⊠연구논문

Nonparametric Estimation of Renewal Function

Hai Sung Jeong
Dept. of Applied Statistics, Seowon University
Jee Hoon Kim
Statistical Research Institute, Seoul National University
Myoung Hwan Na
Dept. of Computer Science and Statistics, Seoul National University

Abstract

We consider a nonparametric estimation of the renewal function. In this paper, we suggest modified methods for Frees's estimator to enhance the efficiency. The methods are based on a piecewise linearization and on the fact that the bounded monotonic functions converging pointwise to the bounded monotonic continuous function converge uniformly. In a simulation study, we show that the modifed methods have the better efficiency than that introduced by Frees.

1. Introduction

In reliability theory, renewal processes describe the model of an item in continuous operation which is replaced at each failure, in a negligible amount of time, by a new, statistically identical item. The most important and basic function in renewal processes is the renewal function. To obtain the renewal function, we should derive the k-fold convolution, $F^{(k)}(t)$ generated from the life distribution function $F(t)$. Let $X_1, X_2, \cdots, X_n$ be non-negative random samples from the life distribution function $F(t)$. Then $F^{(k)}(t)$ is defined by $F^{(k)}(t) = P(X_1 + X_2 + \cdots + X_k \leq t)$ for $k \geq 1$. And the renewal function, $M(t)$ is defined as followings.

$$M(t) = \sum_{k=1}^{\infty} F^{(k)}(t).$$

$M(t)$ means the expected number of renewals in an interval $(0, t]$ for renewal processes with the underlying life distribution $F(t)$. The renewal function plays an important role in many probabilistic models and sequential analysis such that warranty analysis and the determination of the preventive maintenance policy.

There are several studies for calculating the renewal function. Smith and Leadbetter(1963) have given a series expansion method for calculating it when $F(t)$ is the Weibull distribution. Jaquette(1972) have given a method based on the asymptotic expansions of the dominating residues of the Laplace transform of the renewal function, and Cleroux and McConalogue(1976) a method involving numerical algorithm for calculating $F^{(k)}(t)$ recursively from the known density function $f(t)$. However all these methods are parametric approaches. Since, in many cases, the underlying life distribution $F(t)$ is not known, it is desirable to have a nonparametric estimate of the renewal function in order to apply to the field. Frees(1986a, b) has suggested the nonparametric estimator for the renewal function based on random samples without replacement. He defines the nonparametric estimator

$$\hat{M}_F(t) = \sum_{k=1}^{m} \hat{F}_F^{(k)}(t),$$

where

$$\hat{F}_F^{(k)}(t) = \binom{n}{k}^{-1} \sum_{i} I(X_{i1} + \cdots + X_{ik} \leq t), \tag{1.1}$$

and the summation in (1.1) extends over all subsamples without replacement of size k from $\{X_1, X_2, \cdots, X_n\}$. Here, $I(A)$ is the indicator function of the event A. The design parameter m is a positive integer depending on n such that $m \leq n$ and $m \uparrow \infty$ as $n \uparrow \infty$. Often he simply uses $m = n$. Another nonparametric estimator suggested by Grübel and Pitts(1993) is based on the sum of convolutions with replacement. And this is as follows.

$$\hat{M}_{GP}(t) = \sum_{k=1}^{m} \hat{F}_{GP}^{(k)}(t),$$

where

$$\hat{F}_{GP}^{(k)}(t) = n^{-k} \sum_{i} I(X_{i1} + \cdots + X_{ik} \le t), \tag{1.2}$$

and the summation in (1.2) extends over all subsamples with replacement of size k from $\{X_1, X_2, \cdots, X_n\}$. The drawback of this estimator is the considerable amount of computation needed to evaluate it.

In this paper, we suggest modified methods for Frees's estimator to enhance the efficiency. The methods are based on a piecewise linearization and on the fact that the bounded monotonic functions converging pointwise to the bounded monotonic continuous function converge uniformly. Also we compare our methods with Frees's via Monte Carlo simulation and show that the modified methods have the better efficiency than that of Frees.

2. Estimation of $F^{(k)}(t)$ and $M(t)$

The scheme used in this section to estimate $F^{(k)}(t)$, $k \ge 1$, is the modified representation of $\hat{F}_F^{(k)}(t)$ in (1.1), estimator of $F^{(k)}(t)$ as a piecewise linear function. This piecewise linear function is specified by n_k jump points $\xi_{k,1}, \ldots, \xi_{k,n_k}$ from $\hat{F}_F^{(k)}(t)$ in (1.1). The proposed estimator of $F^{(k)}(t)$ is as follows.

$$\hat{F}_{PL}^{(k)}(t) = \frac{2}{\xi_{k,j+2} - \xi_{k,j}} \frac{1}{\binom{n}{k}} \left| (j+1)\left(t - \frac{\xi_{k,j} + \xi_{k,j+1}}{2}\right) + j\left(\frac{\xi_{k,j+1} + \xi_{k,j+2}}{2} - t\right) \right|,$$

$$\text{if } \frac{\xi_{k,j} + \xi_{k,j+1}}{2} \le t < \frac{\xi_{k,j+1} + \xi_{k,j+2}}{2} \quad (j = 0, 1, \ldots, n_k - 1),$$

where we define $\xi_{k,0} = 0$ and $\xi_{k,n_k} = \xi_{k,n_{k+1}}$. Then proposed estimator of $M(t)$ is expressed as

$$\hat{M}_{PL}(t) = \sum_{k=1}^{m} \hat{F}^{(k)}(t).$$

Next we will prove almost sure uniform consistency of our estimator for the case $m = n$.

