REVSTAT – Statistical Journal
Volume 11, Number 3, November 2013, 301–315

TWO NONPARAMETRIC ESTIMATORS OF THE MEAN RESIDUAL LIFE

Authors: Abdel-Razzaq Mugdadi
– Department of Mathematics and Statistics,
Jordan University of Science and Technology,
Irbid, Jordan
aamugdadi@just.edu.jo

Amanuel Teweldemedhin
– Department of Mathematics,
University of Wisconsin – Richland,
Richland Center, WI 53581, U.S.A.
amanuel.teweldemedhi@uwc.edu

Received: March 2012 Revised: January 2013 Accepted: February 2013

Abstract:

- The mean residual life function $L(t)$ can be written based on the vitality function $V(t)$. In this article we propose two methods to estimate $V(t)$. The two methods are based on both the kernel density estimation and the empirical function. In addition, we evaluate the mean square error of the two estimators and we study the consistency for both of them.

Key-Words:

- *mean residual life; kernel; empirical; estimation; mean square error; bandwidth; consistent.*

AMS Subject Classification:

- 62G05, 62N02.

1. INTRODUCTION

The mean residual life MRL is the expected remaining life, $T - t$, given that the item has survived to time t. The unconditional mean of the distribution, $E(T)$, is a special case given by $L(0)$. To determine a formula for this expectation, the conditional probability density function is needed

$$f_{T|T\geq t}(\tau) = \frac{f(\tau)}{P[T \geq t]} = \frac{f(\tau)}{R(t)}\,, \qquad \tau \geq t\,. \tag{1.1}$$

This conditional probability density function is actually a family of probability density functions (one for each value of t), each of which has an associated mean

$$E[T|T \geq t] = \int_t^\infty \tau f_{T|T\geq t}(\tau)\, d\tau = \int_t^\infty \tau \frac{f(\tau)}{R(t)}\, d\tau\,.$$

Thus, in life testing situations, the expected additional lifetime given that a component has survived until time t is called the MRL. Since the MRL function is the expected *remaining* life, t must be subtracted, yielding

$$L(t) = E[T - t|T \geq t] = \frac{1}{R(t)} \int_t^\infty \tau f(\tau)\, d\tau - t\,. \tag{1.2}$$

Thus $L(t)$ can be written as

$$L(t) = V(t) - t\,,$$

where

$$V(t) = \frac{\int_t^\infty s f(s)\, ds}{R(t)} = \frac{M(t)}{R(t)}\,. \tag{1.3}$$

We study the vitality function estimator when $R(t) > 0$, since the vitality function estimator generates the mean residual life function estimator directly by the above equation.

Ratio functions for which nonparametric estimators have been considered include the MRL function and hazard rate among others. One estimation method involves individual estimates of the numerator and denominator. An alternative estimator is to estimate the entire function not the separate pieces. For a discussion of ratio functions estimates see Patil *et al.* [13]. In many reliability studies, the MRL function (corresponding to a lifetime distribution with density $f(t)$, and survival function $R(t)$), is of prime importance. A problem of considerable interest, therefore, is the estimation of mean residual life function. The kernel density estimation is the most popular technique to estimate the probability density function, which is basically can be define as follow: Let $X_1, ..., X_n$ be a random sample from unknown continuous probability density function $f(x)$. The kernel density estimate with appropriate kernel function $k(t)$ and smoothing parameter h is

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right). \tag{1.4}$$

Kernel type estimators of ratio functions, such as the density under random censoring and the hazard rate have been studied by several authors (e.g. Watson and Leadbetter [18], [19], Marron and Padgett [10], Lo *et al.* [9], Sarda and Vieu [16], and Hollander and Proschan [8]).

The basic estimator for $L(t)$ is $\hat{L}(t) = \hat{V}_0(t) - t$, where $\hat{V}_0(t) = \frac{M_e(t)}{R_e(t)}$. $R_e(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{(X_i>t)}$ and $M_e(t) = \frac{1}{n}\sum_{i=1}^{n} X_i 1_{(X_i>t)}$, but Abdous and Berred[1] discussed that $\hat{V}_0(t)$ does not take into account the smoothness of $V(t)$. Guillamon *et al.* [6] studied the estimator $\hat{V}_3(t) = \frac{M_n(t)}{R_n(t)}$ for $V(t)$, where $M_n(t) = \int_t^{\infty} s f_n(s)\, ds$, $f_n(t)$ is the kernel density estimation defined in (1.4) and $V_n(t)$ is the kernel reliability estimator (see Section 3). Other estimators or cases proposed by Mitra and Basu [11], Ruiz and Guillamon [14], Chaubey and Sen [3], and Abdous and Berred [1]. In this paper we propose and study two new estimators for the MRL both based on the kernel estimator and the empirical function. Also, we propose new techniques to select the bandwidth for the estimators. From the simulations, we can conclude that the new estimator is competitive with the basic one but we can't say it is a better one.

