International Journal of Reliability and Application
Vol.1, No.1, pp.81–87, 2000

Parametric Estimation of a Renewal Function

Hai Sung Jeong
Department of Applied Statistics
Seowon University, Chonju 361–742, Korea

Myung Hwan Na
Department of Naval Architecture and Ocean Engineering
Chosun University, Kwangju 501–759, Korea

Abstract. One of the most important quantities in reliability theory is the expected number of renewals of a system during a given interval. This quantity, the renewal function, is used to determine the optimal preventive maintenance policy and to estimate the cost of a warranty. In this paper we study a parametric approach for a renewal funtion. The simulation study is presented to compare the relative performance of the introduced estimators of a renewal function. And we show that the proposed parametric estimator performs well.

Key Words : *renewal function, convolution, parameter, cubic spline.*

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, the expected number of renewals in a given interval. To obtain the renewal function, we should derive the k -fold convolution, $F^{(k)}(t)$ generated from the life distribution function, $F(t)$. This $F^{(k)}(t)$ is defined recursively by the repeated convolution of $F(t)$ itself according to the following scheme

$$\begin{aligned} F^{(1)}(t) &= F(t) \\ F^{(k)}(t) &= (F^{(k-1)} * F)(t) = \int_0^t F^{(k-1)}(x) f(t-x) dx, \quad k \geq 2, \end{aligned}$$

where f is the density funtion, the derivative of F and the simbol * denotes convolution. And the renewal function, $M(t)$ is defined as follows.

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

It is impossible or, if possible, inconvenient 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 such as exponential and Erlang . There are three approaches for calculating the renewal function. The first is to find the renewal funtion in the case that life distribution function $F(t)$ is known. Smith and Leadbetter(1963) have given a series-expansion method for calculating it when $F(t)$ is Weibull distribution. Cléroux and McConalogue(1976) have given a method involving numerical algorithm for calculating $F^{(k)}(t)$ recursively from the known density function. Since, in many cases, the underlying life distribution $F(t)$ is unknown, this method is unrealistic. The second is a nonparametric approach free from assumptions about a life distribution. Frees(1986a, b) has suggested the nonparametric estimator for a renewal fuction based on random samples without replacement. Let $X_1, X_2, \cdots, X_n$ be non-negative random samples with distribution $F(t)$. He defines a nonparametric estimator

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

where

$$\hat{F}^{(k)}(t) = \binom{n}{k}^{-1} \sum_{i} I(X_{i1} + \cdots + X_{ik} \leq t), \tag{2}$$

and the summation in (2) extends 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 $n \uparrow \infty$. Grübel and Pitts(1993) have suggested another nonparametric estimator based on the sum of convolutions with replacement. In this case the summation in the estimator extends n^k times. To enhance the efficiency, modified methods for Frees's estimator are suggested by Jeong, Kim and Na(1997). These methods are based on a piecewise linearization and on the fact that the bounded monotonic functions which converge pointwise to the bounded monotonic continuous function converge uniformly. The drawback of these nonparametric approaches is that the efficiency is relatively low due to the loss of information about a life distribution and that the computational difficulty arises as the sample size increases. As mentioned before, the underlying life distribution $F(t)$ is unknown in most cases. There is another possible way to overcome these problems; Assuming distribution family including life distribution functions which will be identified if unknown characteristics are decided, you have only to think out how to guess these unknown characteristics, which are named unknown parameters. This way is the third approach called a parametric approaches. In this paper we consider the parametric estimation of renewal function.

2. ESTIMATION OF $F^{(k)}(t)$ AND $M(t)$

Let $F(t|\theta)$ be the life distribution function with unknown parameter θ and $F(t|\hat{\theta})$ be the estimator of $F(t|\theta)$. Based on the random samples $X_1, X_2, \cdots, X_n$, the estimator of θ, $\hat{\theta}$ can be obtained by the traditional point estimation method such as maximun likelihood estimation, method of moments etc. We consider the algorithm which gets k-th convolution, $F^{(k)}(t|\hat{\theta})$ from $F(t|\hat{\theta})$ and its derivative, $f(t|\hat{\theta})$ by cubic spline interpolation, and calculates the renewal function with this k-th convolution.

Let $t_0, t_1, \cdots, t_K$ be a knot sequence in $[\ 0, \infty\ )$ where $0 = t_0 < t_1 < \cdots < t_K < \infty$, $K > 1$. Values of $F^{(k)}(t|\hat{\theta})$, $k \geq 2$ at the knots points $t = t_0, t_1, \cdots, t_K$ can be obtained as followings.

$$\begin{aligned} F^{(k)}(0|\hat{\theta}) &= 0, \\ F^{(k)}(t_i|\hat{\theta}) &= \sum_{j=1}^{i} \int_{t_{j-1}}^{t_j} F^{(k-1)}(x|\hat{\theta}) f(t_i - x|\hat{\theta}) dx, \quad i = 1, \cdots, K. \end{aligned}$$

In the above equation, the integral from the point t_{j-1} to t_j can be calculated using the numerical integration formula such as Newton-cotes and Simpson's composite. In this case $F^{(k-1)}(t|\hat{\theta})$ is defined only at $t = t_0, t_1, \cdots, t_K$ and intermidiate values required by the numerical integration formula are given by cubic spline interpolation.

