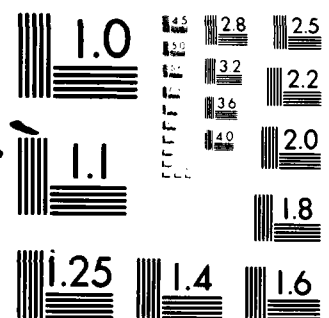AD-A099 430   GEORGE WASHINGTON UNIV WASHINGTON DC INST FOR MANAGE--ETC F/G 12/1
                ESTIMATION OF THE FAILURE RATE. A SURVEY OF NONPARAMETRIC METHO--ETC(U)
                JUL 80   N D SINGPURWALLA, M WONG   N00014-77-C-0263
UNCLASSIFIED   SERIAL-T-422   NL

END
DATE
FILMED
F. 8!
DTIC

1.0

1.1

1.25 1.4 1.6

2.8 2.5

3.2 2.2

3.6

4.0 2.0

1.8

MICROCOPY RESOLUTION TEST CHART

LEVEL

AD A099430

STUDENTS FACULTY STUDY R
ESEARCH DEVELOPMENT FUT
URE CAREER CREATIVITY CO
MMUNITY LEADERSHIP TECH
NOLOGY FRONTIE ESIGN
ENGINEERING APP ENC
GEORGE WASHIN NIV

DTIC FILE COPY

1981

A

81 5 27 042

SCHOOL OF ENGINEERING
AND APPLIED SCIENCE

ESTIMATION OF THE FAILURE RATE:
A SURVEY OF NONPARAMETRIC METHODS.
PART I. NON-BAYESIAN METHODS.

by

Nozer D. Singpurwalla†
Man-Yuen Wong†

Serial-T-422
10 July 1980

The George Washington University
School of Engineering and Applied Science
Institute for Management Science and Engineering

The findings in this report are not to be construed as an official
Department of the Army position, unless so designated by other authorized
documents.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER T-422	2. GOVT ACCESSION NO. AD-A099430	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* ESTIMATION OF THE FAILURE RATE-- A SURVEY OF NONPARAMETRIC METHODS PART I: NON-BAYSIAN METHODS		5. TYPE OF REPORT & PERIOD COVERED SCIENTIFIC
		6. PERFORMING ORG. REPORT NUMBER T-422
7. AUTHOR(*s*) NOZER D. SINGPURWALLA MAN-YUEN WONG		8. CONTRACT OR GRANT NUMBER(*s*) N00014-77-C-0623 DAAG 29-80-C-0067
9. PERFORMING ORGANIZATION NAME AND ADDRESS THE GEORGE WASHINGTON UNIVERSITY, SEAS INSTITUTE FOR MANAGEMENT SCIENCE & ENGINEERING WASHINGTON, DC 20052		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS OFFICE OF NAVAL RESEARCH, ARLINGTON, VA U.S. ARMY RESEARCH OFFICE, TRIANGLE PARK, NC		12. REPORT DATE 10 July 1980
		13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* NONE
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

APPROVED FOR PUBLIC SALE AND RELEASE: ITS DISTRIBUTION IS UNLIMITED.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

FAILURE RATE	KERNEL ESTIMATES
GENERALIZED FAILURE RATE	GENERALIZED JACKKNIFE
HAZARD RATE	ESTIMATION OF THE FAILURE RATE
LIFE TABLE	UNIFORM CONFIDENCE BOUNDS
PRODUCT LIMIT ESTIMATES	

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

In this paper several nonparametric, non-Bayesian methods for estimating the failure rate function on which no monotonicity conditions have been imposed are surveyed. The survey attempts to consolidate and synthesize literature from several diverse areas of application, and endeavors to be as up-to-date as is feasible.

DD ,_{JAN 73} 1473 FORM EDITION OF 1 NOV 65 IS OBSOLETE
 S/N 0102-014-6601

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering
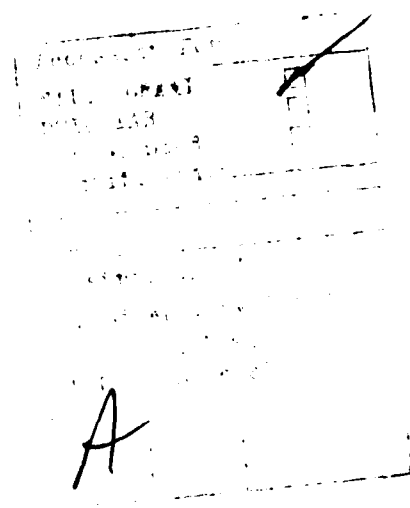
Abstract
of
Serial T-422
10 July 1980

ESTIMATION OF THE FAILURE RATE--
A SURVEY OF NONPARAMETRIC METHODS
PART I: NON-BAYESIAN METHODS

by

Nozer D. Singpurwalla*†
Man-Yuen Wong*

In this paper several nonparametric, non-Bayesian methods for estimating the failure rate function on which no monotonicity conditions have been imposed are surveyed. The survey attempts to consolidate and synthesize literature from several diverse areas of application, and endeavors to be as up-to-date as is feasible.

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

ESTIMATION OF THE FAILURE RATE--
A SURVEY OF NONPARAMETRIC METHODS
PART I: NON-BAYESIAN METHODS

by

Nozer D. Singpurwalla
Man-Yuen Wong

1. Introduction

The failure rate function $h(x) = f(x)/(1-F(x))$, corresponding
to a distribution F with density f , is one of the most important
parameters in reliability theory. A problem of considerable interest,
especially to inventory theorists, logistics planners, reliability en-
gineers, and seismologists, is the estimation of h from a sample of
n independent and identically distributed lifetimes $X_1,\ldots,X_n$.

Several nonparametric methods for estimating h have been proposed in
the literature; many of these methods are based on the assumption that
h is a monotonic function of x . Whereas such an assumption may be
realistic in many applications, estimators of h which are not based
upon any assumptions about h , appear to be more palatable to a large
spectrum of users.

This paper is an attempt to survey the various nonparametric,
and non-Bayesian methods for estimating an h on which no monotonicity
conditions are imposed, and to list the important properties of these
estimates. Bayesian methods for estimating h will be surveyed in a
sequel to this paper.

It is to be emphasized that the literature on the topic of this paper is fairly extensive and scattered, ranging from reliability studies to mortality studies and seismology. Furthermore, it is technically quite detailed, involving a host of regularity conditions, and therefore difficult to synthesize. Consequently, a complete discussion of each method of estimation is not undertaken here. It is hoped that this paper will serve as a guide, and the reader, having read the survey, may then be motivated to read those individual papers which capture his or her interest.

The organization of this survey is as follows. Section 2 pertains to some preliminaries and introduces some notation and terminology. Section 3 briefly discusses a graphical method of estimating h , and Section 4 summarizes estimators of h which are commonly used by biometricians, with a special emphasis on the "life table" and the "product limit estimates." Section 5, which occupies a large portion of this survey, pertains to "kernel estimates" of h and methods of improving these estimates using the "generalized jackknife." Section 6 discusses some recent results involving a "naive estimator" and its smoothing using kernels of fixed and random band widths. In both Sections 5 and 6 some global results for the appropriate estimates are also given. In Section 7, the "generalized failure rate function" is introduced and methods of estimating it are outlined.

2. Preliminaries and Notation

Let X be a nonnegative random variable with absolutely continuous distribution function F , and probability density function f . If X denotes the lifetime of a physical device or a biological organism, then F(x) is the probability of the event that the device has failed by age x , and R(x) = 1-F(x) , called the *reliability* of the device to age x , is the probability that the device survives to age x . The *failure rate function* h(x) at age x > 0 is defined as

$$h(x) = \frac{F(x)}{1 - F(x)} , \quad \text{if } F(x) \neq 1 . \tag{2.1}$$

- 2 -

The function $h(x)$ is also known as the "hazard function," the "mortality intensity," the "age-specific death rate," or the "force of mortality," depending upon the context in which $h(x)$ is used.

The failure rate, as a function of time (age), has a probabilistic interpretation, namely, $h(x)dx$ represents the probability that a device which is surviving at age x will fail in the interval $(x, x+dx)$. On the basis of physical, biological, or engineering considerations, one is at liberty to choose the functional form of $h(x)$ for a particular device or organism. Once this is done, a differential equation in $h(x)$ is obtained, from which $f(x)$ and $h(x)$ can easily be determined. In fact, one can show [see Barlow and Proschan (1975), or Mann, Schafer, and Singpurwalla (1974)], that

$$h(x) = \frac{d}{dx} [-\ln(1-F(x))] , \tag{2.2}$$

and

$$f(x) = h(x)\left[\exp - \int_0^x h(t)dt\right] . \tag{2.3}$$

The purpose of this paper is to discuss the various nonparametric and non-Bayesian methods that have been proposed to estimate $h(x)$, and to state the properties of the proposed estimators.

Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the ordered values of a random sample $X_1, \ldots, X_n$ from a population with absolutely continuous distribution function F and density f. Let $F_n(x)$ be the *empirical sample distribution function* of $X_1, \ldots, X_n$; that is,

$$F_n(x) = \frac{1}{n} \{\text{number of observations among } X_1, \ldots, X_n \leq x\} ,$$

and

$$R_n(x) = \frac{1}{n} \{\text{number of observations among } X_1, \ldots, X_n > x\} .$$

Throughout this paper, the following abbreviations will be used.

If $\{a_n\}$ and $\{b_n\}$ are two sequences, then "$a_n \sim b_n$" is read "a_n is asymptotically equivalent to b_n," and means that the ratio of a_n

- 3 -

to b_n has limit one. The notation "$a_n = o(b_n)$" means that the ratio a_n to b_n has limit zero, and "$a_n = O(b_n)$" means that the absolute value of the ratio is bounded in the limit. The terms $o(b_n)$ and $O(b_n)$ are frequently used to represent some unknown function of n which has the appropriate property.

A sequence of random variables $\{X_n(\omega)\}$ defined on a space Ω is said to *converge with probability one*, or *converge almost surely* to a random variable $X(\omega)$, if

$$\lim_{n \to \infty} X_n(\omega) = X(\omega)$$

for almost all $\omega \in \Omega$; that is,

$$P\left\{\omega; \lim_{n \to \infty} X_n(\omega) = X(\omega)\right\} = 1 .$$

This type of convergence will be denoted by "$X_n \to X$, w.p.1 ."

The sequence of random variables $\{X_n(\omega)\}$ is said to *converge in probability* to a random variable $X(\omega)$, denoted by "$X_n \overset{P}{\to} X$", iff for every $\varepsilon > 0$

$$\lim_{n \to \infty} P\left\{\omega : |X_n(\omega) - X(\omega)| \geq \varepsilon\right\} = 0 .$$

Let $X_n(\omega)$ be a sequence of random variables, and let $\{b_n\}$ be any other sequence. Then the notation "$X_n = o_p(b_n)$" denotes the fact that $(X_n/b_n) \overset{P}{\to} 0$, so that $X_n = o_p(1)$ is another way of writing $X_n \overset{P}{\to} 0$. The notation "$X_n = O_p(b_n)$" denotes the fact that for every $\varepsilon > 0$, there exists a K_ε such that

$$P\left\{\omega : \left|\frac{X_n(\omega)}{b_n}\right| \leq K_\varepsilon\right\} \geq 1 - \varepsilon$$

for all $n > N_\varepsilon$, where N_ε is some integer which depends on ε.

If the sequence of random variables $\{X_n(\omega)\}$ have corresponding distribution functions F_n , and the random variable $X(\omega)$ has distribution function F , then the sequence $\{X_n(\omega)\}$ is said to *converge in distribution*, or *converge in law* to the random variable $X(\omega)$, denoted by "$X_n \overset{\mathcal{D}}{\rightarrow} X$" , if $F_n(\omega) \rightarrow F(\omega)$ at every point ω in the set of continuity points of F .

3. A Graphical Estimate

Motivated by Equation (2.2), Watson and Leadbetter (1964a) have proposed a simple graphical estimator h_g of h , where

$$h_g(x) = \text{the graphical derivative of } \{-\ln(1-F_n(x))\} .$$

The curve $-\ln(1-F_n(x))$ is plotted against x , and a smooth curve drawn through the points by any reasonable method. The slope of this curve at point x is an estimate of $h(x)$.

Clearly, this method involves an unspecified amount of smoothing, and obtaining the slope of a hand-drawn curve is notoriously difficult.

4. Estimates Used in Biometry

Some of the earliest discussions on estimating the failure rate are in studies of the mortality rates of animals and human beings. A survey of this early literature is given by Kimball (1960). Conspicuous among these is the paper by Grenander (1956), who obtains the maximum likelihood estimator of h under the assumption that it is nondecreasing.

Suppose that the domain of x , $[0,\infty)$, is divided into k intervals $[0,\alpha_1)$, $[\alpha_1,\alpha_2)$, ..., $[\alpha_{k-1},\infty)$. Let the j*th* time interval I_j , $j=1,\ldots,k$ be

$$I_j = \alpha_j - \alpha_{j-1} , \text{ with } \alpha_0 = 0 , \alpha_k = \infty .$$

Let n_j be the number of X_i's which fall in I_j , $j=1,\ldots,k$, and let $\alpha_{j-1} + t_{j-1,\ell}$ be the age at death of the ℓth individual in the $j th$ interval, $\ell=1,\ldots,n_j$, with $0 < t_{j-1,\ell} < I_j$; here $t_{j-1,0} = 0$.

Kimball (1960) proposes two different sampling procedures for estimation:

(1) the time intervals I_j are chosen arbitrarily, so that each n_j associated with an interval is a random variable;

(2) the n_j are preassigned, and the I_j's become random.

We shall first consider the estimation of h under sampling scheme (1).

The *actuarial estimate* $h_a(x)$ of $h(x)$, for $x \in [\alpha_{j-1} , \alpha_j)$ is

$$h_a(x) = h_a\left(\frac{\alpha_{j-1} + \alpha_j}{2}\right) = \frac{n_j}{\alpha_j - \alpha_{j-1}} \left(\frac{1}{n - n_1 - \ldots - n_{j-1} - \frac{1}{2}n_j}\right) ;$$

it is the number of deaths per unit time in the interval I_j divided by the average number of survivals in the interval.

A motivation for considering h_a is based on the fact that h_a is related to the approximation

$$h(x) = \frac{\dfrac{dF(x)}{dx}}{1-F(x)} \approx \frac{1}{\text{avg}[1-F(x)]} \frac{\Delta F(x)}{\Delta x} = \frac{\frac{1}{I}\left[F\left(x+\frac{I}{2}\right) - F\left(x-\frac{I}{2}\right)\right]}{\frac{1}{2}\left[2 - F\left(x+\frac{I}{2}\right) - F\left(x-\frac{I}{2}\right)\right]} ,$$

where $\text{avg}(\cdot)$ denotes the average of $(\cdot)$.

Another estimate of h based on preassigned time intervals has been proposed by Sacher (1956), who apparently bases his estimate on the fact that

$$h(x) = -\frac{d}{dx} \ell n[1-F(x)] \approx -\frac{\Delta}{\Delta x} \ell n[1-F(x)]$$

$$= \left\{ -\ell n\left[1 - F\left(x + \frac{I}{2}\right)\right] - \ell n\left[1 - F\left(x - \frac{I}{2}\right)\right] \right\} \Big/ I$$

$$= \frac{1}{I} \ell n\left[\frac{1 - F\left(x - \frac{I}{2}\right)}{1 - F\left(x + \frac{I}{2}\right)}\right] .$$

This leads to the estimate $h_{S_1}(x)$ for $x \in [\alpha_{j-1}, \alpha_j)$, where

$$h_{S_1}(x) = h_{S_1}\left(\frac{\alpha_{j-1} + \alpha_j}{2}\right) = \frac{1}{\alpha_j - \alpha_{j-1}} \ell n\left[\frac{n - n_1 - \ldots - n_{j-1}}{n - n_1 - \ldots - n_j}\right] .$$

Whereas h_{S_1} is almost always positively biased, h_a tends to be negatively biased as x increases. The bias of both estimates varies with I ; in h_{S_1} it increases monotonically, whereas in h_a it sometimes increases monotonically, and at other times it starts off being positive, and then decreases through zero through negative values. Since h_a is essentially a maximum likelihood estimate, it has desirable large sample properties. However, in many real life situations, expecially those involving human experiments, the sample sizes are not large enough to justify this argument.

Under sampling scheme (2), originally conceived by Moran (1951), an unbiased estimate of h has been proposed by Seal (1954) when I_j is small enough so that

$$h(\alpha_{j-1} + \tau) = h(\alpha_{j-1}) , \quad 0 \leq \tau < I_j .$$

Seal's estimate $h_{S_2}(\alpha_{j-1})$ is

$$h_{S_2}(\alpha_{j-1}) = \frac{n_j - 1}{\displaystyle\sum_{k=0}^{n_j-1} \left(n - n_1 - \ldots - n_{j-1} - k\right)\left(t_{j-1,k+1} - t_{j-1,k}\right)} .$$

This estimate is efficient and complete, and has variance approximately equal to $h^2(\alpha_{j-1}) / (n_j - 2)$.

The assumption that $h(\alpha_{i-1} + 1) = h(\alpha_i)$, $0 \le 1 < 1_j$, though suitable for some purposes, fails to be consistent with most data in biometry which suggest that h is usually a nondecreasing function of h . We refer the reader to Kimball (1960) for more details on h_{S_2} and modifications to it.

