The Annals of Statistics
1986, Vol. 14, No. 4, 1366–1378

NONPARAMETRIC RENEWAL FUNCTION ESTIMATION

BY EDWARD W. FREES

University of Wisconsin

The renewal function is a basic tool used in many probabilistic models and sequential analysis. Based on a random sample of size n, a nonparametric estimator of the renewal function is introduced. Asymptotic properties of the estimator such as consistency and asymptotic normality are developed. A discussion of an application to warranty analysis is also provided.

1. Introduction. Let $X_1, X_2, \dots$ be identically and independently distributed random variables with distribution function F. Assume that F has positive mean μ and finite variance σ^2. With $S_k = X_1 + \cdots + X_k$, let $F^{(k)}(t) = P(S_k \le t)$ be the k-fold convolution of F for $k \ge 1$. The renewal function H is defined by

$$H(t) = \sum_{k \ge 1} F^{(k)}(t) \tag{1.1}$$

for $t > 0$. Although we do not assume here that the data come from realizations of a renewal process, the function H is of interest in renewal theory because $H(t)$ is the expected number of renewals in an interval $(0, t]$ for a renewal process with underlying lifetime distribution F. The renewal function plays an important role in many probabilistic models [cf. Feller (1971) and Karlin and Taylor (1975)] and sequential analysis [cf. Woodroofe (1982)].

Most nonparametric estimators of $H(t)$ are based on a realization of a renewal process and on theorems which yield simple approximations of $H(t)$ for asymptotically large values of time t. For example, suppose that nonnegative observations are recorded and that F has an arithmetic distribution. Recall that a distribution function is said to be arithmetic if its support is on $\{0, \pm d, \pm 2d, \dots\}$ for some constant d and otherwise is nonarithmetic. Then, the result

$$\lim_{t \to \infty} H(t) - t/\mu = \left(\sigma^2 + \mu - \mu^2\right)/\left(2\mu^2\right), \tag{1.2}$$

where the limit goes through multiples of d [cf. Feller (1968), page 341], suggests the use of the estimator

$$\hat{H}(t) = t/\hat{\mu} + \left(\hat{\sigma}^2 + \hat{\mu} - \hat{\mu}^2\right)/\left(2\hat{\mu}^2\right), \tag{1.3}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are estimators of μ and σ^2 based on the data recorded up to time t. See Yang (1983) for an application of $\hat{H}(t)$ to continuous sampling plans. See Cox and Lewis (1966) for an early treatment of the statistical analysis of renewal processes. See Brillinger (1975) and Vardi (1982) for more recent discussions.

In this paper, estimators of $H(t)$ for a fixed time t are based on a random sample of size n, $X_1, X_2, \dots, X_n$. Estimation of the distribution function F and linear functionals of F are problems that have been thoroughly investigated in the literature [cf. Serfling (1980), Chapters 2 and 6]. Viewing $H(t)$ defined in (1.1)

Received January 1985; revised December 1985.
AMS 1980 *subject classifications.* Primary 62G05; secondary 60K05, 62M99.
Key words and phrases. U-statistics, reverse martingales, warranty analysis.

as merely the infinite sum of convolutions of F, it seems natural to estimate $H(t)$ based on a sum of estimator of the convolutions of F. As one would suspect, even though estimators of the type in (1.3) are based on recorded observations, they do not perform well for small (relative to μ) times t. This was pointed out by Frees (1986). In that study the author introduced several estimators, both parametric and nonparametric, of $H(t)$ for a fixed time t based on a random sample of size n. We now define a nonparametric estimator which performed particularly well in the simulation portion of that study. Let $\{i_1, i_2, \dots, i_k\}$ be a subset of size k of $\{1, 2, \dots, n\}$ and let $\sum_c$ be the sum over all $\binom{n}{k}$ distinct combinations of $\{i_1, i_2, \dots i_k\}$. Then, an unbiased estimator of $F^{(k)}(t)$ is

$$F_n^{(k)}(t) = \binom{n}{k}^{-1} \sum_c I(X_{i_1} + \cdots + X_{i_k} \le t). \tag{1.4}$$

Here $I(\cdot)$ is the indicator function of a set. Let $m = m(n)$ be a positive integer depending on n such that $m \le n$ and $m \uparrow \infty$ as $n \uparrow \infty$. Then, a nonparametric estimator of the renewal function is

$$H_n(t) = \sum_{k=1}^{m} F_n^{(k)}(t). \tag{1.5}$$

Often we will simply use $m = n$. Some advantages of introducing the design parameter m are discussed in Section 5.

