BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 15, Number 2, June 1987

HAZARD RATE ESTIMATION UNDER THE SIMPLE PROPORTIONAL HAZARDS MODEL

BY

PHILIP E. CHENG (程爾觀)

Abstract. This paper considers nonparametric estimation of the hazard rate function λ of the lifetime distribution under a simple proportional hazards model. Based on the model characterization, two naive estimators $\tilde{\lambda}_n$ and $\hat{\lambda}_n$ of the kernel type are proposed. The former is derived from the empirical cumulative hazard function and the latter from the maximum likelihood estimator of the lifetime survival function. The two are shown to be asymptotically equivalent and are compared with their well-known counterparts, $\tilde{\lambda}$ and $\hat{\lambda}$, which are utilized under the general random right censoring model. The comparisons are made in terms of the ratios of asymptotic variances and the widths of simultaneous confidence bands. It is shown that both $\tilde{\lambda}_n$ and $\hat{\lambda}_n$ are preferred to $\tilde{\lambda}$ and $\hat{\lambda}$, a desirable information under this model.

1. **Introduction.** Let $X_1, \cdots, X_n$ be independent lifetimes of n individuals from a population of interest and $Y_1, \cdots, Y_n$ be the corresponding independent censoring times. The so-called random right censorship model assumes that X_i's and Y_i's are independent and the data observed is the set $\{(Z_i, D_i), i = 1, \cdots, n\}$ where $Z_i = \min(X_i, Y_i)$ and $D_i = I(X_i \le Y_i)$. This model is often useful for survival analysis in some medical studies and engineering reliability studies where the Y_i's play the role of censoring mechanism, for instance, random loss of patients in follow-up studies or random failures due to another cause. In connection to the random right censoring model, a particular model with simple proportional hazads, hereafter referred to as the SPH model, was introduced by Armitage (1959). To be specific mathematically, let $P(X_i \ge t) = S(t)$, $P(Y_i \ge t) = C(t)$ and $P(Z_i \ge t) = L(t)$. Note

Received by the editors March 12, 1986 and in revised form November 18, 1986.
1980 Subject Classification. Primary 62G05, Secondary 62N05.
Key words and Phrases: Censored data, hazard funtion, kernel estimation, simultaneous confidence bands, asymptotic relative effeciency.

that $L(t) = S(t)\,C(t)$ by independence. The SPH model assumes that $C(t) = S(t)^{\beta}$ for all $t \geq 0$ for some positive constant β, equivalently, $S(t) = L(t)^{\alpha}$ with $\alpha = 1/(1+\beta)$. The case $\beta = 0$ or $\alpha = 1$ represents the full-data case without the censoring variables Y_i's, or $Y_i = \infty$ for all i. Practically, in some engineering studies where two types of independent failure mechanisms are competing with each other, it is often of interest to test if they represent independent Poisson processes or the independent risk intensities are proportional when the risk patterns are similar, etc. (See for example Cox (1959)), namely, certain SPH model may be valid.

Studies on estimation of the survival function $S(t)$ under the SPH model have been extensive. For testing $S(t)$ in the two sample problem, Efron (1967) computed the Pitman efficiency of competing test statistics. Koziol and Green (1976), Csörgö and Horváth (1981), Chen, Hollander and Langberg (1982) and Ebrahimi (1985) all concentrated on estimating $S(t)$. Recently, Cheng and Lin (1984) investigated the maximum likelihood estimator of $S(t)$ and the small-sample mean squared errors for this m.l.e. was given by Cheng and Chang (1985). Estimation of the hazard rate function $\lambda(t) = -(d/dt)\log S(t)$ via the product limit estimator of Kaplan and Meier (1958) has been studied by McNichols and Padgett (1985). It appears that the derivation of mean squared errors by McNichols and Padgett lacks an intuitive explanation for the SPH model, for instance, it does not indicate a comparison with the general random right censoring model. This motivates the following study on estimating $\lambda(t)$ under the SPH model.