Theorem 1. Suppose that $n=m$ and $M(t)$ is continuous. Then, for each $t \geq 0$,

$$\sup_{u \in [0,t]} | \widehat{M}_{PL}(u) - M(u) | \rightarrow 0, \quad \text{a.s.}$$

Proof. We see from Theorem 2.3 and corollary 2.1 of Frees (1986 b) that

$$\sup_{u \in [0,t]} | \widehat{M}_{F}(u) - M(u) | \rightarrow 0, \quad \text{a.s.},$$

and that

$$\widehat{M}_{F}(u-) \rightarrow M(u-), \quad \text{a.s.}.$$

Also, since the convergence of bounded monotonic functions on $[0,t]$ converging pointwise to the bounded monotonic continuous function is uniform, it is easily obtained that

$$\sup_{u \in [0,t]} | \widehat{M}_{F}(u-) - M(u-) | \rightarrow 0, \quad \text{a.s.}.$$

Let $A=\{\omega : 0 \leq \omega \leq t$ and ω is a discontinuity point of $\widehat{M}_F\}$. Then we have

$$\begin{aligned}
&\sup_{u \in [0,t]} | \widehat{M}_{F}(u) - \widehat{M}_{PL}(u) | \\
&\quad \leq \sup_{u \in A} | \widehat{M}_{F}(u) - \widehat{M}_{F}(u-) | \\
&\quad \leq \sup_{u \in A} | \widehat{M}_{F}(u) - M(u) | + \sup_{u \in A} | \widehat{M}_{F}(u-) - M(u-) |,
\end{aligned}$$

and so proved.

During our simulation study, we can find another estimator which seems to be more efficient. This another estimator of $F^{(k)}(t)$ is given by

$$\widehat{F}_{PL^*}^{(k)}(t) = \frac{1}{\xi_{k,j+1} - \xi_{k,j}} \frac{1}{\binom{n}{k}+1} \left| (j+1)(t-\xi_{k,j}) + j(\xi_{k,j+1} - t) \right|,$$

$$\text{if} \quad \xi_{k,j} \leq t < \xi_{k,j+1}, \quad (j=1, \cdots n_k - 1),$$

where we define $\xi_{k,0} = 0$. Then this estimator of $M(t)$ can be given by

$$\widehat{M}_{PL^*}(t) = \sum_{k=1}^{m} \widehat{F}_{PL^*}^{(k)}(t).$$

3. Simulation Study

This simulation study is to compare of the Free's estimator $\widehat{M}_F(t)$ and our estimators, $\widehat{M}_{PL}(t)$ and $\widehat{M}_{PL^*}(t)$ by computing the bias, variance and mean squared error(MSE). From this we can see the relative performance among those estimators. $m = 5$ is sufficient for the number of terms(m) required to summation associated with estimators $\widehat{M}_F(t)$, $\widehat{M}_{PL}(t)$ and $\widehat{M}_{PL^*}(t)$, since the value of k-fold convolution is close to 0 as k increases. It is seldom possible or convenient to express the renewal function in an analytical form. An analytical solution can be found if $F(t)$ is a special case of gamma distribution (exponential, Erlang). As is well known, the exponential distribution with $f(t) = \lambda e^{-\lambda t}$, $\lambda > 0$, generates the analytically defined renewal function $M(t) = \lambda t$. And let $F(t)$ be the Erlang distribution with 2 stages with its density, $f(t) = te^{-t}$. Renewal function of this Erlang distribution is generated as the closed form with $M(t) = \frac{t}{2} - \frac{1}{4} + \frac{1}{4}e^{-2t}$. For these two distributions, random variables are generated by using International Mathematical and Statistical Libraries(IMSL). At the 25, 50 and 75 percentiles, the bias, variance and MSE for the three sample size(n) 10, 20 and 30 are calculated with 1000 replications.

We summarize our findings from <Table 1> and <Table 2> as follows:

1) Our estimates always have MSE less than that of Free (1986b) in all cases. In particular, $\widehat{M}_{PL^*}(t)$ is the best among three.

2) Our estimates are more efficient as n decreases.

3) Our estimates have the better performances at lower percentiles.