4.1 The Life Table and Product Limit Estimates

It is now convenient to introduce here a close relative of the failure rate estimate; the "life table estimate of a survival distribution," which is dominantly featured in the literature on survival studies. A key feature of survivorship and life testing data is the presence of "censored" or withdrawn observations. Breslow and Crowley (1974), in a seminal paper on this topic, incorporate a model of random censorship and study the "standard life table estimate" using grouped data. The notation and terminology given below is an adaptation of that used by Breslow and Crowley, who in turn have borrowed it from Efron (1967).

Let X_n^0 , $n=1,\ldots,N$, be the true survival time for the nth individual included in the life table. These are assumed to be independent random variables having a common but unknown distribution function F^0 . The period of observation or follow-up, for the nth individual, will typically be limited by an amount Y_n , $n=1,\ldots,N$. Under the model of random censorship, $Y_1,\ldots,Y_N$ are also assumed to be a random sample drawn *independently* of the X_n^0 from an unknown distribution H . The X_n^0 are said to be *censored on the right* by the Y_n , since one observes only

$$X_n = \min(X_n^0, Y_n) \quad \text{and} \quad \delta_n = I_{[X_n^0 \le Y_n]} ,$$

where the indicator function $\delta_n = 0$, if X_n is censored, and $\delta_n = 1$, otherwise, for $n=1,\ldots,N$.

The observed X_n's constitute a random sample from the distribution function F, where

$$(1-F) = (1-H)(1-F^0) .$$

The distribution function $\tilde{F}$ of an uncensored observation is called a *subdistribution function*. Specifically,

$$\tilde{F}(x) = P\left\{X_n \leq x , \delta_n = 1\right\} = \int_0^x (1-H(s)) dF^0(s) .$$

In life table analysis, interest centers around the distribution function F^0.

The time period of observation $[0,T]$ is partitioned into k intervals $I_j = (\xi_{j-1}, \xi_j]$, with $0 \equiv \xi_0 < \xi_1 < \ldots < \xi_k = T$. The conditional probabilities of failure in the j*th* interval

$$q_j = \frac{F^0(\xi_j) - F^0(\xi_{j-1})}{1 - F^0(\xi_{j-1})}$$

are the parameters of interest, since they are approximations to the failure rate for small values of I_j. If $p_j = 1 - q_j$, then the probability of survival to ξ_m,

$$P_m = 1 - F^0(\xi_m) = \prod_{j=1}^m p_j .$$

Let N_j, D_j, and W_j represent, respectively, the number of items surviving at the start of I_j, the number known to have failed in I_j, and the number censored (or withdrawn surviving) in I_j. Furthermore, suppose that it is possible to subdivide N_j and D_j according to whether an item is due for withdrawal (N_{2j}, D_{2j}) or not due for withdrawal $(N_{1j}$ and $D_{1j})$ in I_j. This subdividion is possible only if one knows the censoring variables Y_n for all N items, knowledge which in most practical situations is hard to come by.

The *standard life table estimate* (SD) used most often in practice [Berkson and Gage (1950), Cutler and Ederer (1958), Gehan (1969)] is given by

$$\hat{q}_j = \frac{D_j}{N_j - \frac{1}{2} W_j} \ .$$

Another estimate of q_j , the *reduced sample estimate* (RS) studied by Kaplan and Meier (1958) is

$$\hat{q}_j = \frac{D_{1j}}{N_{1j}} \ .$$

It should be noted that only the RS estimate is generally consistent for q_j ; however, because it is based on items which are not due for withdrawal in I_j , it is not commonly used. On the other hand, the SD estimate is consistent only under special conditions relating the survival distribution F^0 and the censoring distribution H . These conditions are stated more precisely in the following theorem.

Theorem 4.1 [Breslow and Crowley (1975)]: Let the censoring distribution H be absolutely continuous with density h on $[0,T]$. A necessary and sufficient condition that the SD estimate yield a consistent estimate of F^0 at the end points of each of the k intervals, for any choice of interval end points, is that F^0 satisfy

$$F^0(x) = 1 - [1/(1 + cH(x))]^{\frac{1}{2}}$$

for some constant $c > 0$.

Quite often, a reasonable approximation to the censoring distribution is the uniform distribution; thus $H(x) = x/T$, $0 < x < T$, in which case the distributions yielding a consistent estimate satisfy $F^0(x) = 1 - [1/(1+cx)]^{\frac{1}{2}}$. However, the distribution $1 - [1/(1+cx)]^{\frac{1}{2}}$ has a failure rate function which decreases in time, implying negative aging, and this may be unrealistic in many situations.

Theorem 4.1 therefore suggests that the SD estimate is generally inconsistent, with $\hat{q}_j$ converging to q_j^* , where $q_j^* \neq q_j$ is specified by Breslow and Crowley. The asymptotic normality of the SD estimate has also been established by Breslow and Crowley, who also point out that the individual $\hat{q}_j$'s are asymptotically uncorrelated.

An estimate $\hat{P}_m$ of P_m , the probability of survival to ξ_m , is given by

$$\hat{P}_m = \prod_{j=1}^{m} (1-\hat{q}_j) \ .$$

Since the estimates $\hat{q}_j$ depend on the partition $0 < \xi_1 < \xi_2 < \ldots < \xi_k < T$, and sample size N , it is helpful to denote the dependence of $\hat{P}_m$ on these choices by writing

$$\hat{P}_m = 1 - \hat{F}^0_{k,N}(t) \quad \text{for} \quad t \in [\xi_m, \xi_{m-1}) \ , \quad m=1,2,\ldots,k-1 \ .$$

Kaplan and Meier (1958) introduce the now extensively discussed *product limit estimate* (PL) $\hat{F}^0_N$ of F^0 , where $\hat{F}^0 = \lim_{k \to \infty} \hat{F}^0_{k,N}$, and where the right continuous limit is taken under any nested sequence of partitions such that $\sup_{1 \leq m \leq k} |\xi_m - \xi_{m-1}| \to 0$. Kaplan and Meier show that for $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(N)}$, the PL estimate can be written as

$$1 - F^0_N(t) = \prod_{j=1}^{N} \left(1 - \frac{1}{N-j+1}\right)^{\delta^t_{(j)}} \ ,$$

where

$$\delta^t_{(j)} = \begin{cases} 1 \ , & \text{if } X_{(j)} < t \ , \text{ and } X_{(j)} \text{ is uncensored,} \\ 0 \ , & \text{otherwise.} \end{cases}$$

Strong consistency of the PL estimate is shown by Peterson (1977). In order to obtain the asymptotic distribution of $\hat{F}^0_N$, Breslow and Crowley define the stochastic process $\{Z^*_N(t) \ ; \ 0 < t < T\}$, where $Z^*_N(t) = \sqrt{N} \left(\hat{F}^0_N(t) - F^0(t)\right)$. Then

Theorem 4.2 [Breslow and Crowley (1974)]: Let $T < \infty$ satisfy $F(T) < 1$, and suppose that F^0 and H are continuous. Then the process $\{Z_N^*(t) ; 0 < t < T\}$ converges weakly to a mean 0 Gaussian process $\{Z^*(t) ; 0 < t < T\}$ whose covariance structure for $s \leq t$ is given by

$$\text{Cov}\left(Z^*(s), Z^*(t)\right) = \left(1 - F^0(s)\right)\left(1 - F^0(t)\right) \int_0^s \left(1 - F(x)\right)^{-2} d\tilde{F}(x) .$$

Hall and Wellner (1980) transform the above weak convergence result to a Brownian bridge form, and then obtain uniform confidence bands for F^0.

An alternate method of estimating F^0 under different censoring schemes is given by Turnbull (1976). Wegman, Nour, and Kukuk (1980) give a time series approach to life table construction.

5. Kernel Estimates

Of the various methods used for estimating h, those based on smoothing functions, called *kernels*, happen to be predominant in the literature. Such estimators are known as "kernel estimators of the failure rate," and it appears that their consideration is motivated by the popularity and success of kernel estimators of the (underlying) density function.

The choice of the kernel is very important, and to a large extent determines the properties of the estimators. Watson and Leadbetter (1964b), in an insightful and fundamental paper on this topic, introduced a sequence of functions $\{\delta_n(x)\}$, which they called a δ-*function sequence*, which satisfies the following conditions:

(a) $\delta_n \in L^1$ (i.e., $\int |\delta_n(x)| dx < \infty$) , for all n ;

(b) $\int \delta_n(x) dx = 1$, for all n ;

- 12 -

(c) $\delta_n(x) \to 0$, uniformly in $|x| \geq \lambda$, for some fixed $\lambda > 0$,

(d) $\int_{|x| \geq \lambda} \delta_n(x) dx \to 0$ as $n \to \infty$ for any fixed $\lambda > 0$.

