BOOTSTRAP CONFIDENCE BANDS FOR THE RENEWAL FUNCTION

by

John Crowell and Pranab K. Sen
Western Michigan University
and
University of North Carolina

Let F be a lifetime distribution and H the associated renewal function. Two nonparametric estimators of H are shown to have desirable asymptotic properties without moment restrictions on F. One of these estimators, the empirical renewal function, is then used to construct an asymptotically correct bootstrap confidence band for H on a finite interval. Weak convergence of the bootstrapped empirical renewal function is established using weak convergence of the bootstrapped empirical process.

AMS 1980 subject classifications. Primary 60K05; secondary 62G15.

Key words and phrases. Renewal function estimation, bootstrap confidence band, weak convergence.

Abbreviated title: Renewal Function Confidence Bands

1. Introduction

Let $X, X_1, X_2, \ldots$ be independent and identically distributed nonnegative random variables with distribution function F. With $S_K = X_1 + \ldots + X_k$ and $F^{(k)}(t) = P(S_k \leq t)$ (the k-th convolution of F), the renewal function H is defined by

$$H(t) = \sum_{k=1}^{\infty} F^{(k)}(t), \quad t \geq 0.$$

The renewal function occurs in a number of different applications, including some where F is not a lifetime distribution (See Frees (1986b)). Here attention is restricted to nonnegative random variables, this being the most important special case for applications. The principal objective of this paper is constructing a confidence band for H on a finite interval when F is a lifetime distribution. For large values of t, the renewal function can be estimated by exploiting an asymptotic expression for H(t) as $t \to \infty$, as described in Frees (1986a).

Several recent papers have considered nonparametric estimation of H using the random sample $X_1, \ldots, X_n$. Let I(E) be the indicator of the event E. An estimator introduced by Frees (1986a,b) is of the form

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

where $m \leq n$ and $m \to \infty$ as $n \to \infty$. Here

$$\hat{F}_n^{(k)}(t) = \binom{n}{k}^{-1} \sum_{i} I(X_{i1} + \ldots + X_{ik} \leq t)$$

is the U-statistic for $F^{(k)}(t)$, the summation being over all subsamples of size k taken without replacement from $\{X_1,\ldots,X_n\}$. Frees establishes strong consistency of $\hat{H}_n(t)$ and asymptotic normality of $n^{1/2}[\hat{H}_n(t)-H(t)]$ for fixed $t \geq 0$ under several sets of conditions on m and the moments of F. Let $F_n(x) = n^{-1} \Sigma_{i=1}^n I(X_i \leq x)$, $-\infty < x < \infty$, be the usual sample distribution function and $F_n^{(k)}$ denote its k-th convolution. Frees (1986b) also discusses the estimator

(1.2) $$H_n(t) = \sum_{k=1}^{\infty} F_n^{(k)}(t),$$

which is just the renewal function that arises from sampling with replacement from F_n.

Schneider, Lin, and O'Cinneide (1990) suggest algorithms for computing both (1.1) and (1.2), as well as comparing the performance of the two estimators on data and in a simulation study. They conclude that H_n may often be the preferred estimator because the order of computation is less, although H_n may fail to perform well when F has rapidly decreasing failure rate.

Let T be a finite positive constant and D[0,T] be the space of functions on [0,T] that are right-continuous and have left-hand limits. O'Cinneide and Schneider (1988) show that the process $r_n = \{n^{1/2} [H_n(t) - H(t)],\ t \in [0,T]\}$ converges weakly to a Gaussian process in the Skorohod topology on D[0,T] when F is a continuous lifetime distribution. We prove that under the same conditions the process r_n can be bootstrapped to construct an asymptotically correct confidence band for H on the interval [0,T].

In Section 2 we note that when F is a lifetime distribution no moment conditions are necessary for strong consistency and asymptotic normality of $\hat{H}_n$; the same properties then hold for H_n. The bootstrap procedure is defined in Section 3; Theorem 3.1 gives conditions for the asymptotic validity of the confidence band. Section 4 contains a proof of weak convergence for the bootstrapped version of the process r_n.