A cubic spline approximation to $F^{(k)}(t|\hat{\theta})$ over the entire range can be obtained by defining K different cubic polynomials y_j for each interval $[\ t_{j-1},\ t_j]$, $j = 1, 2, \cdots, K$. If F_j denotes $F^{(k)}(t_j|\hat{\theta})$ for simplicity, then for $t_{j-1} \leq x \leq t_j$, $F^{(k)}(x|\hat{\theta})$ is represented by y_j satisfied by the following four conditions.

$$\begin{aligned} &a)\ \ y_j(t_{j-1}) = F_{j-1}, \\ &b)\ \ y_j(t_j) = F_j, \\ &c)\ \ \frac{dy_{j-1}(t_{j-1})}{dx} = \frac{dy_j(t_{j-1})}{dx}, \\ &d)\ \ \frac{d^2 y_{j-1}(t_{j-1})}{dx^2} = \frac{d^2 y_j(t_{j-1})}{dx^2}. \end{aligned} \qquad (3)$$

Since, in general, y_j , $j = 1, 2, \cdots, K$ are different cubics, there are $4K$ coefficients to determine K different cubics. From the conditions in (3) at the knot points $t = t_1, t_2, \cdots, t_{K-1}$, we can obtain $4(K-1)$ equations. Furthemore we know

$y_1(t_0) = F_0$ and $y_K(t_K) = F_K$. In order to find all coefficients, two additional conditions are needed besides the $4K-2$ conditions obtained above. These two can be obtained with the left and right boundary conditions. Usually three kinds of conditions are recommanded (chapter 4 in de Boor(1978)). The first is that F'_0 and F'_K, the values of the first derivative of $F^{(k)}(t|\hat{\theta})$ at the boundary point are specified. The second is that F''_0 and F''_K, the values of the second derivative of $F^{(k)}(t|\hat{\theta})$ at the boundary point are specified. Natural splines are obtainable by setting these values zero. The last is "not-a-knot" condition. If one knows nothing about boundary point derivatives, then one should try the "not-a-knot" condition. Here, one chooses y_1 and y_K so that $y_1 = y_2$ and $y_K = y_{K-1}$ (i.e., the first and the last interior knots are not active).

Using $F^{(k)}(t|\hat{\theta})$ by cubic spline approximation, we construct the renewal function as follows.

$$\hat{M}_P(t) = \sum_{k=1}^{m} F^{(k)}(t|\hat{\theta}). \tag{4}$$

3. SIMULATION STUDY

The purpose of this simulation study is to compare the relative performance of the estimators for a renewal function by computing the bias and the mean squared error(MSE). Simulation study is performed for three estimators mentioned before. The first is $\hat{M}_{CM}(t)$, an estimator by Cléroux and McConalogue's algorithm stated in the Introduction. Since Cléroux and McConalogue's algorithm assumes the known life distribution, random samples are not needed in Cléroux and McConalogue's algoritmn. Hence in the case of Cléroux and McConalogue, MSE has no meaning. The second is $\hat{M}_F(t)$, Frees' nonparameric estimator in (1). The last is $\hat{M}_P(t)$, our parametric estimator in (4). As is known well, the exponential distribution with parameter λ, $f(t|\lambda) = \lambda e^{-\lambda t}$, $\lambda > 0$, generates the analytically defined renewal function with $M(t) = \lambda t$. And let $F(t|\lambda)$ be the Erlang distribution with 2 stages with its density, $\lambda^2 t e^{-\lambda t}$. The renewal function of this Erlang distribution is generated as the closed form with $M(t) = \lambda t/2 - 1/4 + 1/4e^{-\lambda t}$. For the above two distributions with parameter $\lambda = 1$, random samples are generated by using International Mathematical and Statistical Libraries(IMSL) routine RNGAM. Based on this random samples $X_1, X_2, \cdots, X_n$, the maximun likelihood estimator and the momemt method estimator for λ can be obtained as $\hat{\lambda} = n/\sum_{i=1}^{n} X_i$ for exponential distribution and $\hat{\lambda} = 2n/\sum_{i=1}^{n} X_i$ for Erlang distribution with 2 stages. At the 25, 50, and 75 percentiles, the bias and MSE for the sample size(n) 10, 20, 30, 50 and 100 are calculated with 1000 replications. Due to the computational difficulty, the simulations for $\hat{M}_F(t)$ is carried for sample size $n \leq 30$. The bias and MSE for the sample size 50 and 100 are calculated for the case of $\hat{M}_P(t)$ only. IMSL