In summary, $\{\delta_n(x)\}$ is a sequence of smoothing functions (kernels) tending, as $n \to \infty$, to a "Dirac-δ function." Using the sequence $\{\delta_n(x)\}$, they consider the following two estimators of h :

$$\hat{h}_n(x) = \frac{\hat{f}_n(x)}{1 - \hat{F}_n(x)} \qquad (5.1)$$

and

$$\tilde{h}_n(x) = \sum_{j=1}^{n} \frac{\delta_n(x - X_{(j)})}{n-j+1} \qquad (5.2)$$

where

$$\hat{f}_n(x) = \frac{1}{n} \sum_{j=1}^{n} \delta_n(x - X_{(j)}) \qquad \text{and} \qquad \hat{F}_n(x) = \int_0^x \hat{f}_n(u) du \ .$$

If the sequence $\{\delta_n(x)\}$ is suitably chosen, that is, if $\alpha_n = \int \delta_n^2(x) dx < \infty$ at every point of continuity x of h at which $F(x) < 1$, then $\hat{h}_n(x)$ is shown to be asymptotically unbiased with an asymptotic variance given by

$$\text{Var}[\hat{h}_n(x)] \sim \frac{\alpha_n}{n} \frac{h(x)}{1-F(x)} \ . \qquad (5.3)$$

If in addition to the above $\alpha_n = 0(n)$, then the asymptotic variance $\text{Var}[\hat{h}_n(x)]$ converges to zero in the order of α_n/n ; that is, $\hat{h}_n$ is consistent.

Under the following slightly more restrictive conditions,

(e) $\alpha_n = \int \delta_n^2(x) dx < \infty$, $\int |\delta_n(x)|^{2+\eta} dx < \infty$, for some $\eta > 0$ and

such that $\dfrac{\alpha_n}{n^{\eta/2} \cdot \alpha_n^{1+\eta/2}} \to 0$, as $n \to \infty$;

- 13 -

(f) $\left(\frac{n}{\alpha_n}\right)^{\frac{1}{2}}$ $\{E[\hat{F}_n(x)] - F(x)\} \to 0$, as $n \to \infty$; and

(g) $\left(\frac{n}{\alpha_n}\right)^{\frac{1}{2}}$ $\{E[\hat{F}_n(x)] - F(x)\} \to 0$, as $n \to \infty$;

Watson and Leadbetter (1964b) show that the random variable

$$[1-F(x)]\left[\frac{n}{\alpha_n f(x)}\right]^{\frac{1}{2}}[\hat{h}_n(x) - h(x)] \tag{5.4}$$

is asymptotically normally distributed with mean zero and variance 1, at every continuity point x of $h(x)$.

In contrast to the asymptotic results for the estimator $\hat{h}_n$, small sample properties of $\tilde{h}_n$ have been obtained by Watson and Leadbetter (1964b), who show that

$$E[\tilde{h}_n(x)] = \int \delta_n(x-u)h(u)du - \int \delta_n(x-u)F^n(u)h(u)du , \tag{5.5}$$

and

$$\mathrm{Var}[\tilde{h}_n(x)] = \int \delta_n^2(x-u)h(u)I_n(F(u))du +$$

$$+ 2 \int\int_{0\leq u\leq v} \frac{\delta_n(x-u)\delta_n(x-v)}{1-F(v)} \left[\frac{1-F^n(u)}{1-F(u)} F^n(v) - \frac{F^n(v)-F^n(u)}{F(v)-F(u)}\right] dF(u)dF(v) ,$$

where

$$I_n(F) = \int_0^{1-F} \frac{(F+B)^n-F^n}{B} dB .$$

To study the large sample properties of $\tilde{h}_n$, a further condition is required which restricts the class of distributions considered. Specifically, for a given δ-function sequence, a class C_δ of distribution functions F is defined such that for any fixed x_0 and any fixed $\lambda > 0$, there exists a G_λ such that

$$\frac{|\delta_n(x-x_0)|}{1-F(x)} \leq G_\lambda$$

for all sufficiently large n , and $|x-x_0| \geq \lambda$.

Under the assumption that $F \epsilon C_\delta$, and h is continuous at x_0 , with $F(x_0) < 1$, one can show that $\tilde{h}_n(x_0)$ is an asymptotically unbiased estimator of $h(x_0)$. If, in addition, the δ-function sequence is chosen such that $\alpha_n = \int \delta_n^2(x)dx < \infty$, and $\alpha_n = O(n)$, then

$$\lim_{n \to \infty} \left(\frac{n}{\alpha_n}\right) \text{Var}[\tilde{h}_n(x_0)] = \frac{h(x_0)}{1-F(x_0)} .$$

The assumption $\alpha_n = O(n)$ implies in particular that $\text{Var}[\tilde{h}_n(x_0)] \to 0$; thus $\tilde{h}_n$ is consistent, and in fact its variance converges to 0 in exactly the same way as α_n/n (i.e., $\text{Var}[\tilde{h}_n(x_0)] = O(\alpha_n/n)$). A comparison of this result, with that given by (5.3), suggests that

$$\text{Var}[\tilde{h}_n(x_0)] \sim \text{Var}[\hat{h}_n(x_0)] .$$

In Watson and Leadbetter (1964a), some numerocal investigations were conducted with the main conclusion that in practice, the estimators $\hat{h}$ and $\tilde{h}$ are essentially equivalent.

Murthy (1965) proposes two estimators Z_n and Z_n^* of h by considering $\delta_n(x) = B_n K(B_n x)$, where B_n is a nonincreasing function of n such that $\lim_{n \to \infty} B_n = \infty$ and $\lim_{n \to \infty} B_n/n = 0$, and the kernel K satisfies the conditions

$$K(x) \geqq 0 , \quad K(x) = K(-x) , \quad \lim_{|x| \to \infty} xK(x) = 0 , \quad \text{and} \quad \int K(x)dx = 1 .$$

If

$$f_n(x) = B_n \int K\left(B_n(u-x)\right) dF_n(u) = \frac{B_n}{n} \sum_{j=1}^{n} K\left(B_n(X_j-x)\right)$$

and

$$R_n^*(x) = \int_x^\infty f_n(u)du ,$$

then

$$Z_n(x) = f_n(x)/R_n(x) \quad \text{and} \quad Z_n^*(x) = f_n(x)/R_n^*(x) \ .$$

If x_0 is a continuity point of both F and f, then the estimators $Z_n(x_0)$ and $Z_n^*(x_0)$ are consistent and asymptotically normally distributed.

Recent literature on the topic of this survey has received its impetus from the work of Rice and Rosenblatt (1976), who propose three estimators, $h_n^{(1)}$, $h_n^{(2)}$, and $h_n^{(3)}$ of h. These estimators are similar to those of Watson and Leadbetter, and Murthy, except for the fact that the *modified sample distribution function* (the usual sample distribution function multiplied by $n/(n+1)$) F_n is considered. They consider in good detail both the bias and the covariance properties of expressions which agree with these estimators up to a term which is negligible in order of magnitude with large probability.

Rice and Rosenblatt let $H(x) = -\ln(1-F(x))$, $H_n(x) = -\ln(1-F_n(x))$, and denote their δ-function sequence by $\{\omega_n(x)\}$, where $\omega_n(x) = \frac{1}{b_n} \omega\left(\frac{x}{b_n}\right)$, with ω a kernel which is bounded, *band-limited* (i.e., $\omega(x) = 0$, $|x| \geq A$ for some positive constant A), symmetric, and of integral one, and $b_n \downarrow 0$ with $nb_n \to \infty$, as $n \to \infty$. An estimate of the underlying density function f is given by

$$f_n(x) = \int \omega_n(x-u)dF_n(u) \ .$$

Their estimates of the failure rate function $h(x)$ are

$$h_n^{(1)}(x) = \frac{f_n(x)}{1-F_n(x)} \ ,$$

$$h_n^{(2)}(x) = \int \omega_n(x-u) \frac{dF_n(u)}{1-F_n(u)} = \sum_{j=1}^{n} \omega_n(x-X_{(j)}) \frac{1}{n-j+1} \ ,$$

and

- 16 -

$$h_n^{(3)}(x) = \int \omega_n(x-u)dH_n(u) = \sum_{j=1}^{n} \omega_n(x - X_{(j)}) \ln\left[1 + \frac{1}{n-j+1}\right] .$$

For the estimate $h_n^{(1)}(x)$, Rice and Rosenblatt show that if $F \in C^1$ (i.e., F is continuously differentiable), then

$$h_n^{(1)}(x) = a_n(x)\left(1 + O_p\left(\frac{1}{\sqrt{n}}\right)\right),$$

where

$$a_n(x) = \frac{f_n(x)}{1-F(x)} .$$

Furthermore, if $F \in C^2$, then

$$E\left(a_n(x)\right) = h(x) + \frac{b_n^2}{2} \frac{f''(x)}{1-F(x)} \int u^2 \omega(u)du + O(b_n^2) ;$$

also, $a_n(x)$ is asymptotically normally distributed.