2. Asymptotic Properties of $\hat{H}_n(t)$ and $H_n(t)$

Theorems 2.2 and 3.1 of Frees (1986b) state the strong consistency and asymptotic normality of $\hat{H}_n(t)$ under several sets of conditions on the moments of F and the design parameter m; an underlying assumption is that F has positive mean μ and finite variance σ^2. In the present section it is shown that no moment assumptions are necessary when $F(0-) = 0$ and $F(0) < 1$, as indicated in Theorem 2.1. The asymptotic properties then carry over to H_n without further restrictions on F. (See Theorem 2.2.)

Theorem 2.1: Assume $F(0-) = 0$, $F(0) < 1$, and $\log n = o(m)$. Then for each $t \in \mathbb{R}$

(2.1) $$\hat{H}_n(t) \to H(t) \text{ a.s.}$$

and

(2.2) $$n^{1/2}[\hat{H}_n(t) - H(t)] \xrightarrow{D} N(0,\sigma_t^2),$$

where

$$\sigma_t^2 = \sum_{r,s=1}^{\infty} \text{Cov}\{F^{(r-1)}(t-X),\ F^{(s-1)}(t-X)\}.$$

PROOF: Assuming that F has positive mean μ and finite variance σ^2, a sufficient condition for (2.1) and (2.2) given by Frees (1986b) is that

(2.3) for some $\theta_1 > 0$ and all $|\theta| < \theta_1$, $E \exp(-\theta X^-) < \infty$ and $\log n = o(m)$.

Here $X^- = \min(0,X)$. Let $p = \int e^{-u}\, dF(u)$. Assuming (2.3), the condition that $0 < \mu < \infty$ is used to show that

(2.4) $0 < p < 1$ and there exists a $\theta_2 > 0$ such that $F^{(k)}(t) \leq \exp(\theta_2 t)\, p^k$ for all $k \geq 1$, $t \in \mathbb{R}$.

Frees' proofs of strong consistency and asymptotic normality assuming (2.3) rely solely on the inequality (2.4) and the condition that $\log n = o(m)$; in this case, the assumption of a finite variance can be dropped. If $F(0-) = 0$ and $F(0) < 1$, then it is easily shown that (2.4) holds with $\theta_2 = 1$, even when $EX = \infty$. Taking $\log n = o(m)$, (2.1) and (2.2) follow by the same arguments used by Frees assuming (2.4) holds and $\log n = o(m)$. □

Frees (1986b) noted that by exploiting the relationship between the U-statistic $\hat{F}_n^{(k)}(t)$ and the V-statistic $F_n^{(k)}(t)$ for $k = 1,\ldots,m$, asymptotic properties of $\hat{H}_n(t)$ could be shown to hold for a truncated version of H_n under certain conditions on F. Here we use Theorem 2.1 to establish the asymptotic behaviour of H_n when F is a lifetime distribution.

Theorem 2.2. Assume $F(0-) = 0$ and $F(0) < 1$. Then for each $t \in \mathbb{R}$

(2.5) $$H_n(t) \to H(t) \text{ a.s.}$$

and

(2.6) $$n^{1/2}[H_n(t) - H(t)] \xrightarrow{D} N(0,\sigma_t^2),$$

where σ_t^2 is defined in Theorem 2.1.