Consider the observed random sample (Z_i, D_i), $i = 1, \cdots, n$, let $Z_{(1)}, \cdots, Z_{(n)}$ be the ordered Z_i's and $D_{(1)}, \cdots, D_{(n)}$ be the corresponding indicators. Under the general random right censoring model, a popular estimator of $\lambda(t)$ is

$$\tilde{\lambda}(t) = \sum_{i=1}^{n} K_b(t - Z_{(i)})\, D_{(i)}(n - i + 1)^{-1},$$

where $K_b(t-s) = K((t-s)/b)/b$; $b = b_n$ is a bandwidth sequence which approaches zero as n approaches infinity, and $K(\cdot)$ is the so-called kernel function that integrates to one, usually a probability density function on the line. The estimator $\tilde{\lambda}$ is a convolution smoothing of the formal derivative of the empirical cumulative hazard function $H_n(t) = \sum_{Z_{(i)} \le t} D_{(i)}/(n-i+1)$ discussed by Tanner and Wong (1983), Ramlau-Hansen (1983) and Yandell (1983). In the full-data case, $D_i \equiv 1$, $S(t) = L(t)$ and $\tilde{\lambda}$ reduces to the estimator $\tilde{h}(t) = \sum_{i=1}^{n} K_b(t - Z_{(i)})(n-i+1)^{-1}$, proposed by Watson and Leadbetter (1964), of the hazard rate function h of Z. Many studies on the hazard rate estimators have appeared in the literature. For the full-data case, there is a survey paper by Singpurwalla and Wong (1984), and for the censored data, there is one by McNichols and Padgett (1985).

It is worth noting that the SPH model assumes the basic facts: $\lambda(t) = \alpha h(t)$ with $h(t) = -(d/dt) \log L(t)$ and α is the expected proportion of the uncensored observations. Therefore a natural estimator of λ is $\tilde{\lambda}_n = \alpha_n \tilde{h}$ where $\alpha_n = n^{-1} \sum_{i=1}^{n} D_i$. The use of this estimator may also be justified from the following two aspects. Firstly, consider the fact $S = L^{\alpha}$ and let L_n be the empirical survival function for L, then $L_n^{\alpha_n}$ is a maximum likelihood estimator for S (Cheng and Lin (1984)). Hence a convolution smoothing of $-\alpha_n \log L_n$ by a delta sequence will provide an estimator $\hat{\lambda}_n = \alpha_n \hat{h}$ for λ, where $\hat{h} = K_b * (-\log L_n)$ is also an estimator of h considered by Rice and Rosenblatt (1976) for the full-data case. The performances of $\tilde{\lambda}_n$ and $\hat{\lambda}_n$ are expected to be asymptotically equivalent. Secondly, the well-known product-limit estimator $\hat{S}$ of Kaplan and Meier (1958) is a maximum likelihood estimator for the survival function S. Thus, a kernel convolution smoothing of $-\log \hat{S}$ will also give an estimator of λ, i.e., $\hat{\lambda} = K_b * (-\log \hat{S})$. The estimator of λ given by McNichols and Padgett (1985) is derived along this latter idea and its performance is asymptotically equivalent to that of $\hat{\lambda}$.

The aim of this paper is to show that the performance of $\tilde{\lambda}_n$ (or $\hat{\lambda}_n$) is different from that of $\tilde{\lambda}$ (or $\hat{\lambda}$). Specifically, as estimator of λ under the SPH model, the former is preferred to the latter based on their asymptotic variances although they are all asymptotically unbiased. Section 2 describes the usual assumptions needed for standard asymptotic results and treats the comparison in terms of pointwise asymptotic normality. The ratio of the asymptotic variances also leads to a comparison of simultaneous confidence bands given in Section 3 as a byproduct. The results provide a more precise information not shown in previous studies of the SPH model.

