[-\log \bar{F}(x)] = f(x)/\bar{F}(x), \qquad F(x) < 1,$$

with $\bar{F} = 1-F$. (The quantity $H(x) = -\log \bar{F}(x)$ is called the cumulative hazard function.) In the setting without censoring, parametric models of monotone failure rate have been extensively studied (see Ch. 3 of Barlow and Proschan (1975)). The nonparametric estimation of h(x) was initiated by Watson and Leadbetter (1964a, 1964b). Subsequent research works include Barlow and van Zwet (1971), Ahmad (1976), Rice and Rosenblatt (1976), Ahmad and Lin (1977) and Singpurwalla and Wong (1983). There are essentially 3 variants based on the delta-sequence smoothing introduced by Watson and Leadbetter (1964a, 1964b) and Rice and Rosenblatt (1976) (the third variant):

(3) $$h_n^{(1)}(x) = \int k_n(x-u)dF_n(u)/\bar{F}_n(x), \qquad F_n(x) < 1;$$

(4) $$h_n^{(2)}(x) = \int k_n(x-u)\frac{dF_n(u)}{\bar{F}_n(u)}$$

$$= \sum_{j=1}^{n} k_n(x-X_{(j)}) \cdot \frac{1}{n-j+1} \; ;$$

(5) $$h_n^{(3)}(x) = \int k_n(x-u)dH_n(x)$$

$$= \sum_{j=1}^{n} k_n(x-X_{(j)}) \log \left[1 + \frac{1}{n-j+1}\right],$$

where F_n is the empirical d.f., H_n is the empirical cumulative hazard function, $X_{(j)}$ is the j^{th} order statistic from the sample $\{X_i,\ i=1,\ldots,n\}$; and $\{k_n(\cdot)\}$ is a delta-sequence (see Walter and Blum (1979)), which in the kernel case (see Rosenblatt (1956)) is specialized by taking

(6) $$k_n(v) = \frac{1}{b_n} k\left(\frac{v}{b_n}\right),$$

where k is usually a bounded, symmetric, density function, and $\{b_n\}$ is a so-called band sequence such that $b_n \downarrow 0$, $nb_n \to \infty$ as $n \to \infty$. The method of analysis in the uncensored case in Rice and Rosenblatt (1976) parallels that of kernel density estimation and exploits heavily the strong approximation of the empirical process by a Brownian bridge (Komlós, Major and Tusnády (1975)).

When the data are subjected to random right censoring, the problem becomes more complex, primarily because the estimate of $F(\cdot)$, due to Kaplan and Meier (1958), now takes on a product form:

$$KM_n(x) \equiv \begin{cases} 1 - \prod\limits_{X_{(i)}<x}^{n} \left(\frac{n-i}{n-i+1}\right)^{\delta(i)} & \text{if } x < X_{(n)}; \\ 1 \quad \text{if } x > X_{(n)} \text{ and the largest observation} & \text{is uncensored.} \end{cases}$$

Since many well-studied properties of the empirical d.f. cannot be readily transferred to the Kaplan-Meier estimator, several researchers circumvented the technical difficulty by considering an equivalent problem on the uncensored observations (for example, Blum and Susarla (1980), Burke (1983), Yandell (1983), Liu and Van Ryzin (1985)). Some researchers (for instance Ramlau-Hansen (1983)), employed the method of counting processes studied by Aalen (1978), and Gill (1983). Still others (Földes, Rejtö and Winter (1981), Burke and Horváth (1984)) used a Chung-Smirnov type result on the Kaplan-Meier estimator. To the credit of Tanner and Wong (1983), expressions for the bias and variance in the kernel case (essentially the form $h_n^{(2)}(x)$ given in (4)) were obtained by direct calculations and asymptotic normality was proved by appealing to Hajek's projection. Tanner (1983) and Liu and Van Ryzin (1985) also considered the variable kernel case along the line of the nearest-neighbor method (see Mack and Rosenblatt (1979)). Padgett and McNichols (1984) gave a review of density and failure rate estimators from censored data.

Our present research is motivated by a recent result of Lo and Singh (1984) which establishes a strong uniform approximation of the Kaplan-Meier estimator by an average of i.i.d. random variables with a sufficiently small error. This allows for a more traditional approach to the hazard estimation problem. As constrasted with approaches mentioned in the paragraph above, our method will be a direct one. Although it will become apparent that we could equally well have considered the variants $h_n^{(2)}(x)$, or $h_n^{(3)}(x)$, since there have been fewer investigations carried out for $h_n^{(1)}$ with censored data, the estimator we use will be of the form given by $h_n^{(1)}(x)$ (see (3)) with $F_n(x)$ replaced by a modified version $\Gamma_n(x)$ of the Kaplan-Meier estimator defined as follows to avoid the possibility that $KM_n(x)= 1$:

$$(8)\quad \Gamma_n(x) \equiv \begin{cases} 1 - \prod\limits_{X_{(i)}<x}^{n} \left(\frac{n-i+1}{n-i+2}\right)^{\delta(i)}, & \text{if } x < X_{(n)}; \\ \Gamma_n(X_{(n)}) & \text{if } x > X_{(n)} \text{ and the largest observation is uncensored.} \end{cases}$$

It is easily checked that $\bar{\Gamma}_n(x) > \frac{1}{n+1}$ for all x, and that

$$(9)\qquad \sup_{0<x<T} \left|KM_n(x) - \Gamma_n(x)\right| \overset{a.s.}{=} O\left(n^{-1}\right),$$

for any $0<T < \inf\{t>0\colon L(t) = 1\}$, where $\bar{L}(x) = \bar{F}(x)\cdot\bar{G}(x) = P(T_i>x,\ C_i>x)$. (Hereafter, a.s. will be an abbreviation for "almost surely.")

In Section 2, we state the preliminaries needed for our presentation. In Section 3, we focus our attention on kernel density estimation under censoring via strong approximation. In Section 4, we give the consistency, asymptotic normality and mean squared error expression of our hazard rate estimate. Finally, in the last section, we conclude with relevant comments and some comparison with the nearest neighbor method.

2. Preliminaries

We will concentrate our analysis on the kernel method. We assume throughout our discussion that $L(x) < 1$ for a given point x under consideration. The assumptions we made on the kernel k are as follows:

(k1) $k(x)$ is a symmetric density function.

(k2) $k(x)$ is compactly supported with support $[-c,c]$.

(k3) k is continuous on its support.

(k4) k is of bounded variations with total variation $|k|$.

These assumptions are the "usual" ones encountered in the kernel method of curve estimation. We will comment on the use of kernels with vanishing moments in the last section. The estimate that we consider are modelled after $h_n^{(1)}(x)$ (we continue to label these as $h_n^{(1)}(x)$ for convenience):

$$(10) \qquad h_n^{(1)}(x) \equiv \int \frac{1}{b_n} k\left(\frac{x-u}{b_n}\right) d\,\Gamma_n(u) \,/\, \bar{\Gamma}_n(x)$$

$$\equiv f_n(x) \,/\, \bar{\Gamma}_n(x)$$

where $\{b_n\}$ is a band sequence satisfying initially

(b1) $\qquad b_n \to 0$, as $n \to$.