PROOF: Take m such that $m^6 = o(n)$ and $\log n = o(m)$. Then (2.1) and (2.2) hold. We first show that

(2.7) $$|\hat{H}_n(t) - \sum_{k=1}^{m} F_n^{(k)}(t)| \to 0 \text{ a.s.}$$

and

$$n^{1/2} \sum_{k=m+1}^{\infty} F_n^{(k)}(t) \to 0 \quad \text{a.s.} \tag{2.8}$$

Since $n^k - n(n-1)(n-2)\ldots(n-k+1) \le \Sigma_{j=1}^{k} k^{2j} n^{k-j}$, it follows that for any positive integer r and a positive constant $C > 0$,

$$E|\hat{F}_n^{(k)}(t) - F_n^{(k)}(t)|^r \le [C\, k^2/n]^r, \quad k \le m.$$

Here C depends on r but not on n. Thus by the Minkowski inequality

$$E|\sum_{k=1}^{m} [\hat{F}_n^{(k)}(t) - F_n^{(k)}(t)]|^r \le [C\, m^3/n]^r \tag{2.9}$$

(See Serfling (1980).) Taking r sufficiently large, by applying the Markov inequality and Borel-Cantelli lemma, (2.7) holds.

Next let $p_n = \int e^{-u}\, dF_n(u)$. Applying (2.4) to the distribution F_n (we may take $\theta_2 = 1$ for any lifetime distribution), we see that

$$n^{1/2} \sum_{k=m+1}^{\infty} F_n^{(k)}(t) \le n^{1/2} e^t p_n^{m+1}/(1-p_n). \tag{2.10}$$

Since $p_n \to p$ a.s. by the strong law of large numbers and $\log n = o(m)$, (2.8) holds. Thus we have (2.5).

Using (2.9) and the Markov inequality, it can be shown that

$$n^{1/2} \sum_{k=1}^{m} [\hat{F}_n^{(k)}(t) - F_n^{(k)}(t)] \overset{P}{\to} 0,$$

which with (2.8) implies (2.6). □

3. A Bootstrap Confidence Band

In the remainder of this paper, the product of two functions without their arguments will denote convolution. For example, $F^2(t) = \int_0^t F(t-u)\, dF(u) =$

$F^{(2)}(t)$, $t \geq 0$. The symbol $\xrightarrow{W}$ will represent weak convergence, where unless otherwise specified this is understood to be in the Skorohod topology on $D[0,T]$. For any function f, let $\|f\| = \sup_{t\in[0,T]} |f(t)|$.

Let W^o be a Brownian bridge on [0,1] and $G(t) = \Sigma_{k=0}^{\infty} (k+1)F^{(k)}(t)$, $t \geq 0$, where $F^{(0)}(t) = I(t \geq 0)$. As in the introduction, define $r_n = \{n^{1/2} [H_n(t) - H(t)],\ t \in [0,T]\}$. O'Cinneide and Schneider (1988) have shown that $r_n \xrightarrow{W} (B \circ F)G$ when F is continuous. Here $\circ$ denotes composition of functions.

To construct a confidence band for H on the interval [0,T], one would like some knowledge of the distribution

$$J(x) = P\{\|(B \circ F)G\| \leq x\}, \quad -\infty < x < \infty.$$

If the desired level of confidence was $1-\alpha$ and one knew $c = \inf\{x : J(x) \geq 1-\alpha\}$, then $\{H_n(t) \pm n^{-1/2} c,\ t \in [0,T]\}$ would be a reasonable large sample confidence band for H having the correct asymptotic coverage probability when J is continuous at c. We now describe how by bootstrapping the process r_n we may estimate J and construct an asymptotically correct confidence band for H on [0,T].

Let $q = q_n$ and $B = B_n$ be sequences of positive integers such that $q \to \infty$ and $B \to \infty$ as $n \to \infty$. Given the random sample $X_1,\ldots,X_n$, let $X_1^*,\ldots,X_q^*$ be drawn with replacement from $\{X_1,\ldots,X_n\}$; that is, $X_1^*,\ldots,X_q^*$ are conditionally independent with distribution F_n. The empirical distribution function of the bootstrap sample is then $F_{nq}(t) = q^{-1} \Sigma_{i=1}^{q} I(X_i^* \leq t)$. Let $F_{nq}^{(k)}$ be the k-th convolution of F_{nq}. Define the bootstrapped version of the process r_n by

$$r_n^* = \{q^{1/2}[H_n^*(t) - H_n(t)],\quad t \in [0,T]\}$$

where $H_n^*(t) = \Sigma_{k=1}^{\infty} F_{nq}^{(k)}(t)$. In Section 4 it is shown that if F is a continuous lifetime distribution, then the conditional distribution of r_n^* given $X_1,\ldots,X_n$ converges weakly to $(B \circ F)G$ along almost all sample sequences.