The estimator of $F^{(k)}(t)$, $F_n^{(k)}(t)$, is a U-statistic and thus it is easy to establish that for each $k \ge 1$ and for each $t \ge 0$ that

$$F_n^{(k)}(t) \to F^{(k)}(t), \quad \text{a.s.}$$

However, asymptotic properties of $H_n(t)$ are not immediate since $H_n(t)$ is not a U-statistic. In Section 2, the almost sure (a.s.) consistency is established by showing that $H_n(t)$ is a reverse martingale with respect to an appropriate sequence of sub σ-fields plus some negligible terms. Also in that section for the case $m = n$, we prove a.s. uniform consistency, the Glivenko–Cantelli property, when the uniformity is restricted to bounded subsets of the positive real line. In Section 3, the asymptotic normality, when properly standardized, of $H_n(t)$ is proved via the projection technique popularized by Hájek [cf. Serfling (1980), Chapter 9.2.5]. To keep potential applications for this estimator as broad as possible, we distinguish between the usual renewal theory assumptions of non-negative observations and the more general framework which also permits negative observations. The latter is the situation usually encountered in sequential analysis. In Section 4, we prove the consistency of an estimator of the asymptotic variance. This provides the important result of large sample interval estimates which is illustrated with an example in warranty analysis. We conclude in Section 5 with some general remarks.

2. Almost sure consistency. Let $a \in R$ and g_a be a real valued function defined on $R^+ = [0, \infty)$ such that

$$\int_0^\infty |g_a(u)|\, d\Big(\sum_{k \ge 1} k^a F^{(k)}(u)\Big) < \infty. \tag{2.1}$$

For each $k \geq 1$, define $m^{-1}(k) = \inf\{n\colon m(n) \geq k\}$. We will also use the assumption

$$\int_0^\infty |g_a(u)|\, d\Big(\sum_{k\geq 1}(m^{-1}(k)-k)k^a F^{(k)}(u)\Big) < \infty. \tag{2.2}$$

In this section we establish the following result.

THEOREM 2.1. *Suppose that* (2.1) *holds. Then,*

$$\int_0^\infty g_a(u)\, d\Big(\sum_{k=1}^{m} k^a F_n^{(k)}(u)\Big) \to \int_0^\infty g_a(u)\, d\Big(\sum_{k\geq 1} k^a F^{(k)}(u)\Big), \tag{2.3}$$

in probability.

Further suppose that (2.2) *holds. Then the convergence in* (2.3) *can be strengthened to a.s. convergence.*

Note that in the case $m = n$, we have $m^{-1}(n) = n$ and the requirement (2.2) is vacuous. In Sections 4 and 5, we give applications where $a \neq 0$. To provide motivation for Theorem 2.1, we consider the following corollary for the case $a = 0$ and $m = n$.

COROLLARY 2.1. *Suppose that* $m = n$ *and that* g *is a function defined on* R^+ *such that* $\int_0^\infty |g(u)|\, dH(u) < \infty$. *Then,*

$$\lim_{n\to\infty} \int_0^\infty g(u)\, dH_n(u) = \int_0^\infty g(u)\, dH(u), \quad a.s.$$

Corollary 2.1 indicates that the sequence of random measures associated with the sequence $\{H, H_n,\ n \geq 1\}$ possess a type of ergodic property. The statement of Corollary 2.1 is similar in flavor to the statement of the key renewal theorem of Smith (1958), (1.3). Some of the applications of Smith's key renewal theorem are also present in the estimation context of Corollary 2.1. For example, since $H(t) < \infty$ for each $t > 0$ when μ is positive and σ^2 is finite, we may let $g(u) = I(u \leq t)$ to get that $H_n(t) \to H(t)$ a.s. when $m = n$. When $m \neq n$, from Theorem 2.1 it is easy to see that $H_n(t) \to H(t)$ in probability. This is strengthened in Theorem 2.2 below to a.s. convergence by requiring that $m(n)$ grows sufficiently quickly. Define $X^- = \min(0, X)$.

THEOREM 2.2. *Suppose that* F *has positive mean* μ *and finite variance* σ^2. *Suppose that either, for* $r \geq 2$,

$$E|X^-|^r < \infty \quad \text{and} \quad n = O(m^{r-2}) \tag{2.4}$$

or, for some $\theta_1 > 0$ *and all* $|\theta| < \theta_1$,

$$E\exp(-\theta X^-) < \infty \quad \text{and} \quad \log n = o(m). \tag{2.5}$$

Then, for each $t \geq 0$,

$$H_n(t) \to H(t), \quad a.s. \tag{2.6}$$

REMARKS. When X is a nonnegative random variable, from (2.5) we see that the only moment requirement is that $\sigma^2 < \infty$. The growth conditions on m are to ensure that the bias of $H_n(t)$ in estimating $H(t)$ dies out sufficiently quickly.

Corollary 2.1, together with some results on uniformly convergent measures due to Rao (1962), is also used to prove

THEOREM 2.3. *Suppose that $m = n$ and that F has positive mean μ and finite variance σ^2. Then, for each $t \geq 0$,*

$$\sup_{u \in [0, t]} |H_n(u) - H(u)| \to 0, \quad a.s.$$

REMARKS. Note that Theorems 2.1–2.3 do not require that the support of F be on R^+ and also hold for both arithmetic and nonarithmetic distributions. From (1.2), we see that $H(t)$ asymptotically approaches a straight line with positive slope and hence is not bounded. This also holds for F nonarithmetic since limit theorems similar to (1.2) are available, cf. Feller (1968). Thus, one would not expect Theorem 2.3 to be true when the supremum extends over R^+ and, indeed, counter-examples can be easily constructed.