Thus, the leading bias term of $a_n(x)$, where $a_n(x)$ agrees with $h_n^{(1)}(x)$ up to a term which is negligible in order to magnitude with large probability, is proportional to $f''(x)/(1-F(x))$.

For the estimate $h_n^{(2)}(x)$, if $F \in C^2$, and if x is such that $F(x) > 0$, then

$$E\left[h_n^{(2)}(x)\right] = h(x) + \frac{b_n^2}{2} h''(x) \int u^2 \omega(u)du + O(b_n^2) ,$$

implying that the leading bias term of $h_n^{(2)}(x)$ is proportional to $h''(x)$.

In order to study the estimator $h_n^{(3)}(x)$, Rice and Rosenblatt consider the difference between $h_n^{(2)}(x)$ and $h_n^{(3)}(x)$, and show that it is negligible in order of magnitude with large probability. That is,

- 17 -

$$h_n^{(2)}(x) - h_n^{(3)}(x) = O_p\left(\frac{1}{n}\right) .$$

Thus, like $h_n^{(2)}(x)$, the leading bias term of $h_n^{(3)}(x)$ is also proportional to $h''(x)$.

Using the fact that the leading bias term of $h_n^{(1)}(x)$ is proportional to $f''(x)/(1-F(x))$, it can be shown that the bias of $h_n^{(2)}(x)$ and $h_n^{(3)}(x)$ is larger if $f'(x) = 0$ and $f''(x) > 0$, or if f is almost constant near x ; whereas the bias of $h_n^{(1)}(x)$ is larger if $h'(x) = 0$ and $h''(x) > 0$, or if h is almost constant near x .

Rice and Rosenblatt also consider the interesting case of the observations $X_1, \ldots, X_n$ being dependent. Specifically, if the observations constitute a stationary time series which satisfies some mixing (regularity) conditions, then the estimates $h_n^{(1)}$, $h_n^{(2)}$, and $h_n^{(3)}$ are asymptotically equivalent to $a_n(x) = f_n(x)/(1-F(x))$.

The main contribution of Rice and Rosenblatt (1976) is to obtain a global result on the estimation of the failure rate, specifically on the maximal weighted deviation between an estimate of h , and h . They have accomplished this by directly applying the results of Bickel and Rosenblatt (1973) on the estimation of the density function f [strengthened by Rosenblatt (1976) using the recent results of Komlos, Major, and Tusnady (1975)]. Specifically, they obtain the asymptotic distribution of the global weighted deviation of $h_n^{(1)}(x)$ from $h(x)$, on any finite interval whose length grows with the sample size n and which diverges (tends to infinity) as n tends to infinity.

Consider a sequence $\{\ell(n)\}$, where $\ell(n)$ denotes the length of a finite interval, and where $\ell(n) \to \infty$ as $n \to \infty$, but in such a way that $\log \ell(n) = O(n)$, and let $b_n = n^{-\delta}$, for $0 < \delta < \frac{1}{2}$. Let

- 18 -

$$\bar{M}_n = \max_{|x| \le \ell(n)} \left| \left(nb_n f^{-1}(x)\right)^{\frac{1}{2}} (1-F(x)) \left(h_n^{(1)}(x) - h(x)\right) \right|$$

denote the maximal weighted deviation of $h_n^{(1)}(x)$ from $h(x)$, weighted by $(nb_n f^{-1}(x))^{\frac{1}{2}}(1-F(x))$, over the interval $\ell(n)$. Under some regularity conditions on $\ell(n)$, and the additional assumption that

$$\sup_{|x| \le \ell(n)} (1-F(x))^{-1} = o(n^{\frac{1}{2}}) ,$$

one has the result that

$$\lim_{n \to \infty} P\left\{ (2 \log c(n))^{\frac{1}{2}} \left(\frac{\bar{M}_n}{(\lambda(\omega))^{\frac{1}{2}}} - d_n \right) \le x \right\} = e^{-2e^{-x}} ,$$

where

$$c(n) = 2\ell(n)/b_n , \qquad \lambda(\omega) = \int \omega^2(u)du ,$$

and

$$d_n = \begin{cases} \left(2 \log c(n)\right)^{\frac{1}{2}} + \dfrac{1}{\left(2 \log c(n)\right)^{\frac{1}{2}}} \left[\log \dfrac{K_1(\omega)}{\sqrt{\pi}} + \dfrac{1}{2} \log \log c(n) \right] , \\[2mm] \qquad \text{if} \quad K_1(\omega) = \dfrac{\omega^2(A) + \omega^2(-A)}{2\lambda(\omega)} > 0 ; \\[4mm] \left(2 \log c(n)\right)^{\frac{1}{2}} + \dfrac{1}{\left(2 \log c(n)\right)^{\frac{1}{2}}} \left[\log \dfrac{1}{\pi} \dfrac{K_2(\omega)}{2} \right] \\[2mm] \qquad \text{if} \quad K_1(\omega) \le 0 , \text{ with } K_2(\omega) = \left| \dfrac{\int [\omega^1(u)]^2 du}{\lambda(\omega)} \right|^{\frac{1}{2}} . \end{cases}$$

The above global result can be used to construct uniform confidence bands for the failure rate function $h(x)$ using the estimator $h_n^{(1)}(x)$. A $100\varepsilon\%$ uniform confidence band for $h(x)$, for $|x| \le \ell(n)$ is given by

- 19 -

$$h_n^{(1)}(x) \quad \left(nb_n f^{-1}(x)\right)^{\frac{1}{2}} (1 - F(x)) \left[\gamma + \log c(n)\right]^{\frac{1}{2}} + d_n \Big] .$$

with $\gamma = -\log\left(-\dfrac{\log \varepsilon}{2}\right)$.

Rice and Rosenblatt undertake some simulations to give some indication of the performance of the estimator $h_n^{(1)}$ and its limiting distribution, and claim the results to be encouraging for large sample sizes with narrow bandwidths.

Global results for the estimators $h_n^{(2)}$ and $h_n^{(3)}$ and the resulting uniform confidence bands have been obtained by Sethuraman and Singpurwalla (1981) via considerations of a "naive estimator" of h . This estimator, which is shown to be uniformly asymptotically equivalent to $h_n^{(2)}$ and $h_n^{(3)}$, will be discussed later, in Section 6.

5.1 Improvement of Kernel Estimates--Use of the Generalized Jackknife

Singpurwalla and Wong (1980a) attempt to reduce the bias and improve upon the rate of convergence of the mean square error (MSE) of kernel estimators by considering estimators of the type

$$h_n(x_0) = \frac{1}{b_n} \sum_{j=1}^{n} \frac{1}{n-j+1} K\left(\frac{X_{(j)} - x_0}{b_n}\right) ,$$

where the sequence $\{b_n\}$ is such that $b_n \downarrow 0$ and $nb_n \to \infty$, as $n \to \infty$, and K is a bounded and symmetric kernel of integral one with the additional properties that $\int |K(x)| dx < \infty$ (i.e., $K \in L^1$) , and $\lim_{x \to \infty} |xK(x)| = 0$. Note that $h_n(x_0)$ is identical to (5.2), the estimator $\tilde{h}_n$ of Watson and Leadbetter (1964a, 1964b), and the estimator $h_n^{(2)}$ of Rice and Rosenblatt (1976).

The authors first prove a "pointwise saturation theorem," which says that for the class of nonnegative L^1 kernels K, and with the additional requirements that $x^2 K \in L^1$, $\lim_{n \to \infty} nb^2(n) = \infty$, $h \in C^2$, and $h^{(3)}$ (the third derivative of h) being uniformly bounded, the rate of convergence of the MSE of $h_n(x_0)$ is *at most* $O(n^{-4/5})$, regardless of the smoothness of h. If the condition that the kernel K be nonnegative is relaxed, that is, if the kernel is allowed to take some negative values, then the bias contribution to the asymptotic MSE can, in principle, be eliminated to any desired order, and the rate of convergence of the asymptotic MSE given by $O(n^{-2m/(2m+1)})$ for some $m \geq 3$, can be brought as close to n^{-1} as is desired. It is shown that an indefinite use of the "generalized jackknife method" of Gray and Schucany (1972) can be used to achieve this goal. Furthermore, by eliminating the requirement that $K \in L^1$, one is able to consider kernels for which the rate of conveggence of the MSE of $h_n(x_0)$ is not as slow as $O(n^{-2m/(2m+1)})$. One possible non-L^1 kernel is the "sinc function kernel," $K(x) = \frac{\sin x}{\pi x}$, considered by Singpurwalla and Wong (1980b). If the Fourier transform of h decreases exponentially, then the asymptotic MSE of $h_n(x_0)$ using the sinc function kernel decreases at the rate $\log n/n$, which is very close to the ideal rate of n^{-1}, which cannot be achieved in practice.