Given $X_1,\ldots,X_n$, take $r_{n1}^*,\ldots,r_{nB}^*$ to be conditionally independent copies of r_n^*. Define

$$J_n(x) = B^{-1} \sum_{j=1}^{B} I(\|r_{nj}^*\| \le x), \quad -\infty < x < \infty,$$

and take $c_n = \inf\{x : J_n(x) \ge 1-\alpha\}$. A confidence band for H is then given by

$$\{H_n(t) \pm n^{-1/2} c_n, \quad t \in [0,T]\}.$$

The following theorem confirms that the coverage of the confidence band is asymptotically correct when $q = o(n^2)$.

<u>Theorem 3.1</u>: Assume F is absolutely continuous, $F(0) = 0$, and $F(T) > 0$.

If $q = o(n^2)$, then

$$\lim_{n,q,B\to\infty} P\{\|H_n - H\| \le n^{-1/2} c_n | X_1,\ldots,X_n\} = 1-\alpha$$

along almost all sample sequences.

PROOF: By Theorem 4.1, $J_n(x) \to J(x)$ for every continuity point x of J. If J is continuous at c, then

$$\lim_{n,q,B\to\infty} P\{\|H_n - H\| \le n^{-1/2} C_n | X_1,\ldots,X_n\} = 1-\alpha$$

along almost all sample sequences. When F is continuous and $F(T) > 0$, the distribution J cannot have an atom at zero and is necessarily continuous, which completes the proof. (See Csörgő and Mason (1990).) □

The assumption that $q = o(n^2)$ is not restrictive since in practice one typically takes $q = O(n)$.

Theorem 3.1 allows one to test the null hypothesis that $H(t) = H_o(t)$, $t \in [0,T]$, where H_o is some user specified function. Specifically, one would reject the null hypothesis when $\|H_n - H_o\| > n^{-½} c_n$. Asymptotically this test would have size α under the conditions of Theorem 3.1.

More typically one would like to test that F belongs to some class $\mathcal{F}_o$ of distribution functions, and hence H belongs to the corresponding class of renewal functions $\mathcal{H}_o$. Alternatively, one could test directly that $H \in \mathcal{H}_o$ using the following test: if for somme $H' \in \mathcal{H}_o$ one has $\|H_n - H'\| \leq n^{-½} c_n$, then retain the null hypothesis that $H \in \mathcal{H}_o$. For example, suppose $\mathcal{F}_o$ were the class of exponential distributions with positive mean $\mu > 0$ so that $\mathcal{H}_o =$ $\{H : H(t) = t/\mu,\ t \geq 0,\ \mu > 0\}$. Then the null hypothesis $H \in \mathcal{H}_o$ would be retained if for some $\mu > 0$, $\|H_n(t) - t/\mu\| \leq n^{-½} c_n$, a seemingly reasonable test in this case.

4. Weak Convergence

<u>Theorem 4.1</u>: Assume F is an absolutely continuous distribution such that $F(0) = 0$. If $q = o(n^2)$, then the conditional distribution of r_n^* given $X_1, \ldots, X_n$ converges weakly to $(B \circ F)G$ as $n, q \to \infty$ along almost all sample sequences.