To analyze the asymptotic behavior of $h_n^{(1)}(x)$, it suffices to analyze that of $f_n(x)$. As mentioned earlier our technique is motivated by the strong representation result (Theorem 1) of Lo and Singh (1984). In Lemma 1 we shall show a modified version of their result. We begin with some notations. Let $L_1(t) = P(X_1 \le t, \delta_1 = 1)$. For positive real z and x, and δ taking values 0 or 1, let $\zeta(z,\delta,x) = -g(z\wedge x) + [\bar{L}(z)]^{-1} I(z<x \text{ and } \delta=1)$, where $g(y) = \int_0^y [\bar{L}(s)]^{-2} dL_1(s)$ and $I(\cdot)$ is the indicator function. Let $\zeta_i(x) = \zeta(X_i, \delta_i, x)$. Let T be any point with $L(T) < 1$. Note that the random variables $\zeta_i(x)$ are bounded, uniformly in $0<x<T$, $E\zeta_i(x) = 0$, and $\mathrm{Cov}(\zeta_i(x), \zeta_i(y)) = g(x\wedge y)$ (cf. Lo and Singh (1984)).

Lemma 1. Assuming that F is continuous, one can write

$$(11)\qquad \log \bar{\Gamma}_n(x) - \log \bar{F}(x) = -\frac{1}{n}\sum_{i=1}^{n} \zeta_i(x) + R_n(x), \text{ where}$$

$$(12)\qquad P\Big(\sup_{0<x<T} |R_n(x)| > a_n\Big) = O(n^{-\beta}),$$

for any $\beta > 0$ with $a_n = \theta\cdot[\log n/n]^{3/4}$ for some constant $\theta > 0$ depending on β.

Proof. The proof is given in the Appendix. ∎

Remark: Formula (12) is replaced by

$$(13)\qquad \sup_{0<x<T} |R_n(x)| \overset{a.s.}{=} O\big(n^{-3/4}(\log n)^{3/4}\big)$$

in Theorem 1 of Lo and Singh (1984).

It follows from Borel-Cantelli Lemma that (12) implies (13). Hence Lemma 1 is a stronger result than Theorem 1 of Lo and Singh (1984).

Let $\xi_i(x) = \bar{F}(x)\cdot\zeta_i(x)$.

Lemma 2. Assume that F is continuous, then

$$(14)\qquad \Gamma_n(x) - F(x) = \frac{1}{n}\sum_{i=1}^{n} \xi_i(x) + r_n(x), \quad \text{where}$$

$$(15) \quad \sup_{0<x<T} E|r_n(x)|^{\alpha} = O\{[\frac{\log n}{n}]^{(3/4)\alpha}\} \text{ for any } \alpha>1.$$

Proof. We shall only demonstrate the case $\alpha=1$. Since $\zeta_i(x)$'s are uniformly bounded and $(n+1)^{-1} < \bar{\Gamma}_n(x) < 1$ for all x in [0,T], we have $\sup_{0<x<T} |R_n(x)| = O(\log(n+1))$, and hence

$$(16) \quad \sup_{0<x<T} E|R_n(x)| = \sup_{0<x<T} E[|R_n(x)| \cdot I\{|R_n(x)| > a_n\}]$$

$$+ \sup_{0<x<T} E[|R_n(x)| \cdot I\{|R_n(x)| < a_n\}]$$

$$< \sup_{0<x<T} P\{|R_n(x)| > a_n\} \cdot O(\log(n+1)) + a_n$$

$$= O(a_n), \text{ by Lemma 1.}$$

Similarly, one can show that

$$(17) \quad \sup_{0<x<T} E(R_n(x)^2) = O(a_n^2).$$

Now by Taylor's expression,

$$- [\Gamma_n(x) - F(x)]$$

$$= \bar{\Gamma}_n(x) - \bar{F}(x)$$

$$= \exp\{\log \bar{\Gamma}_n(x)\} - \exp\{\log \bar{F}(x)\}$$

$$= [\log \bar{\Gamma}_n(x) - \log \bar{F}(x)] \cdot \bar{F}(x) + \Delta_n \cdot [\log \bar{\Gamma}_n(x) - \log \bar{F}(x)]^2$$

$$\equiv -\frac{1}{n} \sum_{i=1}^{n} \xi_i(x) + \bar{F}(x) \cdot R_n(x) + \Delta_n \cdot [\log \bar{\Gamma}_n(x) - \log \bar{F}(x)]^2,$$

where Δ_n is between $\overline{\Gamma}_n(x)$ and $\overline{F}(x)$ and is therefore bounded by one. It now follows from (16) and (17) that

$$\sup_{0<x<T} E|\log \overline{\Gamma}_n(x) - \log \overline{F}(x)|^2 = \sup_{0<x<T} E[\frac{1}{n} \Sigma\, \xi_i(x) + R_n(x)]^2$$

$$< \sup_{0<x<T} 2[E(\frac{1}{n} \Sigma\, \xi_i(x))^2 + E(R_n(x)^2)]$$

$$< \sup_{0<x<T} 2n^{-1} \mathrm{Var}\,(\xi_1(x)) + 0(a_n^2)$$

$$= 0(n^{-1}) + 0(a_n^2).$$

Hence

$$\sup_{0<x<T} E|r_n(x)| < \sup_{0<x<T} \overline{F}(x) \cdot E|R_n(x)| + \sup_{0<x<T} E|\log \overline{\Gamma}_n(x) - \log \overline{F}(x)|^2$$

$$= 0(a_n) + 0(n^{-1}) + 0(a_n^2)$$

$$= 0(a_n). \qquad \blacksquare$$

Finally, we state a lemma which by now is a standard device in the kernel estimation literature:

<u>Lemma 3</u>: Assume the kernel k is a bounded density. Let g be an integrable function.

(a) If $\{b_n\}$ is a sequence of positive numbers such that $b_n \to 0$ as $n \to \infty$, then

$$\lim_{n\to\infty} \int \frac{1}{b_n} k(\frac{x-u}{b_n})\, g(u)\, du = g(x) \tag{18}$$

for every continuity point x of g.