6. The Naive Estimator

The "naive estimator*" of h was originally proposed by Grenander (1956b), and has been considered by Marshall and Proschan (1965) and by Barlow, Bartholomew, Bremner, and Brunk (1972, Section

---

*Naming the estimator naive is due to Frank Proschan.

5.3) for estimating a monotonically increasing (or decreasing) function h . It has been used by Singpurwalla (1975) for performing a time series analysis of failure data.

Following the notation of Section 4.1, let X_n^0 , $n=1,\ldots,N$, be the true survival time of the nth item, assumed to be independent identically distributed random variables having a common but unknown distribution F^0 . Let Y_n , $n=1,\ldots,N$, be the withdrawal time of the nth item, which may or may not have a joint distribution H , but which must be independent of the X_n^0's . As before, let

$$X_n = \min(X_n^0, Y_n) , \quad n=1,\ldots,N ,$$

and interest centers arount the distribution function F^0 .

Suppose that k *failures*, $(k \leq N)$ have been observed in all, and let

$$0 \equiv Z_{(0)} \leq Z_{(1)} \leq \cdots \leq Z_{(k)}$$

be the *ordered failure* times. Note that X_n is a failure time whenever $X_n = X_n^0$. The *total time on test at time* t , $T_N(t)$, is defined as

$$T_N(t) = \int_0^t N(u)\,du ,$$

where $N(u)$ is the number of items which are surviving at time u .

The *naive estimator* of $h(z)$, $h_N^*(z)$, is defined as

$$h_N^*(z) = \begin{cases} \dfrac{1}{T_N(Z_{(i)}) - T_N(Z_{(i-1)})} , & Z_{(i-1)} \leq z \leq Z_{(i)} , \; i=1,\ldots,k \\ \\ 0 , & z \geq Z_{(k)} . \end{cases}$$

Note that $h_N^*(z)$ is the reciprocal of the total time on test in the interval $(Z_{(i-1)}, Z_{(i)}]$.

- 22 -

When $k = N$, that is, when there is no censoring, $Z_{(i)} = X^0_{(i)}$, $\forall i$, where $X^0_{(1)} \leq X^0_{(2)} \leq \ldots \leq X^0_{(N)}$ are the order statistics of $X^0_1, \ldots, X^0_N$. In this case, for $X^0_{(i-1)} \leq z < X^0_{(i)}$, the naive estimator becomes

$$h^*_N(z) = \{\text{Total time on test in } (X^0_{(i-1)}, X^0_{(i)}]\}^{-1}$$

$$= [(n-i+1)(X^0_{(i)} - X^0_{(i-1)})]^{-1} .$$

Sethuraman and Singpurwalla (1980) show that the above estimator is asymptotically unbiased, but is *not consistent*, since it has a limiting nondegenerate distribution. Furthermore, since for any m distinct time points $z_1, \ldots, z_m$ the estimators $\{h^*_N(z_1), \ldots, h^*_N(z_m)\}$ are asymptotically independent, the graph of $\{h^*_N(z), z \geq 0\}$ will exhibit wild fluctuations, prohibiting the use of h^*_N as an estimator of h. This behavior of h^*_N is analogous to the behavior of the sample periodogram which is used to estimate the spectrum of a stationary time series. Thus, following the standard technique of "smoothing" the sample periodogram, Sethuraman and Singpurwalla smooth the naive estimator h^*_N using a bandlimited kernel, to obtain a consistent and an asymptotically normal estimate of the failure rate.

The *smoothed estimator* $\bar{h}^*_N(z)$ obtained by smoothing the naive estimator $h^*_N(z)$ using a kernel of bandwidth $2b_N A$ is, for $z \geq b_N A$, given by

$$\bar{h}^*_N(z) = \frac{1}{b_N} \int \omega\left(\frac{x-s}{b_N}\right) h^*_N(s) ds ,$$

where the kernel ω and the sequence $\{b_N\}$ have the properties required by Rice and Rosenblatt (1976) (see Section 5).

The estimate $\bar{h}_N^*(z)$ is consistent, and if the first derivative ω' of ω is such that $\omega' \in L^1$, then $\bar{h}_N^*(z)$ is uniformly close to $h_n^{(3)}(z)$ of Rice and Rosenblatt on bounded intervals of length K, where K is such that $K+A \leqq X_{(N)}^0$. Furthermore, under some regularity conditions on F, $\bar{h}_N^*(z)$ can be approximated by an appropriate Gaussian process. Under the additional requirements that $h \in C^2$, $(\omega'(z))^2 \in L^1$, and that $Nb_N^5 \log b_n \downarrow 0$ as $n \to \infty$, the asymptotic distribution of the global deviation of $\bar{h}_N^*(z)$ from $h(z)$ can be obtained on any finite interval $[b_N A, K]$, in much the same way as in Rice and Rosenblatt [see Theorems 2.2 and 2.4 of Sethuraman and Singpurwalla (1981)]. These results can be used to construct the $100\varepsilon\%$ confidence bands for $h(z)$, for $b_N A \leqq z \leqq K$; these are

$$\bar{h}_N^*(z) \quad \pm \quad \left(\frac{\bar{h}_N^*(z)}{Nb_N \bar{F}_N(z)} \right)^{\frac{1}{2}} \left(\frac{\gamma}{b_N} + \delta_N \right) , \qquad (6.1)$$

where $\gamma = -\log(-(\frac{1}{2})\log\varepsilon)$, $\bar{F}_N = 1 - F_N$, with F_N being the empirical distribution function of $X_{(1)}^0 \leqq X_{(2)}^0 \leqq \cdots \leqq X_{(N)}^0$, and for $C_N = (2\log(K/b_N))^{\frac{1}{2}}$,

$$\delta_N = \begin{cases} (\lambda(\omega))^{\frac{1}{2}} \left[C_N^2 + \log\left(C_N \frac{K_1(\omega)}{\sqrt{2\pi}} \right) \right] / C_N , \\ \qquad\qquad\qquad \text{if } K_1(\omega) > 0 , \\[2ex] (\lambda(\omega))^{\frac{1}{2}} \left[C_N^2 + \log\left(\frac{K_2(\omega)}{\pi} \right) \right] / C_n , \\ \qquad\qquad\qquad \text{if } K_1(\omega) \leqq 0 ; \end{cases}$$

$\lambda(\omega)$, $K_1(\omega)$, and $K_2(\omega)$ have been defined in Section 5.

Since $\bar{h}_N^*(z)$ is uniformly close to $h_n^{(3)}(z)$, and since $h_n^{(3)}(z) - h_n^{(2)}(z) = O_p(1/n)$, the $100\varepsilon\%$ *uniform confidence bands* on $h(z)$ can also be obtained by replacing $\bar{h}_N^*(z)$ in Equation (6.1) by the Rice and Rosenblatt estimators $h_n^{(2)}(z)$ or $h_n^{(3)}(z)$.

6.1 Random Smoothing of the Naive Estimator

It is to be emphasized that the results of Watson and Leadbetter (1964a, 1964b), Murthy (1965), Rice and Rosenblatt (1976), and Sethuraman and Singpurwalla (1980) apply only when there is no censoring, when the sample size becomes large, and under a host of regularity conditions on h. The above conditions are not always easy to satisfy in practice, and so in an attempt towards eliminating them, Miller and Singpurwalla (1980) consider a *randomly smoothed estimator* $h_{N,j}$ of h, defined, for some j and $i=j, j+1, \ldots, k$, as

$$h_{N,j}^*(z) = \begin{cases} [T_N(Z_{(i)}) - T_N(Z_{(i-j)})]^{-1} & , \quad Z_{(i-1)} < z \leq Z_{(i)} \\ 0 & , \quad z > Z_{(k)} \; ; \end{cases}$$

note that when $j=1$, $h_{N,j}^* = h_N^*$, the naive estimator.

Alternate expressions for $h_{N,j}^*$ in terms of h_N^*, for $i=j, j+1, \ldots, k$, are

$$\frac{1}{h_{N,j}^*(z)} = \frac{1}{j} \sum_{m=0}^{j-1} \frac{1}{h_N^*(Z_{(i-m)})} \quad , \quad Z_{(i-1)} < z \leq Z_{(i)} \; ,$$

$$h_{N,j}^*(z) = \frac{\displaystyle\int_{Z_{(i-j)}}^{Z_{(i)}} h_N^*(u) dT_N(u)}{\displaystyle\int_{Z_{(i-j)}}^{Z_{(i)}} dT_N(u)} \quad , \quad Z_{(i-1)} < z \leq Z_{(i)} \; .$$

Since the naive estimator h_N^* is constant over a failure inter-val, it can be viewed as some kind of smoothing over a single failure interval. If this notion of smoothing is extended to cover j failure intervals, then $h_{N,j}^*$ results; this is a motivation for $h_{N,j}^*$.