PROOF: Throughout this section let $U(x) = F^{(0)}(x) = I(x \geq 0)$. Expressions of the form $U/(U-F)$ will represent power series, with convolution taking the place of multiplication. For example,

$$\frac{U}{U-F}(t) = 1 + H(t), \quad t \geq 0.$$

O'Cinneide and Schneider (1988) exploited an identity for $r_n(t)$ to show weak convergence of r_n. We utilize the same identity for r_n^*, namely

$$(4.1) \qquad q^{1/2}[H_n^* - H_n] = q^{1/2}[F_{nq} - F_n] \frac{U}{(U-F_n)^2} + \frac{1}{q^{1/2}} [q^{1/2}(F_{nq} - F_n)]^2 \frac{U}{U-F_{nq}} \frac{U}{(U-F_n)^2} .$$

Adapting the approach used by O'Cinneide and Schneider (1988), we first show that for any $\epsilon > 0$

$$(4.2) \qquad P\{\| \frac{1}{q^{1/2}} [q^{1/2}(F_{nq}-F_n)]^2 \frac{U}{U-F_{nq}} \frac{U}{(U-F_n)^2} \| > \epsilon \ |X_1,\ldots,X_n\} \to 0 \text{ as } n,q \to \infty.$$

By Theorem 2.2

$$(4.3) \qquad \left\| \frac{U}{U-F_n} \right\| \leq 1 + H_n(T) \to 1 + H(T) \quad \text{a.s.}$$

Similarly, $\|U/(U-F_{nq})\| \leq 1 + H_n^*(T)$; we show that there exists a positive constant M such that

$$(4.4) \qquad P\{1 + H_n^*(T) \leq M \mid X_1,\ldots,X_n\} \to 1 \text{ as } n,q \to \infty.$$

By a well known probability inequality (cf. Serfling (1980)), for any $\epsilon > 0$

$$(4.5) \qquad P\{\|F_{nq}-F\| > \epsilon \mid X_1,\ldots,X_n\} \to 0 \text{ as } n,q \to \infty.$$

Let $p_{n*} = \int e^{-u}\, dF_{nq}(u)$ and as in Section 2 (using (2.4))

$$(4.6) \qquad 1 + H_n^*(T) \leq 1 + e^{-T} \sum_{k=1}^{\infty} p_{n*}^k .$$

Choose $\delta > 0$ such that $F(\delta) \leq e^{-\delta/2}$. Then

(4.7) $$p_{n*} \le F_{nq}(\delta)(1-e^{-\delta}) + e^{-\delta}$$

and by (4.5), (4.6) and (4.7), we have that (4.4) holds.

Bickel and Freedman (1981) showed that for almost all sample sequences the conditional distribution of the bootstrapped empirical process $q^{1/2}[F_{nq}-F_n]$ converges weakly to $B \circ F$. Thus (4.2) follows from (4.3) and (4.4).

Letting $G_n(t) = \Sigma_{k=0}^{\infty} (k+1)F_n^{(k)}t$, where $F_n^{(0)}(t) = F^{(0)}(t) = U(t)$, we can write

$$q^{1/2}[F_{nq} - F_n] \frac{U}{(U-F_n)^2}$$

$$= q^{1/2}[F_{nq} - F_n]G + q^{1/2}[F_{nq} - F_n] [G_n - G],$$

where G was defined in Section 3. As noted by O'Cinneide and Schneider (1988), when F is continuous G is a continuous, monotone increasing function on [0,T]. Thus convolution with G is a continuous mapping of D[0,T] into D[0,T]. Using Theorem 4.1 of Bickel and Freedman (1981) and Theorem 5.1 of Billingsley (1968), it follows that $q^{1/2}[F_{nq}-F_n]G \overset{W}{\to} (B \circ F)G$ for almost all sample sequences.