(b) If in addition k is symmetric with finite second moment, and g is twice continuously differentiable at x, then

$$(19) \qquad \int \frac{1}{b_n} k\left(\frac{x-u}{b_n}\right) g(u)\, du = g(x) + \frac{g''(x)}{2} \int v^2\, k(v) dv \cdot b_n^2 + o\left(b_n^2\right).$$

3. Strong Approximation of $f_n(x)$

Using the integration by parts lemma of Földes, Rejtö and Winter (1981) and Proposition 1, under the assumption that k is continuous on its support (condition (k3)), we have that if $x<T$, where $L(T)<1$,

$$(20) \qquad f_n(x) = \frac{1}{b_n} \int k\left(\frac{x-u}{b_n}\right) d\, \Gamma_n(u)$$

$$= \frac{1}{b_n} \int_{-c}^{c} \Gamma_n\left(x - vb_n\right)\, dk(v)$$

$$= \frac{1}{b_n} \int_{-c}^{c} \Big[F(x - vb_n) + \frac{1}{n} \sum_1^n \xi\left(X_i,\delta_i,x - vb_n\right) + r_n\left(x - vb_n\right)\Big]\, dk(v)$$

$$= \int_{-c}^{c} f\left(x - vb_n\right)\, k(v)dv + \frac{1}{nb_n} \sum_1^n \int_{-c}^{c} \xi\left(X_i,\delta_i,x - vb_n\right)dk(v)$$

$$+ \frac{1}{b_n} \int_{-c}^{c} r_n\left(x - vb_n\right)\, dk(v)$$

$$\equiv f(x) + \beta_n(x) + \sigma_n(x) + e_n(x),$$

where

$$(21)\qquad \beta_n(x) = \int_{-c}^{c} f(x - vb_n)k(v)\,dv - f(x)$$

is essentially the bias of $f_n(x)$;

$$(22)\qquad \sigma_n(x) = \frac{1}{nb_n}\sum_{1}^{n}\int_{-c}^{c} \xi(X_i,\delta_i,x - vb_n)\,dk(v)$$

is the random fluctuation component of $f_n(x)$, (we note that the integral is well-defined for n large enough because k is compactly supported), and

$$(23)\qquad e_n(x) = \frac{1}{b_n}\int_{-c}^{c} r_n(x - vb_n)\,dk(v)$$

is the error of the approximation. It is easily checked that

$$(24)\qquad \sup_{0<x<T} |e_n(x)| \overset{a.s.}{=} O\left((\log n/n)^{3/4} \cdot \frac{1}{b_n}\right)$$

by Lemma 2 and the fact that k is of bounded variation (condition (k4).) The process

$$(25)\qquad \bar{\xi}(t) = \frac{1}{\sqrt{n}}\sum_{1}^{n} \xi(X_i,\delta_i,t), \qquad 0<t<T,$$

has mean zero and covariance

$$(26)\qquad r(s,t) \equiv E[\bar{\xi}(s)\,\bar{\xi}(t)] = \bar{F}(s)\,\bar{F}(t)\int_{0}^{s\wedge t} [\bar{L}(u)]^{-2}\,d\,L_1(u),$$

where we recall $\bar{L}(t) = \bar{F}(t)\cdot\bar{G}(t)$, and $L_1(t) = P(X_i<t,\ \delta_i=1)$. One notes that this agrees with the covariance of the Kaplan-Meier process obtained by Breslow and Crowley (1974) and reduces to the usual covariance of the empirical process in the absence of censoring (see Hall and Wellner (1980)). $\sigma_n(t)$ is thus a process with mean zero and covariance

(27) $$E[\sigma_n(t)\,\sigma_n(s)] = \frac{1}{nb_n^2}\int_{-c}^{c}\int_{-c}^{c}\gamma(t-ub_n,\ s-vb_n)\,dk(u)\,dk(v).$$

We now summarize our findings in the following

Proposition 1: Suppose F is absolutely continuous with density $f(x) > 0$ at x. Suppose k is of bounded variation and is continuous. Then $f_n(x)$ admits the strong approximation on the interval [0,T]:

(28) $$f_n(x) = f(x) + \beta_n(x) + \sigma_n(x) + e_n(x),$$

where $\beta_n(x)$, $\sigma_n(x)$, $e_n(x)$ are defined in (21), (22) and (23) respectively, and e_n satisfies (24).

In view of previous lemmas and the above proposition, we have the following consequences:

Corollary 1: (Strong pointwise consistency.) Suppose k satisfies (k1) - (k4), $\{b_n\}$ satisfies (b1), and additionally,

(b2) $(n/\log\log n)^{1/2}\cdot b_n \to \infty$ as $n \to \infty$;

f(x) exists and is continuous at x. Then $f_n(x) \to f(x)$ a.s. as $n\to\infty$.

Corollary 2: (Bias and variance.) Suppose k satisfies (k1) - (k4), $f(x) > 0$, and that f is twice continuously differentiable at x, then

(29) $$E\,f_n(x) = f(x) + \frac{f''(x)}{2}\int_{-c}^{c} v^2\,k(v)\,dv\cdot b_n^2 + o(b_n^2) + O(b_n^{-1}a_n).$$

(30) $$\mathrm{Var}\,f_n(x) = (nb_n)^{-1}\frac{f(x)}{\bar{G}(x)}\int_{-c}^{c}k^2(v)dv + O(n^{-1}) + O((a_n/b_n))^2$$
$$+\,O\left(\frac{a_n}{b_n}(nb_n)^{-1/2}\right).$$

Proof: We shall demonstrate (30) only. Consider first

$$\text{Var } \sigma_n(x) = (nb_n^2)^{-1} \int_{-c}^{c} \int_{-c}^{c} \bar{F}(x-ub_n)\bar{F}(x-vb_n)\cdot$$

$$g[(x-ub_n)\wedge(x-vb_n)]dk(u)dk(v),$$

where $g(y) = \int_0^y [\bar{L}(t)]^{-2} dL_1(t)$ has Lebesque derivative

$$(31) \quad \frac{dg(t)}{dt} = \frac{dL_1(t)}{dt}/[\bar{L}(t)]^2 = f(t)/[\bar{G}(t)\cdot\bar{F}(t)^2].$$

Since k is symmetric, a two-term Taylor expansion argument yields

$$(32) \quad \text{Var } \sigma_n(x) = \sigma_n^*(x) + 0(n^{-1}),$$

where

$$\sigma_n^* = \frac{\bar{F}(x)^2}{nb_n^2} \int_{-c}^{c} \int_{-c}^{c} g[(x-ub_n)\wedge(x-vb_n)]dk(u)dk(v).$$

Using integration by parts, we have

$$\int_{-c}^{c} g[(x-ub_n)\wedge(x-vb_n)]dk(u) = \int_{x-cb_n}^{x-vb_n} k(\frac{x-w}{b_n})dg(w)$$

Thus by Fubini Theorem and a change of variable, we otbain

$$\text{Var } \sigma_n^*(x) = \frac{\bar{F}(x)^2}{nb_n^2} \int_{x-cb_n}^{x+cb_n} k^2(\frac{x-w}{b_n})dg(w)$$

$$= \frac{\bar{F}(x)^2}{nb_n^2} \int_{x-cb_n}^{x+cb_n} k^2(\frac{x-w}{b_n})\cdot \frac{f(w)}{\bar{G}(w)\bar{F}(w)^2}\, dw,$$

where the last equality follows by (31). Finally, another change of variable and expansions applied to $\frac{f}{\bar{G}\,\bar{F}^2}$ lead to the following approximation

$$(33)\quad \mathrm{Var}\ \sigma_n^*(x) = \frac{f(x)}{\bar{G}(x)} \int_{-c}^{c} k^2(v)dv \cdot \frac{1}{nb_n} + O(n^{-1}).$$