The remainder of the section is devoted to the proof of Theorems 2.1–2.3. The technique is to show that $\int_0^\infty g_a(u)\, d(\sum_{k=1}^m k^a F_n^{(k)}(u))$ is a reverse martingale plus negligible terms. Reverse martingales are a natural tool in this context if we note that $F_n^{(k)}(t)$ is a U-statistic. The idea of applying reverse martingales to U-statistics is due to Berk (1966). An application of Doob's (reverse) martingale convergence theorem will then establish Theorem 2.1. Let $\{X_{1n}, X_{2n}, \dots, X_{nn}\}$ be the order statistics associated with $\{X_1, X_2, \dots, X_n\}$. We use $\chi_n = \sigma(X_{1n}, \dots, X_{nn}, X_{n+1}, X_{n+2}, \dots)$, $n \geq 1$, to define the sequence of nonincreasing sub σ-fields which are implicitly used in all of the following reverse martingales. We preface the proof of Theorem 2.1 with a preparatory lemma.

LEMMA 2.1. *Let $g_a(\cdot)$ be as defined in* (2.1). *Then,*

$$RM_n = \int_0^\infty g_a(u)\, d\left(\sum_{k=1}^n k^a F_n^{(k)}(u)\right) + \int_0^\infty g_a(u)\, d\left(\sum_{k>n} k^a I(S_k \leq u)\right) - \int_0^\infty g_a(u)\, d\left(\sum_{k\geq 1} k^a F^{(k)}(u)\right)$$

is a zero mean reverse martingale.

PROOF. It is easy to see that RM_n is χ_n-measurable and integrable. That $E(RM_n|\chi_{n+1}) = RM_{n+1}$ a.s. can be seen by integration by parts and noting that

$$E\big(I(S_k \leq u)|\chi_n\big) = \begin{cases} F_n^{(k)}(u), & k \leq n, \\ I(S_k \leq u), & k > n, \quad \text{a.s.} \end{cases}$$ □

PROOF OF THEOREM 2.1. Assume (2.1) and without loss of generality, assume $g_a(u) \geq 0$. Define

$$Z_{1n} = \int_0^\infty g_a(u)\, d\left(\sum_{k=m+1}^{n} k^a F_n^{(k)}(u) \right)$$

and

$$Z_{2n} = \int_0^\infty g_a(u)\, d\left(\sum_{k>n} k^a I(S_k \leq u) \right).$$

With Lemma 2.1 and (2.1), as a consequence of Doob's (reverse) martingale convergence theorem we have

$$\begin{aligned} &\lim \int_0^\infty g_a(u)\, d\left(\sum_{k=1}^{m} k^a F_n^{(k)}(u) \right) + Z_{1n} + Z_{2n} \\ &\quad = \int_0^\infty g_a(u)\, d\left(\sum_{k \geq 1} k^a F^{(k)}(u) \right), \quad \text{a.s.} \end{aligned} \tag{2.7}$$

Since Z_{2n} is monotone and bounded, $\lim Z_{2n}$ exists. By Fatou's lemma,

$$\begin{aligned} E \lim Z_{2n} &\leq \lim E \int_0^\infty g_a(u)\, d\left(\sum_{k>n} k^a I(S_k \leq u) \right) \\ &= \lim \int_0^\infty g_a(u)\, d\left(\sum_{k>n} k^a F^{(k)}(u) \right) = 0. \end{aligned} \tag{2.8}$$

Since $\lim Z_{2n}$ is nonnegative and has nonpositive expectation, it is zero a.s. Thus, by (2.7) and (2.8), to prove (2.3) we need only show that $Z_{1n} \to 0$ in probability. By the Markov inequality, for $\varepsilon > 0$,

$$P(Z_{1n} > \varepsilon) \leq \varepsilon^{-1} \int_0^\infty g_a(u)\, d\left(\sum_{k=m+1}^{n} k^a F^{(k)}(u) \right) \to 0$$

by (2.1). We now further assume (2.2) to prove the a.s. version of (2.3). By (2.7) and (2.8), we need only show that $\limsup Z_{1n} = 0$ a.s. By the Markov inequality and a change of summation, for $\varepsilon > 0$,

$$\begin{aligned} \sum_{n \geq 1} P(Z_{1n} > \varepsilon) &\leq \varepsilon^{-1} \sum_{n \geq 1} \int_0^\infty g_a(u)\, d\left(\sum_{k=m+1}^{n} k^a F^{(k)}(u) \right) \\ &= \varepsilon^{-1} \int_0^\infty g_a(u)\, d\left(\sum_{k \geq 1} (m^{-1}(k) - k) k^a F^{(k)}(u) \right) < \infty \end{aligned}$$