It remains to be shown that for any $\epsilon > 0$

(4.8) $$P\{\|q^{1/2}[F_{nq}-F_n][G_n-G]\| > \epsilon \mid X_1,\ldots,X_n\} \to 0 \quad \text{as} \quad n,q \to \infty.$$

As pointed out in Section 2,

$$p_n = \int e^{-u}\, dF_n(u) \to p = \int e^{-u}\, dF(u) \quad \text{a.s.}$$

Thus using (2.4) and dominated convergence,

(4.9) $$G_n(t) \to G(t) \quad \text{a.s.} \quad \text{for all } t \in [0,T].$$

We next define continuous versions of the empirical functions F_n and F_{nq}. Let $t_1 < t_2 < \ldots < t_n$ be the points of discontinuity of F. With $t_o = 0$, define

$$F_n^o(t) = \begin{cases} F_n(t_i) + n^{-1}(t-t_i)/(t_{i+1}-t_i), & t_i \le t < t_{i+1}, \quad i = 0,\ldots,n-1; \\ 1, & t \ge t_n. \end{cases}$$

Similarly define F_{nq}^o by linearly interpolating between the successive discontinuities of F_{nq}.

Since F is continuous and $q = o(n^2)$,

$$\|q^{½}[F_n^o - F_n]\| \le \frac{q^{½}}{n} \to 0, \tag{4.10}$$

and for any $\epsilon > 0$ it can be shown that

$$P\{\|q^{½}[F_{nq}-F_{nq}^o]\| > \epsilon \mid X_1,\ldots,X_n\} \to 0. \tag{4.11}$$

Thus

$$q^{½}[F_{nq}^o - F_n^o] \overset{W}{\to} B \circ F \quad \text{a.s.}, \tag{4.12}$$

where this weak convergence takes place in the uniform topology on C[0,T], the space of continuous functions on [0,T].

By (4.9) and (4.10),

$$\|q^{½}[F_n^o-F_n]\ [G_n-G]\| \le \frac{q^{½}}{n}\ [G_n(T) + G(T)] \to 0 \quad \text{as} \quad n,\ q \to \infty. \tag{4.13}$$

Similarly employing (4.9) and (4.11) yields that for any $\epsilon > 0$

$$P\{\|q^{½}[F_{nq}-F_{nq}^o]\ [G_n-G]\| > \epsilon \mid X_1,\ldots,X_n\} \to 0. \tag{4.14}$$

Thus we would like to establish that

$$P\{\|q^{½}[F_{nq}^o-F_n^o][G_n-G]\| > \epsilon \mid X_1,\ldots,X_n\} \to 0 \tag{4.15}$$

as $n,\ q \to \infty$.

Let $x \in C[0,T]$ and $\{x_n\}$ any sequence in $C[0,T]$ such that $x_n \to x$ in the uniform metric. Then given $X_1,\dots,X_n$, $x_n[G_n-G] \to 0$ in the Skorohod metric on $D[0,T]$, where 0 is the function which is identically zero on $[0,T]$. Thus by (4.12) and Theorem 5.5 of Billingsley (1968), given $X_1,\dots,X_n$

$$q^{\frac{1}{2}}[F^o_{nq} - F^o_n][G_n-G] \overset{W}{\to} 0$$

in the Skorohod topology on $D[0,T]$. Thus (4.15) holds, and by (4.13), (4.14), and (4.15) the desired conclusion (4.8) holds as well. □

REFERENCES

Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* 9, 1196-1217.

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

Csörgő, S. and Mason, D.M. (1990). Bootstrapping empirical functions. *Ann. Statist.* 17, 1447-1471.

Frees, E.W. (1986a) Warranty analysis and renewal function estimation. *Naval Res. Logist. Quart.* 33, 361-372.

Frees, E.W. (1986b). Nonparametric renewal function estimation. *Ann. Statist.* 14, 1366-1378.

O'Cinneide, C. and Schneider, H. (1988). Weak convergence of the sample renewal function. Working paper. Department of Mathematics, University of Arkansas, Fayetteville.

Schneider, H., Lin, B. and O'Cinneide, C. (1990). Comparison of nonparametric estimators for the renewal function. *Appl. Statist.* 39, 55-61.

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

DEPARTMENT OF MATHEMATICS & STATISTICS
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MI 49008

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NC 27599-3260