2. **Pointwise asymptotic properties.** The derivation of biases and variances of $\hat{\lambda}$ and $\hat{\lambda}_n$ can not be easily carried out due to the logarithmic functional form however, Lemma 2 below asserts that the asymptotic biases and variances of $\tilde{\lambda}_n$ and $\hat{\lambda}_n$ are equal and so are those of $\tilde{\lambda}$ and $\hat{\lambda}$. It suffices to study the comparison between $\tilde{\lambda}$ and $\tilde{\lambda}_n$ say, for the pointwise properties. The exact expressions of bias and variance for the estimators $\tilde{h}$ and $\tilde{\lambda}$ are given by Watson and Leadbetter (1964) and Tanner and Wong (1983) respectively. To obtain the bias and variance for the estimator $\tilde{\lambda}_n$ we shall make use of the characterization of the SPH model due to Armitage (1959): the SPH model holds, i.e., $S = L^{\alpha}$ for certain positive constant α less than one if and only if the random variables Z and D are independent. As an immediate consequence, both $\tilde{h}$ and $\hat{h}$ are independent of α_n. The following standard conditions on the survival function S and the kernel delta sequence K_b will be imposed as needed. The distributional assumptions are

S1. S, C and L are continuous and $L = SC$.

S2. The density f of X is continuous on $[0, T]$ where $L(T) > 0$.

S3. The second derivative of f exists and is bounded on $[0, T]$.

S4. $f^{1/2}$ is absolutely continuous and has bounded derivative on $[0, T]$.

The kernel and bandwidth conditions are:

K1. $K_b(x-y) = K((x-y)/b)/b$. K is a bounded probability density function and vanishes off $[-A, A]$, $0 < A < \infty$.

K2. K is symmetric and has a derivative K' satisfying $\int (K'(x))^2\, dx < \infty$.

K3. The bandwidth sequence $b = b_n \to 0$ and $nb \to \infty$ as $n \to \infty$.

K4. $b \log\log n \to 0$ and $(nb)^{-1/2} \log n \to 0$ as $n \to \infty$.

K5. $nb^5(\log b^{-1}) \to 0$ as $n \to \infty$.

LEMMA 1. *Assume* S_1, S_2, K_1 *and* K_3. *Then for* $0 < t \leq T$

(i) $E\tilde{\lambda}_n(t) = E\tilde{\lambda}(t)$, *and*

(ii) $\lim_n \operatorname{Var} \tilde{\lambda}_n(t)/\operatorname{Var} \tilde{\lambda}(t) = \alpha$.

Proof. From formula (3.1) of Watson and Leadbetter and Theorem 1 of Tanner and Wong we have

$$\begin{aligned} E\tilde{\lambda}_n(t) &= \alpha E\tilde{h}(t) \\ &= \alpha \int [1-(1-L(y))^n]\, h(y)\, K_b(t-y)\, dy \\ &= E\tilde{\lambda}(t). \end{aligned}$$

To prove (ii), we follow Watson and Leadbetter and let $I_n(L) = \int_0^L x^{-1}[(x+1-L)^n - (1-L)^n]\, dx$. Observe that for t fixed,

$$\begin{aligned} \operatorname{Var} \tilde{\lambda}_n(t) &= E\alpha_n^2\, E(\tilde{h}(t))^2 - \alpha^2 (E\tilde{h}(t))^2 \\ &= \alpha^2 \operatorname{Var} \tilde{h}(t) + \alpha(1-\alpha)[\operatorname{Var} \tilde{h}(t) + (E\tilde{h}(t))^2] n^{-1} \\ &= \alpha^2 \operatorname{Var} \tilde{h}(t) + O(1/n) \\ &= \alpha(\mathrm{I}) + (\mathrm{II}) + O(1/n), \end{aligned}$$

where

$$(\mathrm{I}) = \int I_n(L(y))\, \lambda(y) K_b^2(t-y)\, dy,$$

and

$$\begin{aligned} (\mathrm{II}) = 2 \iint_{0 \leq y \leq z} &\{(1-L(z))^n [1-(1-L(y))^n] \\ &- (L(y)[(1-L(z))^n - (1-L(y))^n][L(y) - L(z)]^{-1})\} \\ &\cdot \{\lambda(y)\, \lambda(z)\, K_b(t-y)\, K_b(t-z)\} dy\, dz. \end{aligned}$$