Table 1: Bias and MSE of estimators with 1000 replications †

n	t‡	$M(t)$	$\hat{M}_{CM}(t)$	$\hat{M}_F(t)$		$\hat{M}_P(t)$	
			Bias	Bias	MSE	Bias	MSE
10	.2877	.2877	.000010	.0057	.0346	.0075	.0004
	.6931	.6931	-.000061	.0024	.1032	.0421	.0025
	1.3863	1.3863	-.004161	.0033	.2690	.1106	.0137
20	.2877	.2877	.000010	.0029	.0181	.0043	.0003
	.6931	.6931	-.000061	-.0037	.0512	.0191	.0011
	1.3863	1.3863	-.004161	-.0126	.1302	.0470	.0036
30	.2877	.2877	.000010	.0021	.0112	.0026	.0003
	.6931	.6931	-.000061	-.0031	.0315	.0125	.0009
	1.3863	1.3863	-.004161	-.0110	.0848	.0306	.0024

n	t	$\hat{M}_P(t)$		n	t	$\hat{M}_P(t)$	
		Bias	MSE			Bias	MSE
50	.2877	.0013	.0003	100	.2877	.0008	.0003
	.6931	.0073	.0008		.6931	.0036	.0007
	1.3863	.0175	.0017		1.3863	.0062	.0014

† 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.

routine CSDEC and CSVAL are used to compute the cubic spline interpolant. And "not-a-knot" is used for the type of boundary condition. $m = 5$ is sufficient for the number of terms required to summation associated with estimators $\hat{M}_{CM}(t)$, $\hat{M}_F(t)$ and $\hat{M}_P(t)$ since the value of k-fold convolution is close to 0 as k increases. Thus we use $m = 5$.

We summarize our findings from TABLE 1 and TABLE 2 as follows;
1) $\hat{M}_{CM}(t)$ always has less bias than $\hat{M}_F(t)$ and $\hat{M}_P(t)$ in all cases expect the case of $n = 100$ in TABLE 2, which has been expected.
2) $\hat{M}_P(t)$ outperforms $\hat{M}_F(t)$ with respect to the MSE in all cases.
3) Although $\hat{M}_P(t)$ has higher bias than $\hat{M}_F(t)$ in the most cases of $n \leq 30$, the differences in MSEs have much more to compansate for that.
4) The more n increases, the less the bias and MSE of $\hat{M}_P(t)$ is. This is a good property as an estimator.
5) For the case of $n = 100$, the differences in the biases between $\hat{M}_P(t)$ and $\hat{M}_{CM}(t)$ are little. This means that our proposed estimator $\hat{M}_P(t)$ performs similarly compare to $\hat{M}_{CM}(t)$ obtained under the assumption of the known life distribution function.

Table 2: Bias and MSE of estimators with 1000 replications †

n	t‡	$M(t)$	$\hat{M}_{CM}(t)$	$\hat{M}_F(t)$		$\hat{M}_P(t)$	
			Bias	Bias	MSE	Bias	MSE
10	.1913	.2672	-.000542	-.0036	.0231	.0051	.0003
	1.6783	.5979	-.000717	-.0008	.0564	.0178	.0009
	2.6926	1.0975	-.000778	.0032	.1141	.0470	.0033
20	.1913	.2672	-.000542	.0007	.0131	.0010	.0003
	1.6783	.5979	-.000717	.0043	.0286	.0075	.0007
	2.6926	1.0975	-.000778	.0081	.0588	.0252	.0018
30	.1913	.2672	-.000542	.0017	.0087	-.0007	.0003
	1.6783	.5979	-.000717	.0090	.0198	.0036	.0006
	2.6926	1.0975	-.000778	.0144	.0402	.0181	.0014

n	t	$\hat{M}_P(t)$		n	t	$\hat{M}_P(t)$	
		Bias	MSE			Bias	MSE
50	.9613	-.0001	.0003	100	.9613	-.0005	.0003
	1.6783	.0022	.0006		1.6783	.0005	.0006
	2.6926	.0092	.0012		2.6926	.0042	.0011

† 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.

In summary, our proposed parametric estimator $\hat{M}_P(t)$ for a renewal function performs well.

REFERENCES

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

de Boor, C. (1978), *A Practical Guide to Spline*, Springer-Verlag, New York.

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

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

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.

Jeong, H. S., Kim, J. H. and Na, M. H. (1997), *Nonparametric Estimation of Renewal Function* , Journal of the Korean Society for Quality Management, Vol 23 No 4, pp. 99-105.

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