Next, observe that by Lemma 2, for n large enough, since $b_n \to 0$ as $n \to \infty$, we have

$$(34)\quad \mathrm{Var}\ e_n(x) \le E\left(e_n^2(x)\right)$$

$$= \frac{1}{b_n^2} \int_{-c}^{c} \int_{-c}^{c} E[r_n(x-ub_n)r_n(x-vb_n)]dk(u)dk(v)$$

$$= O\left((a_n/b_n)^2\right).$$

Thus (30) follows by applying (32), (33), (34) and Schwartz's inequality to an expansion of Var $f_n(x)$ via (28). ∎

<u>Corollary 3:</u> (Asymptotic normality.) Suppose k satisfies (k1) -(k4), $b_n = o(n^{-1/5})$, and

$$(b3)\quad \frac{n^{1/2}}{(\log n)^{3/2}} \cdot b_n \to \infty \text{ as } n \to \infty.$$

Then

$$\sqrt{nb_n}\,[f_n(x) - f(x)] \xrightarrow{d} N\left(0, \frac{f(x)}{\bar{G}(x)} \int_{-c}^{c} k^2(v)dv\right)$$

as $n\to\infty$. Here $\xrightarrow{d}$ means convergence in distribution.

Remark: Putting $b_n = O(n^{-\alpha})$, the conditions in Corollary 3 say $1/5 < \alpha < 1/2$.

4. Kernel estimation of the hazard rate

We begin by stating the strong consistency of $h_n^{(1)}(x)$:

Theorem 1. Let k satisfy (k1) - (k4), $\{b_n\}$ satisfies (b1)(b2).

(a) If f is continuous, then $h_n^{(1)}(x) \to h(x)$ a.s. as $n \to \infty$.

(b) If f is uniformly continuous, then for any T with $L(T) < 1$, $h_n^{(1)}(x) \to h(x)$ uniformly a.s. on $[o,T]$ as $n \to \infty$.

Proof. Since $\Gamma_n^{(x)}$ estimates $F(x)$ uniformly a.s., the pointwise result (a) is a direct consequence of Corollary 1. For (b), the proof is also standard, noting that by Csörgö and Horváth (1983),

$$\sup_{0<x<T} |KM_n(x) - F(x)| \overset{a.s.}{=} O(n^{-1/2} \log \log n). \qquad \blacksquare$$

In order to establish the asymptotic normality of $h_n^{(1)}(x)$, write

$$\sqrt{nb_n}\,[h_n^{(1)}(x) - h(x)] = \sqrt{nb_n}\,\{f_n(x)\,[\frac{1}{\bar{\Gamma}_n(x)} - \frac{1}{\bar{F}(x)}]$$

$$+ [f_n(x) - f(x)] / \bar{F}(x)\}.$$

It suffices to show the first term on the right converges to zero in probability. Now

$$\text{(35)} \qquad \sqrt{nb_n}\ \{f_n(x)\,[\frac{\bar{F}(x) - \bar{\Gamma}_n(x)}{\bar{\Gamma}_n(x)\,\bar{F}(x)}]\} = \sqrt{n}\,[\bar{F}(x) - \bar{\Gamma}_n(x)] \cdot$$

$$\cdot \sqrt{b_n}\, f_n(x)[\bar{\Gamma}_n(x)\,\bar{F}(x)]^{-1}.$$

Since $\sqrt{n}[\overline{\Gamma}_n(x) - \overline{F}(x)]$ tends in distribution to a normal random variable, $f_n(x)$ converges to $f(x)$ a.s. by Corollary 1, and clearly $[\overline{\Gamma}_n(x)\ \overline{F}(x)]^{-1}$ converges to $[\overline{F}(x)]^{-2}$ a.s., we have by Slutsky's Theorem that the expression in (35) tends to zero in probability. To summarize, we have

<u>Theorem 2</u>: Suppose F is absolutely continuous with density $f(x) > 0$, suppose the kernel k satisfies (k1)-(k4), and suppose the band sequence $\{b_n\}$ satisfies (b3) and additionally that $b_n = o(n^{-1/5})$. Then we have

(36) $$\sqrt{nb_n}\,[h_n^{(1)}(x) - h(x)] \xrightarrow{d} N\left(0, \frac{h(x)}{\overline{L}(x)} \int_{-c}^{c} k^2(v)dv\right)$$

as $n \to \infty$.

<u>Remark 1</u>.

Tanner and Wong (1983) tackled the asymptotic normality question by Hajek's projection. Their centering constant is $E\, h_n^{(2)}(x)$, thus bypassing the bias issue. They also imposed a compatibility condition on the kernel k with respect to both F and G. Such a condition is met by kernels satisfying (k1)-(k4).

We now turn our attention to the study of MSE of $h_n^{(1)}(x)$. Write

(37) $$E[h_n^{(1)}(x) - h(x)]^2 = E[I + II + III]^2,$$

where

$$I = f_n(x)\left[\frac{1}{\overline{\Gamma}_n(x)} - \frac{1}{\overline{F}(x)}\right],$$

$$II = [f_n(x) - E\, f_n(x)] / \overline{F}(x),$$

$$III = [E\, f_n(x) - f(x)] / \overline{F}(x).$$

We will show that the main contribution comes from $E(II^2)$ and $E(III^2)$, all other terms in the quadratic expansion being of smaller order. Note also that III is deterministic. Now

$$(38)\quad E(II^2) = [\bar{F}(x)]^{-2} \cdot \operatorname{Var} f_n(x)$$

$$= \frac{h(x)}{\bar{L}(x)} \int_{-c}^{c} k^2(v)dv \cdot \frac{1}{nb_n} + O(n^{-1}) + O[(\frac{a_n}{b_n})]^2 + O(\frac{a_n}{b_n} \cdot (nb_n)^{-1/2}).$$

$$(39)\quad E(III^2) = [\bar{F}(x)]^{-2}\, \beta_n^2(x) + O([\,\frac{a_n}{b_n}\,]^2) + \beta_n O(\frac{a_n}{b_n})$$

$$= [\frac{f''(x)}{2\bar{F}(x)} \int_{-c}^{c} v^2 k(v)dv]^2 \cdot b_n^4 + o(b_n^4) + O([\,\frac{a_n}{b_n}\,]^2) + O(a_n b_n).$$