It was shown by Tanner and Wong that $\operatorname{Var} \tilde{\lambda}(t) = (\mathrm{I}) + (\mathrm{II})$ and that $(\mathrm{I}) \simeq \left(\int K^2\right) \lambda(t)/nbL(t)$ and $(\mathrm{II}) = o(1/nb)$. The above approximations together with K3 implies (ii).

One may observe from Lemma 1 that the smaller the parameter α is, the better the estimator $\tilde{\lambda}_n$ is, compared with $\tilde{\lambda}$. As an immediate consequence, the pointwise asymptotic normality of the estimators $\tilde{\lambda}_n$, $\tilde{\lambda}$ and $\tilde{h}$ can be put together in a manifestly comparable form.

THEOREM 1. *Let* S1 – S3 *and* K1 – K3 *hold. Assume also* $nb^5 \to 0$ *as* $n \to \infty$. *Then for* $0 < t < T$ *and*

$$\beta_n(t) = \left[nbL(t)/h(t)\left(\int K^2\right)\right]^{1/2},$$

the following weak convergence results are valid

(i) $\alpha^{-1/2}\beta_n(t)(\tilde{\lambda}(t) - \lambda(t)) \to N(0, 1)$,

(ii) $\beta_n(t)(\tilde{h}(t) - h(t)) \to N(0, 1)$,

(iii) $\alpha^{-1}\beta_n(t)(\tilde{\lambda}_n(t) - \lambda(t)) \to N(0, 1)$.

Proof. Notice that the assumptions together with (2.16) of Rice and Rosenblatt ensure that $\beta_n(t)(E\tilde{h}(t) - h(t)) \to 0$ as $n \to \infty$. It is here for the bias part the added condition $nb^5 \to 0$ is used. By Lemma 1 (i) $\beta_n(t)(E\tilde{\lambda}(t) - \lambda(t)) \to 0$ follows similarly. Since it follows from Tanner and Wong that $\alpha^{-1/2}\beta_n(t)(\tilde{\lambda}(t) - E\tilde{\lambda}(t)) \to N(0, 1)$ and, implicitly in the absence of censoring, $\beta_n(t)(\tilde{h}(t) - E\tilde{h}(t)) \to N(0, 1)$, hence (i) and (ii) are established. For (iii), write $\tilde{\lambda}_n(t) - \lambda(t) = \alpha_n(\tilde{h}(t) - h(t)) + (\alpha_n - \alpha)h(t)$ where $\alpha_n \to \alpha$ and $\beta_n(t)(\alpha_n - \alpha) = O(b^{1/2})$ in probability. By Slutsky's theorem and (ii) we conclude (iii).

As mentioned in the first paragraph of Section 2 we give the following two facts which imply that all the results of Lemma 1 and Theorem 1 are also valid if any estimator(s) with tilta "~" is replaced with the corresponding one(s) with hat "^". Consequently, a complete picture of the pointwise asymptotic properties is provided. The first fact is

LEMMA 2. *Let* K1, K3 *and* S1 – S2 *hold. Then for* $0 < t < T$

$$\limsup_n n|\tilde{h}(t) - \hat{h}(t)| \le Ag(t)\|K\|/L(t)^2.$$

almost surely, where $\|K\| = \sup_x |K(x)|$ *and g is the density function of the random variable Z.*

Proof.