2. THE FIRST ESTIMATOR $\widehat{V}_1(t)$

In the first propose estimator we use kernel estimate for the numerator function and empirical estimator of the survival function in the denominator. Thus,

$$\widehat{V}_1(t) = \frac{\int_t^{\infty} s f_n(s)\, ds}{R_e(t)} = \frac{M_n(t)}{R_e(t)}\ , \tag{2.1}$$

where

$$f_n(t) = \frac{1}{n} \sum_{i=1}^{n} k_h(t - X_i)\ , \tag{2.2}$$

and

$$M_n(t) = \int_t^{\infty} s f_n(s)\, ds \tag{2.3}$$

is the kernel estimate of the numerator, and

$$R_e(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{(X_i>t)} , \tag{2.4}$$

is the frequency count of a set divided by n. Thus, we can estimate the MRL $L(t)$ by:

$$\widehat{L}_1(t) = \widehat{V}_1(t) - t . \tag{2.5}$$

2.1. Properties of $\widehat{V}_1(t)$

In this section, we evaluate the Bias, the variance and the Mean Square Error (MSE) of $\widehat{V}_1(t)$. In addition, we derive the optimal bandwidth that minimizes the Asymptotic Mean Square Error (AMSE) and we study the consistency of $\widehat{V}_1(t)$.

Proposition 2.1. *For any t with $R(t) > 0$,*

$$\widehat{V}_1(t)-V(t) = \frac{1}{nR(t)}\left(\sum_{i=1}^{n}\left(\int_t^\infty s\,k_h(s-X_i)\,ds - V(t)\,1_{(X_i>t)}\right)\right)\Big(1+o(1)\Big) . \tag{2.6}$$

Proof:

$$\begin{aligned}
\widehat{V}_1(t) - V(t) &= \frac{M_n(t)}{R_e(t)} - V(t)\\
&= \left(\frac{M_n(t) - V(t)\,R_e(t)}{R(t)}\right)\left(1 + \frac{R(t)-R_e(t)}{R_e(t)}\right)\\
&= \frac{1}{nR(t)}\left(\sum_{i=1}^{n}\left(\int_t^\infty s k_h(s-X_i)\,ds \,-\, V(t)\,1_{(X_i>t)}\right)\right)\Big(1+o(1)\Big) . \;\square
\end{aligned}$$

Lemma 2.1. *Let $\widehat{V}_1(t)$ be as (2.1), then*

$$\begin{aligned}
E\Big(\widehat{V}_1(t)\Big) &= V(t) + \frac{h^2}{2R(t)}\,\mu_2(k)\int_t^\infty s f''(s)\,ds + o(h^2) ,\\
Var\Big(\widehat{V}_1(t)\Big) &= \frac{1}{nR^2(t)}\Big(\Gamma_2(t) + h\,\gamma_2(t)\,\alpha(k)\Big) + o\big(h^2\big) + o\Big(\frac{h}{n}\Big) ,\\
MSE\Big(\widehat{V}_1(t)\Big) &= \frac{1}{n}\frac{\Gamma_2(t)}{R^2(t)} + h\frac{\gamma_2(t)\,\alpha(k)}{nR^2(t)} + h^4\frac{\mu_2^2(k)}{4R^2(t)}\left(\int_t^\infty s f''(s)\,ds\right)^2 + o\Big(h^4+\frac{h}{n}+h\Big),
\end{aligned}$$

where $\Gamma_i(t) = \int_t^\infty \gamma_i(s)\,ds$, $\gamma_i(t) = t^i f(t)$, $\mu_2(k) = \int_{-\infty}^\infty s^2 k(s)\,ds$, *and* $\alpha(k) = \int_{-\infty}^\infty 2\,s\,W(s)\,k(s)\,ds < \infty$.