To evaluate $E(I)$, let us first consider

$$(40)\quad E[\,\frac{f_n(x)}{\bar{\Gamma}_n(x)}] = E\{\,\frac{f_n(x)}{\bar{F}(x)}\,[1 + \frac{\bar{\Gamma}_n(x) - \bar{F}(x)}{\bar{F}(x)}\,]^{-1}\,\}$$

$$= E\,[\frac{f_n(x)}{\bar{F}(x)}] - E\{\,\frac{f_n(x)}{\bar{F}(x)}\,[\frac{\bar{\Gamma}_n(x) - \bar{F}(x)}{\bar{F}(x)}] \cdot (1 + \varepsilon_n)^{-2}\,\}$$

for some ε_n between 0 and $[\bar{\Gamma}_n(x) - \bar{F}(x)]/\bar{F}(x)$ by Taylor's expansion for large enough n since $\bar{\Gamma}_n(x) > \frac{1}{n+1}$ for all x. Since $\|f_n(x)\|_\infty < M$ for some $0<M<\infty$ by (k2) and (k3), we have that

$$(41)\quad E|I| = E\left|\frac{f_n(x)}{\bar{F}(x)^2}\left[\frac{\bar{\Gamma}_n(x) - \bar{F}(x)}{(1+\varepsilon_n)^2}\right]\right| < \frac{M}{\bar{F}(x)^2}\cdot E\left|\frac{\bar{\Gamma}_n(x) - \bar{F}(x)}{(1+\varepsilon_n)^2}\right|$$

$$< \frac{M}{\bar{F}(x)^2}\{E[\bar{\Gamma}_n(x) - \bar{F}(x)]^2\}^{1/2}\{E[1+\varepsilon_n]^{-4}\}^{1/2}$$

by Schwarz inequality. Now

$$(42)\qquad 0 < E[1+\varepsilon_n]^{-4} < E\left[\frac{\bar{F}(x)}{\bar{\Gamma}_n(x)}\right]^4 + 1$$

$$= E\left[1 - \frac{\bar{\Gamma}_n(x) - \bar{F}(x)}{\bar{\Gamma}_n(x)}\right]^4 + 1,$$

and

$$(43)\qquad E\left[\frac{\bar{\Gamma}_n(x) - \bar{F}(x)}{\bar{\Gamma}_n(x)}\right]^4 < \left[\frac{d_n}{\bar{F}(x) - d_n}\right]^4 \cdot P(|\bar{\Gamma}_n(x) - \bar{F}(x)| < d_n)$$

$$+ \left[\frac{2}{\frac{1}{n+1}}\right]^4 \cdot P(|\bar{\Gamma}_n(x) - \bar{F}(x)| > d_n)$$

$$< O[d_n^4 + n^4 \cdot n^{-5}]$$

$$= O(n^{-1}),$$

where the last inequality follows from Lemma 2 and the exponential bound in Lemma 1 of Lo and Singh (1984) and $d_n = \tau \cdot (\log n/n)^{1/2}$ for some $\tau>0$. Hence from Holder's inequality, we have

$$(44)\qquad E[1+\varepsilon_n]^{-4} = O(1).$$

Apply Lemma 2 once more, one can show that

(45) $E[\bar{\Gamma}_n(x) - \bar{F}(x)]^2 = O(n^{-1})$.

It now follows from (41), (44) and (45) that $E|I| = O(n^{-1/2})$. The term $E(I^2)$ can be shown in a similar fashion to be of order $O(n^{-1})$. Hence from (29) and (38),

(46) $E|I\cdot III| = |III|\cdot E|I| = O(n^{-1/2}b_n^2) + O(n^{-1/2}(a_n/b_n))$

(47) $E|I\cdot II| = O(n^{-1/2}\cdot (nb_n)^{-1/2}) + O(n^{-1/2}\cdot (a_n/b_n))$

$$+ O(n^{-1/2}\cdot (a_n/b_n)^{1/2}\cdot(nb_n)^{-1/4}).$$

Let b_n be of the form cn^{-p}, where c,p are both positive constants. For $0<p<1/4$, $O(n^{-1/2}b_n^2)$ and $O(n^{-1/2}(nb_n)^{-1/2})$ are the dominating terms in $E|I\cdot III|$ and $E|I\cdot II|$ respectively.

For $1/4 < p < 1/2$, $O(n^{-1/2}(a_n/b_n))$ and $O(n^{-1/2}(a_n/b_n)^{1/2}(nb_n)^{-1/4})$ are the dominating terms in $E|I\cdot III|$ and $E|I\cdot II|$ respectively.

For $p > 1/2$, $O(n^{-1/2}(a_n/b_n))$ is the dominating term in both $E|I\cdot III|$ and $E|I\cdot II|$.

Since b_n^4 dominates $n^{-1/2}b_n^2$ for $p < 1/4$, and $(nb_n)^{-1}$ dominates $n^{-1/2}b_n^2$ for $p > 1/6$, the term $O(n^{-1/2}b_n^2)$ is always dominated by either b_n^4 or $(nb_n)^{-1}$ for any $p > 0$.

Also, $(nb_n)^{-1}$ always dominate $n^{-1/2}(nb_n)^{-1/2}$, $n^{-1/2}(a_n/b_n)$ and $n^{-1/2}(a_n/b_n)^{1/2}(nb_n)^{-1/4}$ for any $p > 0$.

$E(II^2)$ and $E(III^2)$ will be the main contribution to the MSE of $h_n^{(1)}(x)$. Futhermore, if $\{b_n\}$ satisfies (b1) and (b2), $(nb_n)^{-1/2}$ will dominate (a_n/b_n). Also either b_n^4 or $(nb_n)^{-1}$ will dominate a_nb_n. We now state our finding:

<u>Theorem 3</u>: Suppose f is twice continuously differentiable at x, $f(x) > 0$, the kernel k satisfies (k1)-(k4), and the band sequence $\{b_n\}$ satisfies (b1) and (b2). Then

$$(48) \quad MSE[h_n^{(1)}(x)] = [\frac{f''(x)}{2\bar{F}(x)} \int_{-c}^{c} v^2 k(v)dv]^2 \cdot b_n^4 + [\frac{h(x)}{\bar{L}(x)} \int_{-c}^{c} k^2(v)dv] \cdot \frac{1}{nb_n}$$

$$+ o(b_n^4 + \frac{1}{nb_n}).$$

5. <u>Concluding Comments</u>

(a) We have seen in the above discussion the use of Lo and Singh's (1984) strong represnetation of the Kaplan-Meier estimator in analyzing kernel estimation of hazard rate functions. We have chosen to consider the estimates given by $h_n^{(1)}(x)$ as contrasted with $h_n^{(2)}(x)$ studied by Yandell (1983). Our variance expression and asymptotic normality results are similar to theirs, although we have employed a more traditional approach. The bias for the three variants appear to be different in the scale constant but not the rate.