A reason for considering $h_{N,j}^*$ is that it leads to an *exact* distribution-free confidence region in the case of *finite* samples, with censoring. Also, its asymptotic theory does not involve the usual regu-larity conditions. Furthermore, $h_{N,j}^*(z)$ is a consistent estimator of $f(z)$ at every continuity point z of h . A disadvantage of $h_{N,j}^*$ is that it yields confidence limits on a *smoothed version* of h , $h_{T_{N,j}}^*$, where

$$h_{T_{N,j}}^*(z) = \frac{\int_{Z_{(i-j)}}^{Z_{(i)}} h(u)dT_N(u)}{T_N(Z_{(i)}) - T_N(Z_{(i-j)})} , \quad Z_{(i-1)} < z \le Z_{(i)} ,$$

rather than h itself. This disadvantage is of concern only if h changes too rapidly.

Specifically, a $100\varepsilon\%$ upper confidence bound for $h_{T_{N,j}}^*(z)$, for $Z_{(i-1)} < z \le Z_{(i)}$, is

$$h_{N,j}^*(z) \left(1 + \frac{C_{j,k,\varepsilon}^+}{j} \right) ,$$

where $C_{j,k,\varepsilon}^+$ is the critical value such that

$$P \left\{ \sup_{j \le i \le k} \sum_{j=i-j+1}^{i} (U_m - 1) \le C_{j,k,\varepsilon}^+ \right\} = 1-\varepsilon ,$$

and where the U_i's are independent and identically distributed random variables with density e^{-u} .

- 26 -

The lower and the two-sided confidence bounds for $h^*_{T_{N,j}}$ can be

similarly obtained, in terms of $C^-_{j,k,\varepsilon}$ and $C_{j,k,\varepsilon}$, respectively,

where $C^-_{j,k,\varepsilon}$ and $C_{j,k,\varepsilon}$ are the infimum and the supremum of the

absolute value of $\sum_{m=i-j+1}^{i} (U_m - 1)$, respectively.

At present there do not exist any analytical methods for comput-

ing $C^+_{j,k,\varepsilon}$, $C^-_{j,k,\varepsilon}$, and $C_{j,k,\varepsilon}$; however, these can be easily ob-

tained via a Monte Carlo method, for any desired values of j, k, and

ε. The need for such a Monte Carlo for every required combination of

j, k, and ε could be viewed as another disadvantage of using $h^*_{N,j}$.

Miller and Singpurwalla (1980) also develop the asymptotic theory

for $h^*_{N,j}$, as $N \to \infty$ with $j/N \to \ell$, $0 < \ell < 1$, and present the confi-

dence bounds for $h^*_{T_{N,j}}$ based on this theory. However, in the light of

the results of Watson and Leadbetter and Sethuraman and Singpurwalla,

there are no advantages in using these bounds, and so they are not

discussed here.

7. **The Generalized Failure Rate Function and Its Estimation**

Barlow and Van Zwet (1970) generalize the definition of the fail-

ure rate function by considering a known absolutely continuous distribu-

tion function G with density function g, and define the *generalized

failure rate function* of F, for all $x > 0$ such that $g[G^{-1}F(x)] > 0$,

by

$$r(x) = \frac{f(x)}{g[G^{-1}F(x)]}.$$

Note that when G is taken to be the exponential distribution with

parameter $r=1$, then $r(x) = h(x)$, the failure rate function given by

(2.1), whereas if G is taken to be the uniform distribution on $[0,1]$, then $r(x) = f(x)$, the density function of F. A motivation for considering the generalized failure rate function is not within the scope of this survey, and is given by Barlow, Bartholomew, Bremner, and Brunk (1972). For now it suffices to know that h is a special case of r.

Ahmad (1976) considers a kernel estimate of $f(x)$,

$$\hat{f}(x) = \frac{1}{b_n} \int K\left(\frac{x-u}{b_n}\right) dF_n(u) = \frac{1}{nb_n} \sum_{j=1}^{n} K\left(\frac{x-X_{(j)}}{b_n}\right) ,$$

where the kernel K and the sequence $\{b_n\}$ satisfy some regularity conditions. He introduces the following kernel estimate of $r(x)$,

$$\hat{r}(x) = \frac{\hat{f}(x)}{g[G^{-1}(\hat{F}(x))]} ,$$

where $\hat{F}(x) = \int_{-\infty}^{x} \hat{f}(u) du$, and shows that $\hat{r}(x)$ is consistent.

Shaked (1978) proposes the following estimator of $r(x)$, for any fixed x :

$$\rho_n(x) = \frac{F_n\left(x + \frac{1}{2} C_n\right) - F_n\left(x - \frac{1}{2} C_n\right)}{C_n g G^{-1}\left\{\frac{1}{n}\left[\frac{1}{C_n} \sum_{j \in R(x)} \left(X_{(j)} - x + \frac{1}{2} C_n\right) + M(x)\right]\right\}} ,$$

where C_n is a positive constant, $R(x) = \left\{i : x - \frac{1}{2} C_n \le X_{(i)} \le x + \frac{1}{2} C_n\right\}$, and $M(x) = $ number of $X_{(i)} : X_{(i)} \ge x + \frac{1}{2} C_n$.

To gain some insight into $\rho_n(x)$, note that if G is the exponential distribution, then the estimator is simply the ratio of the (normalized) number of failures in the interval $[x - \frac{1}{2} C_n , x + \frac{1}{2} C_n]$ to the total time on test in the interval.

When $C_n = cn^{-a}$, where $c > 0$ and $0 < a < 1$, and if r is smooth in the neighborhood of x, Shaked (1978) shows that $\rho_n(x)$ is asymptotically normal. In particular, for $1/5 < a < 1$,

$$\frac{\sqrt{cf(x)}}{r(x)} n^{(1-a)/2} \left(\rho_n(x) - r(x) \right)$$

has an asymptotic standard normal distribution, whereas when $1/7 < a \leq 1/5$,

$$\frac{\sqrt{cf(x)}}{r(x)} n^{(1-a)/2} \left[\rho_n(x) - r(x) - \frac{C_n^2 r(x)}{24f(x)} \left(f''(x) - \frac{f^1(x)r^2(x)g^1(G^{-1}F(x))}{f(x)} \right) \right]$$

has an asymptotic standard normal distribution. Note that the above expression is an extension of (5.4) applied to $\rho_n(x)$.

Barlow and Van Zwet (1969, 1970) and Shaked (1978) also consider methods of estimating $r(x)$ when some monotonicity conditions are imposed on it. However, these are not given here, since they do not belong to the purview of this survey.

Acknowledgment

BIBLIOGRAPHY

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

AHMAD, I. A., and PIN-ERH LIN (1977). Nonparametric estimation of a vector-valued bivariate failure rate. *Ann. Statist.* 5 1027-1038.

ARUNKUMAR, S. (1969). Estimation of the change point of the generalized failure rate function. Research Report ORC 69-16, University of California, Berkeley.

BAIN, L. J. (1978). *Statistical Analysis of Reliability and Life Testing Models.* Marcel Dekker, New York.

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

BARLOW, R. E., and R. CAMPO (1975). Total time on test processes and applications to failure data analysis. In *Reliability and Fault Tree Analysis* (R. E. Barlow, J. B. Fussell, and N. D. Singpurwalla, eds.). SIAM, Philadelphia.

BARLOW, R. E., and F. PROSCHAN (1965). *Mathematical Theory of Reliability.* Wiley, New York.

BARLOW, R. E., and F. PROSCHAN (1969). A note on a test for monotone failure rate based on incomplete data. *Ann. Math. Statist.* 40 595-600.

BARLOW, R. E., and F. PROSCHAN (1975). *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart, and Winston, New York.

BARLOW, R. E., and W. R. VAN ZWET (1969). Asymptotic properties of isotonic estimators for the generalized failure rate function, Part

II: Asymptotic distribution. Research Report ORC 69-10, University of California, Berkeley.

BARLOW, R. E., and W. R. VAN ZWET (1970). Asymptotic properties of isotonic estimators for the generalized failure rate function, Part I: Strong consistency. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.), 159-173. Cambridge University Press.

BARLOW, R. E., and W. R. VAN ZWET (1971). Comparison of several nonparametric estimators of the failure rate function. In *Operations Research and Probability* (D. Grouchko, ed.), 375-399. Gordon and Breach, New York.

BERKSON, J. (1954). Estimation of the interval rate in actuarial calculations (summary). *J. Amer. Statist. Assoc.* $\underset{\sim}{47}$ 363.

BERKSON, J., and R. P. GAGE (1950). Calculation of survival rates for cancer. *Proc. Mayo Clinic* $\underset{\sim}{25}$ 270-286.

BICKEL, P. J., and M. ROSENBLATT (1973). On some global measures of the derivations of density function estimates. *Ann. Statist.* $\underset{\sim}{1}$ 1071-1095.