$$\tilde{h}(t) - \hat{h}(t) = \sum_{i=1}^{n} K_b(t - Z_{(i)})\left[\frac{1}{n - i + 1} - \log\left(1 + \frac{1}{n - i + 1}\right)\right].$$

Since $|x - \log(1 + x)| \le x^2/2$ if $0 \le x \le 1$, we find that

$$|\hat{h}(t) - \tilde{h}(t)| \le \sum_{i=1}^{n} K_b(t - Z_i)/2R(Z_i)^2$$

where $R(x) = \#\{Z_j \ge x\}$, the number of individuals at risk at time x. Henece,

$$n|\tilde{h}(t) - \hat{h}(t)| \le [n/2bR(Z_i)^2]\sum_{i=1}^{n} K((t - Z_i)/b)$$

$$\le [\|K\|/2bn^{-2} R(t + bA)^2] \sum_{i=1}^{n} I[t - bA \le Z_i \le t + bA]/n.$$

The right-hand side of the last inequality converges w. p. 1 to $Ag(t)\|K\|/L(t)^2$ by strong law of large numbers and the fact that $b \to 0$ as $n \to \infty$.

Analogous to Lemma 2, it follows from Lemma 1 of Breslow and Crowley (1974) that $|\tilde{\lambda}(t) - \hat{\lambda}(t)| = O(1/n)$ w. p. 1 under the same conditions. This is the second fact.

3. **Simultaneous confidence bands.** The method of Bickel and Rosenblatt (1973) for deriving simultaneous confidence bands for an unknown density function has been applied to estimation of the hazard rate function. Rice and Rosenblatt (1976) and Sethuraman and Singpurwalla (1981) obtained the same confidence bands based on two different estimators in the full-data case. Along the same line of derivation, Yandell (1983) gave confidence bands for λ based on $\tilde{\lambda}$ for randomly right-censored data. The purpose of this section is to give confidence bands for λ based on the estimators $\hat{\lambda}_n$ and $\tilde{\lambda}_n$ so that a comparison between the confidence bands from $\hat{\lambda}_n$ (or $\tilde{\lambda}_n$) and that from $\hat{\lambda}$ (or $\tilde{\lambda}$) can be made under the SPH model.

First of all, one important fact must be noted. It can be easily checked that the two facts at the end of Section 2, Lemma 2 and its immediate parallel, also hold true uniformly over the interval $[0, T]$. This is due to that $\|\tilde{\lambda} - \hat{\lambda}\| \le 2A\|K\|/R(T + bA)$ and $\|\tilde{h} - \hat{h}\| \le nA\|g\|\cdot\|K\|/[R(T + bA)]^2$ almost surely, where $\|\cdot\|$ denotes supremum modulus over the interval $[0, T]$ except that $\|K\|$ is defined over its own domain. Incidentally, $L(T + bA) > 0$

if n is sufficiently large since L is continuous and $L(T)>0$. The above fact implies that confidence bands (over $[0, T]$) based on $\hat{\lambda}_n$ and $\tilde{\lambda}_n$ are asymptotically equivalent and so are the bands based on $\hat{\lambda}$ and $\tilde{\lambda}$. Therefore, it suffices to compare the bands derived from $\hat{\lambda}_n$ and $\tilde{\lambda}$. Before giving the bands based on $\hat{\lambda}_n$ we need to state two existing results. The following notations will be used. Let $r_n=(2\log(T/b))^{1/2}$ and $d_n=r_n+(\log \gamma)/r_n$, where

$$\begin{aligned} \gamma &= r_n[K^2(-A)+K^2(A)](8\pi)^{-1/2}\left(\int K^2\right)^{-1}, \quad \text{if } \ K(A)>0; \\ &= \left(\int (K')^2 \Big/ \int K^2\right)^{1/2} \Big/ 2\pi \quad \text{if } \ K(A)=0. \end{aligned}$$

Define $A(t)=\lambda(t)\,L^{-1}(t)\left(\int K^2\right)$, $B(t)=h(t)\,L^{-1}(t)\left(\int K^2\right)$ and $h_n^*(t)=\int K_b(t-s)\,dH(s)$. The two results are as follows.