by (2.2). This is sufficient for the proof by the Borel–Cantelli lemma. □

PROOF OF THEOREM 2.2. We first assume (2.4). By a straightforward extension of Lemmas 1 and 2 of Heyde (1964) [see also Gut (1974), Theorem 2.1], $E|X^-|^r < \infty$ implies

$$\sum_{k \geq 1} k^{r-2} F^{(k)}(t) < \infty. \tag{2.9}$$

Thus, since $n = O(m^{r-2})$ implies $m^{-1}(n) = O(n^{r-2})$, we have

$$\sum_{k \geq 1} m^{-1}(k) F^{(k)}(t) < \infty. \tag{2.10}$$

This is sufficient for (2.2) with $a = 0$ and $g_a(u) = I(u \leq t)$. By Theorem 2.1 this proves the result under (2.4). Now assume (2.5). Since $\mu > 0$, there exists $\theta_2 > 0$ such that

$$p = \int \exp(-\theta_2 u)\, dF(u) < 1.$$

By the Markov inequality,

$$F^{(k)}(t) = P(-\theta_2 S_k \geq -\theta_2 t) \leq \exp(\theta_2 t) p^k. \tag{2.11}$$

Since $\log n = o(m)$ means there is a sequence $\delta_n \to 0$ such that $\log n \leq \delta_n m$, we have $m^{-1}(n) \leq \exp(\delta_{m^{-1}(n)} n)$. This and (2.11) are sufficient for (2.10) which proves the result under (2.5). □

Proof of Theorem 2.3. Let $A = \{\omega \colon \omega \leq t$ and ω is a discontinuity point of $H\}$. Since the set of discontinuity points of $F^{(k)}(\cdot)$ is countable for each $k \geq 1$, A is countable and we can let $\{a_i\}$ be some enumeration of A. Define $g(u) = \sum_{i \geq 1} I(u = a_i)$. Since $\int_0^\infty g(u)\, dH(u) < H(t) < \infty$, by Theorem 2.1 we have

$$\sum_{i \geq 1} \sum_{k=1}^{n} \left(F_n^{(k)}(a_i) - F_n^{(k)}(a_i -)\right) \to \sum_{i \geq 1} \left(H(a_i) - H(a_i -)\right).$$

Thus, without loss of generality, we may assume that $H(u)$ is continuous for $u \leq t$. The result is now immediate from Corollary 2.1 and Theorems 4.2 and 6.1 of Rao (1962). □

3. Asymptotic normality.

Define

$$\begin{aligned} \xi_{rs}(c) = \operatorname{Cov}\big(&F^{(r-c)}(t - (X_1 + \cdots + X_c)), \\ &F^{(s-c)}(t - (X_1 + \cdots + X_c))\big). \end{aligned} \tag{3.1}$$

In this section we prove the following result.

Theorem 3.1. *Assume $\sigma^2 < \infty$. Suppose that either, for $r \geq 2$,*

$$E|X^-|^r < \infty \quad \text{and} \quad n = O(m^{2r-4}) \tag{3.2}$$

or (2.5) *holds. Then, for each $t > 0$,*

$$\sqrt{n}\,(H_n(t) - H(t)) \to_D N(0, \sigma_1^2),$$

where

$$\sigma_1^2 = \sum_{s=1}^{\infty} \sum_{r=1}^{\infty} rs\,\xi_{rs}(1) < \infty.$$

Remarks. Note that Theorem 3.1 holds for F both arithmetic and nonarithmetic. See Dynkin and Mandelbaum (1983) for a general discussion of weak convergence of symmetric statistics of possibly infinite order. The specific application in Theorem 3.1 does not seem to fit into their broad framework.

From the definition of $F_n^{(k)}(t)$ in (1.4), an easy calculation shows that

$$E\big(F_n^{(k)}(t)|X_1\big) = (k/n)F^{(k-1)}(t - X_1) + (1 - k/n)F^{(k)}(t). \tag{3.3}$$

Define a truncated version of $H(t)$, $H^*(t) = \sum_{k=1}^m F^{(k)}(t)$. We define the projection of $H_n(t)$ on $H^*(t)$ by

$$\hat{H}_n(t) = \sum_{j=1}^{n} E\big(H_n(t)|X_j\big) - (n-1)H^*(t).$$

From (3.3), we have

$$\hat{H}_n(t) - H^*(t) = n^{-1}\sum_{j=1}^{n}\sum_{k=1}^{m} k\big\{F^{(k-1)}(t - X_j) - F^{(k)}(t)\big\}. \tag{3.4}$$

The idea of projecting $H_n(t)$ onto the original independent observations is due to Hoeffding. Since $\hat{H}_n(t)$ is just the sum of n independent random variables, the usual theorems for double arrays of independent random variables are used to obtain a limiting asymptotic distribution for $\hat{H}_n(t)$. We then show that the moments of $H_n(t) - \hat{H}_n(t)$ are small in the appropriate sense to get an identical asymptotic distribution for $H_n(t)$.