(b) Tanner (1983) mentioned that a nearest-neighbor approach may be preferable to the fixed band sequence approach from an extensive simulation experiment. This observation appears to have some theoretical support judging from the recent work of Liu and Van Ryzin (1985) which essentially used an asymmetric nearest-neighbor window. Both their findings (Theorems 4.3 and 4.4) and the findings of some other researchers on nearest neighbor density estimation with censored data (for instance Mielniczuk (1984)) suggest that the censoring mechanism may have no effect on the variance for nearest-neighbor estimates. This may be an advantage in terms of constructing a confidence interval at a fixed point or a simultaneous confidence band if one wants to test for goodness-of-fit. Nevertheless, one cautions that the bias behavior of the Liu and Van Ryzin variable histogram estimator suffers essentially the same drawback as nearest-neighbor estimators in that it may be quite large at the tail regions of F.

(c) A number of researchers in kernel estimation have studied the effects of kernels which may have vanishing moments. Its use, coupled with the assumption of a higher degree of smoothness of h(x), can make the convergence of the bias to zero faster. This point of view was taken in Singpurwalla and Wong (1983). Of course one pays the price that the estimator so constructed may take on negative values if the sample size is not "large enough." For this reason we have kept the non-negativity of the kernel in this paper.

6. Appendix:

Proof of Lemma 1.

Let $L_n(t) = n^{-1} \sum_{i=1}^{n} I(x_i \le t)$ and

$L_{1n}(t) = n^{-1} \sum_{i=1}^{n} I(x_i \le t, \delta_i = 1)$ be the empirical distribution and subdistribution function respectively. If one checks the proof of Theorem 1 of Lo and Singh (1984) carefully, one will find that $R_n(x)$ is composed of three terms,

$R_n(x) = R_{n1}(x) + R_{n2}(x) + R_{n3}(x)$, where

$$R_{n1}(x) = \log\bar{F}_n(x) + \int_0^x [\bar{L}_n(s)]^{-1} d\, L_{1n}(s),$$

$$R_{n2}(x) = \int_0^x ([\bar{L}(s)]^{-1} - [\bar{L}_n(s)]^{-1})\, d\, L_1(s) +$$

$$\int_0^x ([\bar{L}(s)]^{-2} [\bar{L}(s) - \bar{L}_n(s)]\, d\, L_1(s)$$

$$= - \int_0^x \frac{[\bar{L}_n(s) - \bar{L}(s)]^2}{[\bar{L}(s)]^2\, \bar{L}_n(s)}\, d\, L_1(s),$$

$$R_{n3}(x) = \int_0^x ([\bar{L}(s)]^{-1} - [\bar{L}_n(s)]^{-1})\, d\, (L_{1n}(s) - L_1(s)).$$

Lemma 1 then follows from Lemmas 4, 5 and 6 below.

To prove Lemmas 4, 5, 6, let $q = P(X_1 \le T, \delta_1 = 1) < 1$, $\varepsilon_0 = P(X_1 > T) = \bar{L}(T) > 0$.

Lemma 4. For any $a>0$, and $0<b<1$, we have

$$P\{\sup_{0<x<T} |R_{n1}(x)| > an^{-b}\} < 2e^{-n(1-q)/36}$$

for large n.

Proof.

$$|R_{n1}(x)| < \Sigma^* |\log(\frac{n-i+1}{n-i+2}) + \frac{1}{n} \cdot \frac{1}{\bar{L}_n(X_{(i)})}|,$$

where Σ^* sums over all i such that $X_{(i)}<t$ and $\delta_{(i)} = 1$,

$$= \Sigma^* |\log(1 - \frac{1}{n-i+2}) + \frac{1}{n-i}|$$

$$= \Sigma^* |\log(1 - \frac{1}{n-i+2}) + \frac{1}{n-i+2}| + \Sigma^* (\frac{1}{n-i} - \frac{1}{n-i+2})$$

$$< \Sigma^* \frac{(n-i+2)^{-2}}{1-(n-i+2)^{-1}} + \Sigma^* (\frac{1}{n-i} - \frac{1}{n-i+2}), \text{ since}$$

$$-\frac{x^2}{1-x} < \log(1-x) + x < 0 \text{ for } 0<x<1,$$

$$= \Sigma^* (\frac{1}{n-i+1} - \frac{1}{n-i+2}) + \Sigma^* (\frac{1}{n-i} - \frac{1}{n-i+2})$$

$$< \sum_{i=1}^{K} (\frac{1}{n-i+1} - \frac{1}{n-i+2}) + \sum_{i=1}^{K} (\frac{1}{n-i} - \frac{1}{n-i+2}), \text{ where}$$

$$K = \sum_{i=1}^{n} I(X_i < T, \delta_i = 1)$$

$$= (\frac{1}{n-K+1} - \frac{1}{n+1}) + (\frac{1}{n-K} + \frac{1}{n-K+1} - \frac{1}{n} - \frac{1}{n+1})$$

$$= \frac{2K}{(n+1)(n-K+1)} + \frac{K}{n(n-K)}$$

$$< \frac{3K}{n(n-K)}.$$

$$P\left(\frac{3K}{n(n-K)} > an^{-b}\right) = P\left(K > \frac{an^{2-b}}{3 + an^{1-b}}\right)$$

$$< P(K > (1+q)n/2)), \text{ for large } n$$

$$= P(K - nq > (1-q)n/2)$$

$$< 2e^{-n(1-q)/36},$$

where the last inequality follows from Lemma 1 of Lo and Singh (1984) by letting $\eta_i = I(X_i < T, \delta_i = 1)-q$, $c=1$, $d=n(1-q)/6$, $\sigma^2 = q(1-q)$ and $z = d/6 = n(1-q)/36$

Lemma 4 now follows immediately. ∎

<u>Lemma 5</u>. For any $\varepsilon > 0$, $P\{\sup_{o<x<T} |R_{n2}(x)|>\varepsilon\{ < a'e^{-nb'\varepsilon}$ for some positive constants a' and b'.

<u>Proof</u>. Note that $0 < R_{n2}(x) < R_{n2}(T)$. We have

$$|R_{n2}(T)| < \varepsilon_o^{-2} [\bar{L}_n(T)]^{-1} \|\bar{L}_n-\bar{L}\|_T^2,$$ where $\|\cdot\|_T$ is the sup-norm of a function over the interval [0,T].

Hence, $P(|R_{n2}(T)| > \varepsilon)$

$$< P\{[\bar{L}_n(T)]^{-1} \cdot \|\bar{L}_n-\bar{L}\|_T^2 > \varepsilon\, \varepsilon_o^2 ,\ \bar{L}_n(T) > \varepsilon_o/2\} + P\{\bar{L}_n(T) < \varepsilon_o/2\}$$

$$< P\{\|\bar{L}_n-\bar{L}\|_T^2 > \varepsilon\, \varepsilon_o^3/2\} + P\{\bar{L}_n(T) < \varepsilon_o/2\}$$

$$= I + II.$$

Lemma 2 of Dvoretzky, Kiefer and Wolfowitz (1956) implies that

$$I < \text{constant} \cdot e^{-n\varepsilon\varepsilon_0^3}, \text{ and}$$

$$II = P\{\bar{L}_n(T) - \varepsilon_0 < -\varepsilon_0/2\}$$

$$< \text{constant} \cdot e^{-n\varepsilon_0^2/2}.$$

Lemma 5 is thus proved. ■

Lemma 6. If F is continuous, for any $\beta > 0$, there exists constant $\eta > 0$ such that

$$P\left(\sup_{0<x<T} |R_{n3}(x)| > \eta \cdot (\log n/n)^{3/4}\right) = 0(n^{-\beta}).$$

Proof. We shall give the proof for the case when G is also assumed to be continuous, and hence L is continuous.