LEMMA 3. (Yandell) *Let* S1 − S4 *and* K1 − K4 *hold. Assume* b lon $n\to 0$. *Then for each real value* x

$$P\{r_n(M_n-d_n)<x\}\longrightarrow \exp(-2e^{-x})$$

where $M_n=\|(nb/A(t))^{1/2}[\tilde{\lambda}(t)-E\tilde{\lambda}(t)]\|$.

LEMMA 4. (Sethuraman and Singpurwalla) *Under the same assumptions of Lemma* 3 *without the condition* b log $n\to 0$,

$$P\{r_n({}_hM_n-d_n)<x\}\longrightarrow \exp(-2e^{-x})$$

where ${}_hM_n=\|(nb/B(t))^{1/2}[\hat{h}(t)-h_n^*(t)]\|$.

We remark that Theorem 2.2 of Sethuraman and Singpurwalla (1981) is also valid with $\|\cdot\|$ of Lemma 4 taken over $[0, T]$ instead of over $[bA, T]$ only, because the Wiener process in their proof may be defined to be zero over $(-\infty, 0]$. However, in Lemmas 3 and 4 above, if $E\tilde{\lambda}(t)$ and $h_n^*(t)$ are replaced with $\lambda(t)$ and $h(t)$ respectively, and condition K5 is imposed, then the appropriate domain for $\|\cdot\|$ is $[bA, T]$ instead of $[0, T]$. This is due to the nonnegligible bias near the boundary value $t=0$ if the hazard rate function does not vanish off smoothly near $t=0$. It is worth noting that such phenomenon is also discussed by Falk

(1984) and Rice (1984) in the context of kernel estimation of density and regression functions.

THEOREM 2. *Let* S1 − S4 *and* K1 − K4 *hold. Then for each real value* x

$$P\{r_n(\widehat{M}_n - d_n) < x\} \longrightarrow \exp(-2e^{-x})$$

where $\widehat{M}_n = \|(nb/aA(t))^{1/2}[\hat{\lambda}_n(t) - ah_n^*(t)]\|$.

Proof. Utilizing the fact $\lambda(t) = \alpha h(t)$ under the SPH model, we observe the following:

$$\begin{aligned}
&(nb/\alpha A(t))^{1/2}[\hat{\lambda}_n(t) - \alpha h_n^*(t)] \\
&\quad = (nb/\alpha A(t))^{1/2}[\alpha\hat{h}(t) - \alpha h_n^*(t)] \\
&\qquad + (nb/\alpha A(t))^{1/2}(\alpha_n - \alpha)\,\hat{h}(t) \\
&\quad = (nb/B(t))^{1/2}[\hat{h}(t) - h_n^*(t)] + O((b \log\log n)^{1/2}).
\end{aligned}$$

By Lemma 4 and K4, the proof is completed.

As remarked before Theorem 2, the bias part can be treated in a similary way.

COROLLARY. *If* K5 *holds in addition to the assumptions of Theorem* 2, *then*

$$P\{r_n({}_{\lambda}\widehat{M}_n - d_n) < x\} \longrightarrow \exp(-2e^{-x})$$

where ${}_{\lambda}\widehat{M}_n = \sup_{bA \le x \le T} (nb/\alpha A(t))^{1/2} |\hat{\lambda}_n(t) - \lambda(t)|$.

In view of the construction of confidence bands (p. 1077 Bickel and Rosenblatt (1973)) it follows from Lemma 3 and Corollary above that the widths of the confidence bands based on $\hat{\lambda}_n$ (or $\tilde{\lambda}_n$) is approximately $\alpha^{1/2}$ times as wide as that based on $\hat{\lambda}$ (or $\tilde{\lambda}$) uniformly over the interval $[bA, T]$. Parallel to Theorem 1, this concludes that if the SPH model is valid, an increase of asymptotic relative efficiency is achieved with the proposed estimator $\hat{\lambda}_n$ (or $\tilde{\lambda}_n$) versus the usual choice $\hat{\lambda}$ (or $\tilde{\lambda}$).