From (3.1) and (3.4), we have

$$\operatorname{Var}(\hat{H}_n(t)) = n^{-1}\sum_{r,s=1}^{m} rs\xi_{rs}(1). \tag{3.5}$$

To calculate $\operatorname{Var}(H_n(t))$, we first examine the covariance between $F_n^{(r)}(t)$ and $F_n^{(s)}(t)$. Let $\{a_1, a_2, \dots, a_r\}$ and $\{b_1, b_2, \dots, b_s\}$ be two subsets of $\{1, 2, \dots, n\}$ that have $c \le \min(r, s)$ elements in common. Then,

$$EI(X_{a_1} + X_{a_2} + \cdots + X_{a_r} \le t)I(X_{b_1} + X_{b_2} + \cdots + X_{b_s} \le t) - F^{(r)}(t)F^{(s)}(t) = \xi_{rs}(c).$$

Thus,

$$\operatorname{Cov}\big(F_n^{(r)}(t), F_n^{(s)}(t)\big) = \binom{n}{r}^{-1}\sum_{c=1}^{r}\binom{s}{c}\binom{n-s}{r-c}\xi_{rs}(c),$$

since the number of distinct choices for two subsets of size r and s, respectively, having exactly c elements in common is $\binom{n}{s}\binom{s}{c}\binom{n-s}{r-c}$. Thus,

$$\operatorname{Var}(H_n(t)) = \sum_{r,s=1}^{m}\binom{n}{r}^{-1}\sum_{c=1}^{r}\binom{s}{c}\binom{n-s}{r-c}\xi_{rs}(c). \tag{3.6}$$

To calculate $\operatorname{Cov}(H_n(t), \hat{H}_n(t))$, we first note that

$$\operatorname{Cov}\big(F_n^{(r)}(t), F^{(s-1)}(t - X_1)\big) = r/n\xi_{rs}(1).$$

Hence, from (1.5) and (3.4),

$$\operatorname{Cov}(H_n(t), \hat{H}_n(t)) = \sum_{r,s=1}^{m} rs/n\xi_{rs}(1). \tag{3.7}$$

Thus, from (3.5)–(3.7),

$$nE(H_n(t) - \hat{H}_n(t))^2 = \sum_{r,s=1}^{m} \left\{ n\binom{n}{r}^{-1} \sum_{c=1}^{r} \binom{s}{c}\binom{n-s}{r-c}\xi_{rs}(c) - rs\xi_{rs}(1) \right\}.$$

We now present a series of simple lemmas which, when taken together, provide a proof of Theorem 3.1.

LEMMA 3.1. *Assume* $\sigma^2 < \infty$ *and either* (3.2) *or* (2.5). *Then* $\sigma_1^2 < \infty$ *and*

$$\lim \sqrt{n} \sum_{k>m} F^{(k)}(t) = 0. \tag{3.8}$$

PROOF. We first assume (3.2). Since $E|X^-|^r < \infty$, we have (2.9). Thus, using (3.2),

$$\sqrt{n} \sum_{k>m} F^{(k)}(t) = O\left(\sum_{k>m} k^{r-2} F^{(k)}(t) \right),$$

which proves (3.8) under (3.2). Now assume (2.5). With (2.11),

$$\sqrt{n} \sum_{k>m} F^{(k)}(t) \le \exp(\theta_2 t)\sqrt{n} \sum_{k>m} p^k \to 0,$$

by an easy application of l'Hospital's rule since $\log n = o(m)$. To show $\sigma_1^2 < \infty$, note that for $c \ge 1$,

$$|\xi_{rs}(c)| \le \min(F^{(r)}(t), F^{(s)}(t)) \tag{3.9}$$

and use (2.9) and (2.11). □

LEMMA 3.2. *Under the assumptions and notation of Theorem* 3.1,

$$\sqrt{n}(\hat{H}_n(t) - H(t)) \to_D N(0, \sigma_1^2).$$

PROOF. From (3.8), $\sqrt{n}(H(t) - H^*(t)) \to 0$ and thus, sufficient for the proof of the lemma is

$$\sqrt{n}(\hat{H}_n(t) - H^*(t)) \to_D N(0, \sigma_1^2). \tag{3.10}$$

To prove (3.10), from (3.4), define

$$U_{nj} = n^{-1} \sum_{k=1}^{m} \left(F^{(k-1)}(t - X_j) - F^{(k)}(t) \right).$$

Now $\{U_{nj};\ j = 1, \dots, n,\ n \ge 1\}$ is a double array of random variables that are independent within rows. Now $EU_{nj} = 0$ and, by (3.5),

$$\operatorname{Var}\left(\sum_{j=1}^{n} U_{nj} \right) = n^{-1} \sum_{r,s=1}^{m} rs\xi_{rs}(1).$$

It is easy to check that the usual uniform asymptotic negligibility and Lindeberg conditions hold [cf. Serfling (1980), Section 1.9.3]. □