BRESLOW, N., and J. CROWLEY (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* $\underset{\sim}{2}$ 437-453.

CHANDRA, M. (1977). Small sample behavior of a smoothed estimator of the failure rate function--a Monte Carlo investigation. Technical Memorandum Serial TM-66506, Institute for Management Science and Engineering, The George Washington University.

CUTLER, S. J., and F. EDERER (1958). Maximum utilization of the life table method in analyzing survival. *J. Chron. Dis.* $\underset{\sim}{8}$ 699-712.

GEHAN, E. (1969). Estimating survival functions from the life table. *J. Chron. Dis.* 21 629-644.

GRAY, H. L., and W. R. SCHUCANY (1972). *The Generalized Jackknife Statistics.* Marcel Dekker, New York.

GRENANDER, U. (1956a). On the theory of mortality measurement, I. *Skand. Aktuarietidskr.* 39 90-96.

GRENANDER, U. (1956b). On the theory of mortality measurement, II. *Skand. Aktuarietidskr.* 39 125-153.

HALL, W. J., and J. A. WELLNER (1980). Confidence bands for survival curves from censored data. *Biometrika* 67 133-143.

KAPLAN, E. L., and P. MEIER (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 53 457-481.

KIMBALL, A. W. (1960). Estimation of mortality intensities in animal experiments. *Biometrics* 16 505-521.

KOMLES, J., P. MAJOR, and G. TUSNADY (1975). An approximation of partial sums of independent random variables and the sample distribution function. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* 32 111-131.

MANN, N. R., R. E. SCHAFER, and N. D. SINGPURWALLA (1970). *Methods for Statistical Analysis of Reliability and Life Data.* Wiley, New York.

MARSHALL, A. W., and F. PROSCHAN (1965). Maximum likelihood estimation for distributions with monotone failure rate. *Ann. Math. Statist.* 36 69-77.

MILLER, D. R., and N. D. SINGPURWALLA (1980). Failure rate estimation using random smoothing. *Sankhyā*, to appear.

MILLER, R. G. (1974). The jackknife--a review. *Biometrika* 61 1-15.

MORAN, P. A. P. (1951). Estimation methods for evolution processes. *J. Roy. Statist. Soc. Ser. B* 13 141-146.

MURTHY, V. K. (1965). Estimation of jumps, reliability and hazard rate. *Ann. Math. Statist.* 36 1032-1040.

PARZEN, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* 33 1065-1076.

PETERSON, A. V. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. *J. Amer. Statist. Assoc.* 72 854-858.

PRAKASA RAO, B. L. S. (1970). Estimation for distributions with monotone failure rate. *Ann. Math. Statist.* 41 507-519.

RICE, J., and M. ROSENBLATT (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.

ROSENBLATT, M. (1971). Curve estimates. *Ann. Math. Statist.* 42 1815-1842.

ROSENBLATT, M. (1976). On the maximal deviation of k-dimensional density estimates. *Ann. Probability* 4 1009-1015.

SACHER, G. A. (1956). On the statistical nature of mortality, with especial reference to chronic radiation mortality. *Radiology* 67 250-257.

SEAL, H. L. (1954). The estimation of mortality and other decremental probabilities. *Skand. Aktuarietidskr.* 37 137-162.

SETHURAMAN, J., and N. D. SINGPURWALLA (1981). Large sample estimates and uniform confidence bounds for the failure rate function based on a naive estimator. *Ann. Statist.*, to appear.

SINGPURWALLA, N. D. (1975). Time series analysis and forecasting of
failure rate processes. In *Reliability and Fault Tree Analysis*
(R. E. Barlow, J. B. Fussell, and N. D. Singpurwalla, eds.).
SIAM, Philadelphia.

SINGPURWALLA, N. D., and MAN-YUEN WONG (1980a). Improvement of kernel
estimators of the failure rate function using the generalized
jackknife. Technical Paper Serial T-415, Institute for Manage-
ment Science and Engineering, The George Washington University.

SINGPURWALLA, N. D., and MAN-YUEN WONG (1980b). Fourier integral esti-
mate of the failure rate function and its mean square error
properties. Technical Paper Serial T-416, Institute for Manage-
ment Science and Engineering, The George Washington University.

SHAKED, M. (1979). An estimator for the generalized hazard rate func-
tion. *Comm. Statist.—Theory and Methods*, A8 17-33.

TAPIA, R. T., and J. R. THOMPSON (1978). *Nonparametric Probability
Density Estimation*. The Johns Hopkins University Press.

TURNBULL, B. W. (1976). The empirical distribution function with arbi-
trarily grouped, censored, and truncated data. *J. Roy. Statist.
Soc. Ser. B* 38 (3) 290-295.

WATSON, G. S., and M. R. LEADBETTER (1964a). Hazard analysis I. *Bio-
metrika* 51 175-184.

WATSON, G. S., and M. R. LEADBETTER (1964b). Hazard analysis II.
Sankhyā Ser. A 26 101-116.

WEGMAN, E. J., E. S. NOUR, and C. KUKUK (1980). A time series approach
to life table construction. *Comm. Statist.*, to appear.

WOODROOFE, U. (1970). On choosing a delta sequence. *Ann. Math. Statist.*
41 1665-1671.

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

JANUARY 1981

	Copies		Copies
Statistics and Probability Program (Code 436) Office of Naval Research Arlington, VA 22217	3	Office of Naval Research Scientific Liaison Group Attn: Scientific Director American Embassy - Tokyo APO San Francisco 96503	1
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12	Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1
Office of Naval Research New York Area Office 715 Broadway - 5th Floor New York, New York 10003	1	Commandant of the Marine Corps (Code AX) Attn: Dr. A. L. Slafkosky Scientific Advisor	
Commanding Officer Office of Naval Research Eastern/ Central Regional Office Attn: Director for Science		Washington, DC 20380	1
666 Summer Street Boston, MA 02210	1	Director National Security Agency Attn: Mr. Stahly and Dr. Maar (R51) Fort Meade, MD 20755	2
Commanding Officer Office of Naval Research Western Regional Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1	Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1
Commanding Officer Office of Naval Research Branch Office Attn: Director for Science 536 South Clark Street Chicago, Illinois 60605	1	U. S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1

THE GEORGE WASHINGTON UNIVERSITY

BENEATH THIS PLAQUE
IS BURIED
A VAULT FOR THE FUTURE
IN THE YEAR 2056

THE STORY OF ENGINEERING IN THIS YEAR OF THE PLACING OF THE VAULT AND ENGINEERING HOPES FOR THE TOMORROWS AS WRITTEN IN THE RECORDS OF THE FOLLOWING GOVERNMENTAL AND PROFESSIONAL ENGINEERING ORGANIZATIONS AND THOSE OF THIS GEORGE WASHINGTON UNIVERSITY.

BOARD OF COMMISSIONERS DISTRICT OF COLUMBIA
UNITED STATES ATOMIC ENERGY COMMISSION
DEPARTMENT OF THE ARMY UNITED STATES OF AMERICA
DEPARTMENT OF THE NAVY UNITED STATES OF AMERICA
DEPARTMENT OF THE AIR FORCE UNITED STATES OF AMERICA
NATIONAL ADVISORY COMMITTEE FOR AERONAUTICS
NATIONAL BUREAU OF STANDARDS U S DEPARTMENT OF COMMERCE
AMERICAN SOCIETY OF CIVIL ENGINEERS
AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS
THE SOCIETY OF AMERICAN MILITARY ENGINEERS
AMERICAN INSTITUTE OF MINING & METALLURGICAL ENGINEERS
DISTRICT OF COLUMBIA SOCIETY OF PROFESSIONAL ENGINEERS INC
THE INSTITUTE OF RADIO ENGINEERS INC
THE CHEMICAL ENGINEERS CLUB OF WASHINGTON
WASHINGTON SOCIETY OF ENGINEERS
FAULKNER KINGSBURY & STENHOUSE ARCHITECTS
CHARLES H TOMPKINS COMPANY BUILDERS
SOCIETY OF WOMEN ENGINEERS
NATIONAL ACADEMY OF SCIENCES NATIONAL RESEARCH COUNCIL

THE PURPOSE OF THIS VAULT IS INSPIRED BY AND IS DEDICATED TO
CHARLES HOOK TOMPKINS DOCTOR OF ENGINEERING
BECAUSE OF HIS ENGINEERING CONTRIBUTIONS TO THIS UNIVERSITY TO HIS
COMMUNITY TO HIS NATION AND TO OTHER NATIONS.

BY THE GEORGE WASHINGTON UNIVERSITY.

ROBERT V FLEMING   CLOYD H MARVIN
CHAIRMAN OF THE BOARD OF TRUSTEES   PRESIDENT
JUNE THE ELEVENTH
1956

To cope with the expanding technology, our society must be assured of a continuing supply of rigorously trained and educated engineers. The School of Engineering and Applied Science is completely committed to this objective.