< Table 1 > Bias, Variance and MSE of estimators with 1000 replications †

n	t ‡	$\hat{M}_F(t)$ Bias	$\hat{M}_F(t)$ Var	$\hat{M}_F(t)$ MSE	$\hat{M}_{PL}(t)$ Bias	$\hat{M}_{PL}(t)$ Var	$\hat{M}_{PL}(t)$ MSE	$\hat{M}_{PL^*}(t)$ Bias	$\hat{M}_{PL^*}(t)$ Var	$\hat{M}_{PL^*}(t)$ MSE	$\frac{MSE(\hat{M}_F)}{MSE(\hat{M}_{PL^*})}$
10	.2877	.0057	.0346	.0346	.0025	.0304	.0304	.0399	.0280	.0296	1.17
	.6931	.0024	.1032	.1032	.0003	.0964	.0964	.0156	.0903	.0905	1.14
	1.3863	.0033	.2690	.2690	.0064	.2607	.2607	.0167	.2482	.2485	1.08
20	.2877	.0029	.0180	.0181	.0014	.0165	.0165	.0169	.0160	.0163	1.11
	.6931	.0037	.0512	.0512	.0052	.0496	.0497	.0018	.0479	.0479	1.07
	1.3863	.0126	.1300	.1302	.0160	.1296	.1299	.0250	.1268.	1274	1.02
30	.2877	.0021	.0112	.0112	.0010	.0106	.0106	.0108	.0103	.0104	1.08
	.6931	.0031	.0315	.0315	.0036	.0309	.0309	.0024	.0302	.0302	1.04
	1.3863	-.0110	.0846	.0848	.0121	.0846	.0847	.0191	.0832	.0836	1.01

† Random variables are generated from exponential distribution with $f(t) = e^{-t}$, $t>0$

‡ These values are obtained from 25, 50 and 75 percentiles of given distribution.

< Table 2 > Bias, Variance and MSE of estimators with 1000 replications †

n	t ‡	$\hat{M}_F(t)$ Bias	$\hat{M}_F(t)$ Var	$\hat{M}_F(t)$ MSE	$\hat{M}_{PL}(t)$ Bias	$\hat{M}_{PL}(t)$ Var	$\hat{M}_{PL}(t)$ MSE	$\hat{M}_{PL^*}(t)$ Bias	$\hat{M}_{PL^*}(t)$ Var	$\hat{M}_{PL^*}(t)$ MSE	$\frac{MSE(\hat{M}_F)}{MSE(\hat{M}_{PL^*})}$
10	.9613	.0036	.0231	.0231	.0034	.0205	.0205	.0343	.0185	.0197	1.17
	1.6783	.0008	.0564	.0564	.0021	.0506	.0506	.0128	.0467	.0468	1.21
	2.6926	.0032	.1141	.1141	.0025	.1104	.1104	.0125	.1024	.1025	1.11
20	.9613	.0007	.0131	.0131	.0020	.0121	.0121	.0162	.0116	.0118	1.11
	1.6783	.0043	.0286	.0286	.0022	.0274	.0274	.0056	.0262	.0263	1.09
	2.6926	.0081	.0588	.0588	.0056	.0578	.0578	.0033	.0561	.0561	1.05
30	.9613	.0017	.0087	.0087	.0021	.0083	.0083	.0112	.0080	.0082	1.06
	1.6783	.0090	.0197	.0198	.0081	.0193	.0194	.0094	.0187	.0188	1.05
	2.6926	.0144	.0399	.0402	.0133	.0393	.0395	.0065	.0387	.0387	1.04

† Random variables are generated from Erlang distribution with $f(t) = te^{-t}$, $t>0$.

‡ These values are obtained from 25, 50 and 75 percentiles of given distribution.

References

[1] Cleroux, R. and McConalogue, D.J.(1976) "A Numerical Algorithm for Recursively-Defined Convolution Integrals Involving Distribution Functions," *Management Science*, 22, pp. 1138-1146.

[2] Frees, E.W.(1986a) "Warranty Analysis and Renewal Function Estimatoin," *Naval Research Logistics Quarterly*, 33, pp. 361-372.

[3] Frees, E.W.(1986b) "Nonparametric Renewal Function Estimatoin," *The Annals of Statistics*, 14, pp. 1366-1378.

[4] Grübel, R. and Pitts S.M.(1993) "Nonparametric Estimation in Renewal Theory I: The Empirical Renewal Function," *The Annals of Statistics*, 21, pp. 1431-1451.

[5] Jaquette, D.L.(1972) "Approximations to the Renewal Function $m(t)$," *Operations Research*, 20, pp. 722-727.

[6] Smith, W.L. and Leadbetter, M.R.(1963) "On the Renewal Function for the Weibull Distribution," *Technometrics*, 5, pp. 393-396.

[7] Woodroofe, M.(1982) *Nonlinear Renewal Theory in Sequantial Analysis*. SIAM, Philadelphia.