LEMMA 3.3. *Under the assumptions and notation of Theorem* 3.1,

$$\sum_{r,s=1}^{m}\left\{n\binom{n}{r}^{-1}\sum_{c=1}^{r}\binom{s}{c}\binom{n-s}{r-c}\xi_{rs}(c)-rs\xi_{rs}(1)\right\}\to 0.$$

PROOF. Sufficient for the proof is

$$n\sum_{r,s=1}^{m}\binom{n}{r}^{-1}\sum_{c=2}^{s}\binom{s}{c}\binom{n-s}{r-c}\xi_{rs}(c)\to 0 \tag{3.11}$$

and

$$\sum_{r,s=1}^{m}\left(n\binom{n}{r}^{-1}\binom{s}{1}\binom{n-s}{r-1}-rs\right)\xi_{rs}(1)\to 0. \tag{3.12}$$

To prove (3.12), we have

$$\left|n\binom{n}{r}^{-1}s\binom{n-s}{r-1}-rs\right|=\left|rs\left\{\binom{n-s}{r-1}\Big/\binom{n-1}{r-1}-1\right\}\right|\le 2rs.$$

Thus, we have (3.12) since $\sigma_1^2<\infty$ and by the dominated convergence theorem. Similarly, (3.11) is proved using (3.9) and an application of the dominated convergence theorem. □

4. Interval estimates and an example. A local asymptotic normality result such as Theorem 3.1 is appealing because it gives information about the rate of convergence of $H_n(t)$ to $H(t)$. However, in applications it is also desirable to give interval estimates of $H(t)$. In this section we present an estimator of σ_1^2 and prove its weak consistency. This result and Theorem 3.1 immediately provide a confidence interval for $H(t)$.

Let $\xi_{rs}=E(F^{(r-1)}(t-X)F^{(s-1)}(t-X))$. We wish to estimate

$$\sigma_1^2=\sum_{r,s\ge 1}rs\xi_{rs}(1)=\sum_{r,s\ge 1}rs\xi_{rs}-\left(\sum_{r\ge 1}rF^{(r)}(t)\right)^2.$$

To define an estimator of ξ_{rs}, let $\{i_1,i_2,\dots,i_{r+s-1}\}$ be a subset of $\{1,2,\dots,n\}$, not necessarily ordered. Let Σ_p denote the summation over all permutations of subsets of size $r+s-1$ from $\{1,2,\dots,n\}$. An unbiased estimator of ξ_{rs} is

$$\hat{\xi}_{rs}=(n-r-s+1)!/n!\sum_{p}I\left(X_{i_1}+X_{i_2}+\cdots+X_{i_r}\le t\right)$$
$$I\left(X_{i_1}+X_{i_{r+1}}+\cdots+X_{i_{r+s-1}}\le t\right).$$

In this section, we prove the following result.

THEOREM 4.1. *Let $m_1=m_1(n)$ be a positive integer depending on n such that $m_1\le n$ and $m_1\uparrow\infty$ as $n\uparrow\infty$. Then, with $\sigma_1^2<\infty$ and*

$$\sigma_n^2=\sum_{r,s=1}^{m_1}rs\hat{\xi}_{rs}-\left(\sum_{k=1}^{m_1}kF_n^{(k)}(t)\right)^2, \tag{4.1}$$

we have $\sigma_n^2\to\sigma_1^2$, in probability.

COROLLARY 4.1. *Assume that the conditions of Theorem* 3.1 *hold and let* σ_n^2 *be defined as in* (4.1). *Then*

$$\sqrt{n}/\sigma_n(H_n(t) - H(t)) \to_D N(0,1)$$

and thus, for $\alpha \in (0,1)$,

$$\lim_{n \to \infty} P\left(H_n(t) - z_{\alpha/2}\sigma_n/\sqrt{n} \le H(t) \le H_n(t) + z_{\alpha/2}\sigma_n/\sqrt{n}\right) = 1 - \alpha,$$

where $z_{\alpha/2}$ *is the upper* $\alpha/2$ *quantile of the standard normal distribution.*

PROOF OF THEOREM 4.1. From Theorem 2.1, with $a = 1$ and $g_a(u) = I(u \le t)$, we have

$$\sum_{k=1}^{m_1} kF_n^{(k)}(t) \to \sum_{k \ge 1} kF^{(k)}(t), \quad \text{in probability.}$$

Thus, we need only show that

$$\sum_{r,s=1}^{m_1} rs\hat{\xi}_{rs} \to \sum_{r,s \ge 1} rs\xi_{rs}, \quad \text{in probability.}$$

Let

$$R_n = \sum_{r+s-1 \le n} rs\hat{\xi}_{rs} + \sum_{r+s-1 > n} rsE\{I(S_r \le t)I(S_{r+s-1} - S_{r-1} \le t)|\chi_n\}$$
$$- \sum_{r,s \ge 1} rs\xi_{rs}.$$

As in the proof of Theorem 2.1, the remainder of the proof is to show that R_n is a reverse martingale, apply Doob's (reverse) martingale convergence theorem, and show the excess terms are negligible. □

To illustrate how to calculate the estimator, we used observations of the time to failure of a unit of electronic ground support equipment which were previously used by the author (1986). The data can be found in Juran and Gryna (1970), page 171, and Kolb and Ross (1980), page 170. In Figure 1 an estimate of the failure rate curve is given which suggests an early failure rate of about 20 hours. The estimate of the failure rate was based on Epanečnikov's method. The calculations were done on a VAX 11/750 owned and operated by the Department of Statistics at the University of Wisconsin–Madison. Now, it is not unusual for the manufacturer of equipment to enter into an agreement to replace the equipment for a certain length of time, say, W. This type of agreement is called a warranty and W is the duration of the warranty. In this example, one reasonable warranty duration is the end of the early failure period and thus we give a point and interval estimate of $H(20)$.