The proof parallels those of Lemma 2 of Lo and Singh (1984) with more rigorous probability statements. The proof when G is arbitrary can be done similarly as the remark on page 10 of Lo and Singh (1984). We shall now proceed with the proof when both F and G are assumed to be continuous.

Divide the interval $[0,T]$ into subintervals $[x_i, x_{i+1}]$, $i=0,\ldots,k_n$, where $k_n = o(n/\log n)^{1/2})$ and $0 = x_0 < x_1 < \ldots < x_{k_n+1} = T$ are such that $L(x_{i+1}) - L(x_i) < c_1 \cdot (\log n/n)^{1/2}$. This is possible because L is assumed to be continuous. From any $0<x<T$, we have

$$|R_{n3}(x)| < k_n \cdot \sup_{0<t<T} |[\bar{L}_n(t)]^{-1} - [\bar{L}(t)]^{-1}| \cdot \max_{1<i<k_n} |(\bar{L}_{1n}-\bar{L}_1)(x_{i+1}) -$$

$$(\bar{L}_{1n}-\bar{L}_1)(x_i)| + 2 \max_{0<i<k_n} \sup_{x_i<t<x_{i+1}} |[\bar{L}_n(t)]^{-1} - [\bar{L}_n(x_i)]^{-1}$$

$$- [\bar{L}(t)]^{-1} + [\bar{L}(x_i)]^{-1}|$$

$$= A + B,$$

from the proof of Lemma 2 of Lo and Singh (1984).

To estimate B, we further subdivide each $[x_i, x_{i+1}]$ into subintervals $[x_{ij}, x_{i(j+1)}]$, $j=1,\dots,m_n$ such that $L(x_{i(j+1)}) - L(x_{ij}) < c_2 \cdot (\log n/n)^{3/4}$, for all i, j and $m_n = O((n/\log n)^{1/4})$. Consider

$$\Delta_i = \sup_{x_i<t<x_{i+1}} |[\bar{L}_n(t)]^{-1} - [\bar{L}_n(x_i)]^{-1} - [\bar{L}(t)]^{-1} + [\bar{L}(x_i)]^{-1}|$$

$$< \sup_{x_i<t<x_{i+1}} |[\bar{L}_n(t) - \bar{L}(t)][\bar{L}(t)]^{-2} - [\bar{L}_n(x_i) - \bar{L}(x_i)][\bar{L}(x_i)]^{-2}|$$

$$+ 2\|\bar{L}_n - \bar{L}\|_T \cdot \|(\bar{L}_n \bar{L})^{-1} - (\bar{L})^{-2}\|_T$$

$$< \sup_{x_i<t<x_{i+1}} |\bar{L}(x_i)]^{-2}|\bar{L}_n(t) - \bar{L}(t) - \bar{L}_n(x_i) + \bar{L}(x_i)|$$

$$+ \sup_{x_i<t<x_{i+1}} |\bar{L}_n(t) - \bar{L}(t)|\cdot|\bar{L}^{-2}(t) - \bar{L}^{-2}(x_i)| + 2\|\bar{L}_n - \bar{L}\|_T^2 \cdot \varepsilon_o^{-2} [\bar{L}_n(T)]^{-1}$$

$$< \max_{1<j<m_n} \varepsilon_o^{-2} |\bar{L}_n(x_{ij}) - \bar{L}_n(x_{ij}) - \bar{L}_n(x_i) + \bar{L}(x_i)| + c_2 (\log n/n)^{3/4}$$

$$+ 2\varepsilon_o^{-4} \cdot \|\bar{L}_n - \bar{L}\|_T \cdot c_1(\log n/n)^{1/2} + 2\varepsilon_o^{-2}\|\bar{L}_n - \bar{L}\|_T^2 [\bar{L}_n(T)]^{-1}.$$

We have

$$P\{\Delta_i > [c_2 + 9\varepsilon_o^{-2}(c_1\beta)^{1/2}] \cdot (\log n/n/)^{3/4}\}$$

$$< P\{\max_{1<j<m_n} |\bar{L}_n(x_{ij}) - \bar{L}(x_{ij}) - \bar{L}_n(x_i) + \bar{L}(x_i)| > 3(c_1\beta)^{1/2}(\log n/n)^{3/4}\}$$

$$+ P\{\|\bar{L}_n - \bar{L}\|_T > (3/2)\, \varepsilon_o^2\, (\beta/c_1)^{1/2} \cdot (\log n/n)^{1/4}$$

$$+ P\{\|\bar{L}_n - \bar{L}\|_T^2[\bar{L}_n(T)]^{-1} > (3/2)(c_1\beta)^{1/2}(\log n/n)^{3/4}\}$$

$$= I + II + III.$$

From the proof of Lemma 5, III $< a'e^{-b'n^{1/4}}$, for some positive constants a' and b'. From Lemma 2 of Dvoretzky, Kiefer and Wolfowitz (1956), II $<$ $a^* \cdot e^{-b^* n^{1/2}}$, for some positive constants a^* and b^*. As for I, for any fixed j, use Lemma 1 of Lo and Singh (1984) with $\eta_i = \bar{L}_n(x_{ij}) - \bar{L}(x_{ij}) - \bar{L}_n(x_i) + \bar{L}(x_i)$, c=1, $\sigma^2 < c_1 \cdot (\log n/n)^{1/2}$, $z = \beta \log n$, $d = (c_1\beta)^{1/2} n^{1/4} (\log n)^{3/4}$. We have $cz < d$ for large n, and $nz\sigma^2 = d^2$.

Bonferoni inequality then implies I $< 2m_n\, e^{-\beta \log n} = 2m_n\, n^{-\beta}$.