Acknowledgments. The careful inspection of the referees leading to the present version is deeply appreciated. This research was partiailly supported by the National Science Council, Republic of China. The author is also grateful for the research facilities supported by the Biostatistics Department, Uuiversity of Michigan while the revision is fulfilled.

REFERENCES

1. P. Armitage, *The comparison of survival curves,* J.R. Statist. Soc. **A** 122 (1959), 279-300.

2. P.J. Bickel and M. Rosenblatt, *On some global measures of the deviations of density function estimates,* Ann. Statist. 1 (1973), 1071-1095.

3. N. Breslow and J. Crowley, *A large sample study of the life table and product limit estimate under random censorship,* Ann. Statist. 2 (1974), 437-453.

4. Y.Y. Chen, M. Hollander and N.A. Langberg, *Small-sample results for the Kaplan-Meier estimator,* J. Amer. Statist. Assoc. 77 (1982), 141-144.

5. P.E. Cheng and G.D. Lin, *Maximum likelihood estimation of survival function under the Koziol-Green proportional hazards model,* T.R. No. 84-05, Inst. Statist., Acad. Sinica. To appear in Statist. and Prob. Letters (1984).

6. P.E. Cheng and Y.C. Chang, *PLE versus MLE for survival functions under a simple proportional hazards model,* J. Chinese Statist. Asso. 23 (1985), Special Issue, 57-73.

7. D.R. Cox, *The analysis of exponentially distributed life-times with two types of failure,* J.R. Statist. Soc. **B**21 (1959), 411-421.

8. S. Csörgö and L. Horváth, *On the Koziol-Green model for random censorship.* Biometrika, 68 (1981), 391-401.

9. Nader Ebrahimi, *Nonparametric estimation of survival functions for incomplete observations when the life time distribution is proportionally related to the censoring time distribution.* Comm. Statist. Theor. Math. 14 (1985), 2887-2898.

10. B. Efron, *The two-sample problem with censored data.* Pro. Fifth Berkeley Symp. Math. Statist. Probability 4 (1967), 831-853.

11. M. Falk, *Kernel estimation of a density at an unknown endpoint of its support.* So. Afr. Statist. J. 18 (1984), 91-96.

12. J.A. Koziol and S.B. Green, *A Cramér-von Mises statistics for randomly censored data,* Biometrika 63 (1976), 465-474.

13. D.T. McNichols and W.J. Padgett, *Nonparametric methods for harard rate estimation from right-censored samples,* J. Chinese Statist. Assoc. 23 (1985), Special Issue, 1-15.

14. H. Ramlau-Hansen, *Smoothing counting process intensities by means of kernel functions,* Ann. Statist. 11 (1983). 453-466.

15. J. Rice, *Boundary modifications for kernel regression,* Comm. Statist. **A**13 (1984), 893-900.

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

17. J. Sethuraman and N.D. Singpurwalla, *Large sample estimates and uniform confidence bounds for the failure rate function based on a naive estimator,* Ann. Statist. 9 (1981), 628-632.

18. N.D. Singpurwalla and M.Y. Wong, *Estimation of the failure rate—a survey of nonparametric methods, Part I: Non-Bayesian Methods,* Comm. Statist. **A**12 (1983), 559-588.

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

20. G.S. Watson and M.R. Leadbetter, *Hazard analysis, II,* Sankhyā, Ser. **A**26 (1964), 110-116.

21. B.S. Yandell, *Nonparametric inference for rates with censored survival data,* Ann. Statist. 11 (1983), 1119-1135.

Institute of Statistics
Academia Sinica
Taipei, Taiwan