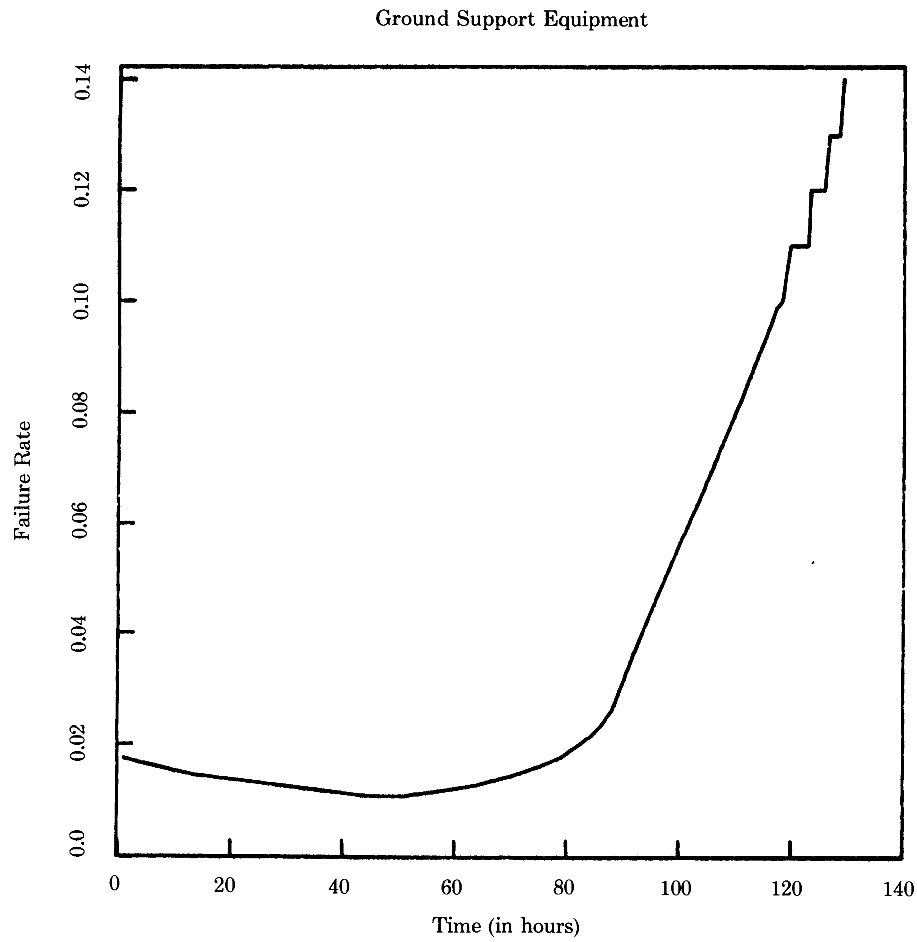

FIG. 1. *Ground support equipment.*

From Table 1, the estimate is $H_n(20) = \sum_{k=1}^{5} F_{105}^{(k)}(20) = 0.46194$. Frees (1986) compared this estimate with other estimators of $H(20)$ and found it reasonable. To calculate the estimated variance of this estimator, from Table 2 we have

$$\sigma_n^2 = \sum_{r,s=1}^{4} rs\hat{\xi}_{rs}(1) = 0.75385.$$

Thus, an approximate 95% interval estimate of $H(20)$ is $H_{105}(20) \pm 2\sigma_n/\sqrt{105}$ which is roughly 0.46 ± 0.17 or $(0.29, 0.63)$.