So far we have shown that, for any positive β there exists a positive constant ω such that

$$P\{\Delta_i > \omega \cdot (\log n/n)^{3/4}) < 2m_n\, n^{-\beta} + a'e^{-b'n^{1/4}} + a^*e^{-b^*n^{1/2}},$$

$$= 2m_n\, n^{-\beta} + O(n^{-\beta}).$$

Applying Bonferoni inequality once more, we have

$$P\{\beta > \omega \cdot (\log n/n)^{3/4}) < (k_n + 1)m_n \cdot O(n^{-\beta}), \text{ for } \beta > 1,$$

$$= O(n^{-\beta}), \text{ for } \beta > 0.$$

To estimate A, use the fact that $|\bar{L}_1(x) - \bar{L}_1(y)| < |\bar{L}(x) - \bar{L}(y)|$ for any x and y. Apply Lemma 1 of Lo and Singh (1984) again as we did for the term I above, and we have

$$P(\Omega_i) = P\{\max_{1<i<k_n} |(\bar{L}_{1n} - \bar{L})(x_{i+1}) - (\bar{L}_{1n} - \bar{L}_1)(x_i)| \text{ constant} \cdot (\log n/n)^{3/4}\}$$

$$= k_n \cdot O(n^{-\beta}).$$

Hence

$$P(A > \text{constant} \cdot (\log n/n)^{3/4})$$

$$= P\{k_n \cdot \|(\bar{L}_n)^{-1} - (\bar{L})^{-1}\|_T \cdot \Omega_1 > \text{constant} \cdot (\log n/n)^{3/4}\}$$

$$\leq P\{k_n \cdot \|\bar{L}_n - \bar{L}\|_T \cdot [\bar{L}_n(T)]^{-1} \cdot \Omega_1 > \text{constant} \cdot (\log n/n)^{3/4}\}$$

$$\leq P\{[\bar{L}_n(T)]^{-1} \cdot \Omega_1 > \text{constant} \cdot (\log n/n)^{3/4}\} +$$

$$P\{k_n \cdot \|(\bar{L}_n) - (\bar{L})\|_T > \beta^{1/2}\}$$

$$= P\{\Omega_1 > \text{constant} \cdot (\log n/n)^{3/4}\} + P\{\bar{L}_n(T) < (\varepsilon_o/2)\}$$

$$+ P\{k_n \cdot \|\bar{L}_n - \bar{L}\|T > \beta^{1/2}\}$$

$$= k_n \cdot O(n^{-\beta}) + \text{constant} \cdot e^{-n\varepsilon_o^2/2} + \text{constant} \cdot e^{-2\beta \log n}$$

for arbitrary $\beta > 1$, where the second term was computed in Lemma 5 and the third term comes from Lemma 2 of Dvoretzky, Kiefer and Wolfowitz (1956).

$$= O(n^{-\beta}) \text{ for arbitrary } \beta > 0.$$

We have thus shown Lemma 6. ∎

References

Aalen, O. (1978). "Nonparametric inference for a family of counting processes." A.S. 6, 701-726.

Ahmad, I. A. (1976). "Uniform strong convergence of a generalized failure rate estimate." Bull. Math. Statist. 17, 77-84.

Ahmad, I. A. and Lin, P. E. (1977). "Nonparametric estimation of a vector-valued bivariate failure rate." A.S. 5, 1027-1038.

Barlow, R. E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing: Probability Models. Holt, Rinehart and Winston, New York.

Barlow, R. E. and van Zwet, W. (1971). "Comparison of several nonparametric estimators of the failure rate function." Oper. Res. Reliab. (Crouchko, ed.), 375-399.

Blum, J. R. and Susarla, V. (1980). "Maximal deviation theory of density and failure rate function estimates based on censored data." Mult. Analysis V, Krishnaiah, P. R., ed., 213-222.

Breslow, N. and Crowley, J. (1974). "A large sample study of the life table and product-limit estimates under random censorship." A.S. 2, 437-443.

Burke, M. D. (1983). "Approximations of some hazard rate estimators in a competing risks model." Stoch. Proc. Appl. 14, 157-174.

Burke, M. D. and Horváth, L. (1984). "Density and failure rate estimation in a competing risks model." Sankhyā Ser A. 46, 135-154.

Csörgö, S. and Horváth, L. (1984). "The rate of strong uniform consistency for the product-limit estimator." Z. Wahr. verw. Geb. 62, 411-426.

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). "Asymptotic minimiax character of the sample distribution function and of the classical multinomial estimator." Ann. Math. Statist. 27, 642-669.

Földes, A., Rejtö, L. and Winter, B. B. (1981). "Strong consistency properties of nonparametric estimators for randomly censored data. II: Estimation of density and failure rate." Period. Math. Hungar. 12, 15-29.

Gill, R. (1983). "Large sample behavior of the product-limit estimator on the whole line." A.S. 11, 49-58.

Hajek, J. (1968). "Asymptotic normality of simple linear statistics under alternatives. Ann. Math. Statist. 39, 325-346.

Hall, W. J. and Wellner, J. A. (1980). "Confidence bands for a survival curve from censored data." Biometrika 67, 133-143.

Kaplan, E. L. and Meier, P. (1958). "Nonparametric estimation from incomplete observations." JASA 53, 457-481.

Komlos, J., Major, P., and Tusnady, G. (1975). "An approximation of partial sums of independent RV's and the sample DF. I." Z. Wahr. verw. Gebiete 32, 111-131.

Liu, R.Y.C. and Van Ryzin, J. (1985). "A histogram estimator of the hazard rate with censored data." A.S. 13, 592-605.

Lo, S.-H. and Singh, K. (1984). "The product-limit estimator and the bootstrap: Some asymptotic representations." To appear in Z. Wahr. verw. Gebiete.

Mack, Y. P. and Rosenblatt, M. (1979). "Multivariate k-nearest neighbor density estimate." J. Mult. Analysis 9, 1-15.

Mielniczuk, J. (1984). "Some asymptotic properties of kernel estimators of a density function in case of censored data." Manuscript.

Padgett, W. J. and McNichols, D. T. (1984). "Nonparametric density estimation from censored data." Commun. Statist. - Theor. Meth. A13, 1581-1613.

Ramlau-Hansen, H. (1983). "Smoothing counting process intensities by means of kernel functions." A.S. 11, 453-466.

Rice, J. and Rosenblatt, M. (1976). "Estimation of the log survivor function and hazard function." Sankhyā Ser. A, 38, 60-78.

Rosenblatt, M. (1956). "Remarks on some nonparametric estimates of a density function." Ann. Math. Statist. 27, 832-837.

Singpurwalla, N. D. and Wong, M.-Y. (1983). "Kernel estimators of the failure rate function and density function: An analogy." JASA 78, 478-481.

Tanner, M. A. (1983). "A note on the variable kernel estimator of the hazard function from randomly censored data." A.S. 11, 994-998

Tanner, M. A. and Wong, W. H. (1983). "The estimation of the hazard function from randomly censored data by the kernel method." A.S. 11, 989-993.

Walter, G. and Blum, J. R. (1979). "Probability density estimation using delta sequences." A.S. 7, 328-340.

Watson, G. S. and Leadbetter, M. R. (1964a). "Hazard analysis. I." Biometrika 51, 175-184.

____________________________. (1964b). "Hazard analysis. II." Sankhyā, Ser. A 26, 101-116.

Yandell, B. S. (1983). "Nonparametric inference for rates and densities with censored survival data." A.S. 11, 1119-1135.

END
10-86
DTIC