TABLE 1
Convolution estimates for failure of a unit of electronic ground support equipment

k	1	2	3	4	5	6	7	8
$F_{105}^{(k)}(20)$	0.35238	0.09048	0.01684	0.00208	0.00016	0	0	0

TABLE 2
Covariance estimates for failure of a unit of electronic ground support equipment

	$\hat{\xi}_{rs}(1)$			
r/s	1	2	3	4
1	0.22821	0.05860	0.01091	0.00135
2		0.01733	0.00490	0.00061
3			0.00355	0.00048
4				0.00037

5. Concluding remarks. The renewal function arises in a wide variety of applications of probabilistic models such as in reliability theory, inventory theory, and continuous sampling plans. In this paper we have presented the asymptotic theory of nonparametric estimation of this key function based on a random sample. The simulation study of Frees (1985) showed that $H_n(t)$ performed well for small ($n \le 30$) sample sizes. The techniques of this paper may also be useful in sequential analysis. For example, an important parameter in sequential estimation is the expected value of the first passage time

$$\tau = \inf\{n \ge 1: S_n > 0\}.$$

From, for example, Woodroofe (1982), Corollary 2.4, we have

$$E(\tau) = \exp\left\{\sum_{k \ge 1} k^{-1} P(S_k \le 0)\right\}, \quad \text{when } \mu > 0.$$

Thus, by Theorem 2.1, with $a = -1$,

$$\tau_n = \exp\left\{\sum_{k=1}^{n} k^{-1} F_n^{(k)}(0)\right\}$$

is a consistent estimator of $E(\tau)$. We intend to explore other applications of nonparametric renewal function estimation in sequential analysis in another paper.

Alternatively, one could estimate $H(t)$ by using the empirical distribution function $\tilde{F}_n^{(1)}(t) = F_n^{(1)}(t)$. Estimates of $F^{(k)}(t)$ can be recursively defined by the relationship

$$\tilde{F}_n^{(k)}(t) = \int \tilde{F}_n^{(k-1)}(t-u)\, d\tilde{F}_n^{(1)}(u).$$

Although $\tilde{F}_n^{(k)}(t)$ is a biased estimate of $F^{(k)}(t)$ (for $k \ge 2$), it does have the advantage of being the nonparametric maximum likelihood estimator. Further, $\tilde{F}_n^{(k)}(t)$ is a V-statistic and thus is closely related to the U-statistic $F_n^{(k)}(t)$. By using some of the several well-known results on this relationship [cf. Serfling (1980)], under mild conditions on F it can be shown that $\tilde{H}_n(t)$, defined as in (1.5) with $\tilde{F}_n^{(k)}(t)$ in lieu of $F_n^{(k)}(t)$, is also consistent and has the same asymptotic distribution. However, additional care must be taken with the estimator

$$\tilde{H}(t) = \sum_{k \geq 1} \tilde{F}_n^{(k)}(t).$$

For example, suppose $q = P(X_i \leq 0) > 0$. Then, for $t > 0$ and for each n,

$$q^n = P(X_1 \leq 0, \ldots, X_n \leq 0) \leq P\big(\tilde{F}_n^{(k)}(t) = 1\big), \qquad k = 1, 2, \ldots .$$

Thus, for each n, $\tilde{H}(t) = \infty$ with positive probability.

In many situations the choice of the design parameters m and m_1 is dictated by practical considerations, as in the example in Section 4. Theorems 2.2 and 3.1 give some theoretical guidelines for the choice of m. However, the convolution $F^{(k)}(t)$ dies out quickly as k approaches infinity [cf. (2.9) and (2.11)], and typically m can be small compared to the sample size. A similar argument can be made for m_1. This is important since the amount of computations increases quickly as m (or m_1) increases.

Acknowledgments. The author would like to thank the Associate Editor and referees for comments which greatly improved the scope of the paper.

REFERENCES

BERK, R. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37** 51–58.

BRILLINGER, D. (1975). Statistical inference for stationary point processes. In *Stochastic Processes and Related Topics* (M. L. Puri, ed.) 55–99. Academic, New York.

COX, D. and LEWIS, P. (1966). *The Statistical Analysis of Series of Events.* Methuen, London.

DYNKIN, E. and MANDELBAUM, A. (1983). Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.* **11** 739–745.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3rd ed. Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

FREES, E. (1986). Warranty analysis and renewal function estimation. *Naval Res. Logist. Quart.* **33** 361–372.

GUT, A. (1974). On the moments and limit distributions of some first passage times. *Ann. Probab.* **2** 277–308.

HEYDE, C. (1964). Two probability theorems and their application to some first passage problems. *J. Austral. Math. Soc. Ser. B* **4** 214–222.

JURAN, J. and GRYNA, F. (1970). *Quality Planning and Analysis.* McGraw-Hill, New York.

KARLIN, S. and TAYLOR, H. (1975). *A First Course in Stochastic Processes.* Academic, New York.

KOLB, J. and ROSS, S. (1980). *Product Safety and Liability.* McGraw-Hill, New York.

RAO, R. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33** 659–680.

SERFLING, R. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

SMITH, W. (1958). Renewal theory and its ramifications (with discussion). *J. Roy. Statist. Soc. Ser. B* **20** 243–302.

VARDI, Y. (1982). Nonparametric estimation in renewal processes. *Ann. Statist.* **10** 772–785.

WOODROOFE, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* SIAM, Philadelphia.

YANG, G. (1983). A renewal-process approach to continuous sampling plans. *Technometrics* **25** 59–67.

DEPARTMENT OF STATISTICS AND
SCHOOL OF BUSINESS
UNIVERSITY OF WISCONSIN
1155 OBSERVATORY DRIVE
MADISON, WISCONSIN 53706