Proof: Using (2.6)

$$
\begin{aligned}
E\Big(\widehat{V}_1(t) - V(t)\Big) &= \frac{1}{R(t)}\left(E\left(\int_t^\infty s k_h(s-X)\,ds\right) - V(t)\,E\Big(1_{(X>t)}\Big)\right) \\
&= \frac{1}{R(t)}\left(\Gamma_1(t) + \frac{h^2}{2}\mu_2(k)\int_t^\infty s f''(s)\,ds - V(t)\int_t^\infty f(s)\,ds\right) + o(h^2) \\
&= \frac{h^2}{2R(t)}\,\mu_2(k)\int_t^\infty s f''(s)\,ds + o(h^2)\ .
\end{aligned}
$$

Thus,

$$
E\Big(\widehat{V}_1(t)\Big) = V(t) + \frac{h^2}{2R(t)}\,\mu_2(k)\int_t^\infty s f''(s)\,ds + o(h^2)\ .
$$

Also, from (2.6), and after some reduction,

$$
\begin{aligned}
Var\Big(\widehat{V}_1(t) - V(t)\Big) &= \\
&= \frac{1}{nR^2(t)}\left(E\left(\int_t^\infty s k_h(s-X)\,ds - V(t)\,1_{(X>t)}\right)^2\right. \\
&\quad \left. - \left[E\left(\int_t^\infty s k_h(s-X)\,ds - V(t)\,1_{(X>t)}\right)\right]^2\right) \\
&= \frac{1}{nR^2(t)}\left(E\left(\int_t^\infty s k_h(s-X)\,ds\right)^2\right. \\
&\quad \left. - 2V(t)\,E\left(1_{X>t}\int_t^\infty s k_h(s-X)\,ds\right) + V^2(t)\,E\Big(1_{X>t}\Big)\right) + o\left(\frac{h}{n}\right) \\
&= \frac{1}{nR^2(t)}\Big(\Gamma_2(t) + h\gamma_2(t)\,\alpha(k) - 2V(t)\,A + V(t)\,\Gamma_1(t)\Big) + o\left(\frac{h}{n}\right).
\end{aligned}
$$

But

$$
\begin{aligned}
A &= E\left(1_{(X>t)}\int_t^\infty s k_h(s-X)\,ds\right) \\
&= \int_t^\infty \int_{\frac{t-y}{h}}^\infty (y+hx)\,k(x)\,dx\,f(y)\,dy \\
&= \frac{1}{2}\,\Gamma_1(t) + h\,\Gamma_1'(t)\int_0^\infty s k(s)\,ds + o\big(h^2\big)\,.
\end{aligned}
$$

So that

$$
\begin{aligned}
&Var\Big(\widehat{V}_1(t) - V(t)\Big) = \\
&= \frac{1}{nR^2(t)} \Bigg(\Gamma_2(t) + h\gamma_2(t)\,\alpha(k) - 2V(t)\Bigg[\frac{1}{2}\,\Gamma_1(t) + h\Gamma_1'(t)\int_0^\infty sk(s)\,ds\Bigg] \\
&\quad + V(t)\,\Gamma_1(t)\Bigg) + o\bigg(h^2 + \frac{h}{n}\bigg) \\
&= \frac{1}{nR^2(t)} \Bigg(\Gamma_2(t) + h\gamma_2(t)\,\alpha(k) - \Gamma_1(t)V(t) - 2\,hV(t)\,\Gamma_1'(t)\int_0^\infty sk(s)\,ds\;\Gamma_1(t)\Bigg) \\
&\quad + o\bigg(h^2 + \frac{h}{n}\bigg) \\
&= \frac{1}{nR^2(t)} \Bigg(\Gamma_2(t) + h\gamma_2(t)\,\alpha(k) - 2\,hV(t)\,\Gamma_1'(t)\int_0^\infty sk(s)\,ds\Bigg) + o\bigg(h^2 + \frac{h}{n}\bigg)\,.
\end{aligned}
$$

Thus,

$$
Var\Big(\widehat{V}_1(t)\Big) = \frac{1}{nR^2(t)}\Big(\Gamma_2(t) + h\gamma_2(t)\,\alpha(k)\Big) + o(h^2) + o\bigg(\frac{h}{n}\bigg)\,.
$$

Therefore,

$$
\begin{aligned}
MSE\Big(\widehat{V}_1(t)\Big) &= Bias^2\Big(\widehat{V}_1(t)\Big) + Var\Big(\widehat{V}_1(t)\Big) \\
&= \frac{1}{n}\frac{\Gamma_2(t)}{R^2(t)} + h\frac{\gamma_2(t)\,\alpha(k)}{nR^2(t)} + h^4\frac{\mu_2^2(k)}{4R^2(t)}\Bigg(\int_t^\infty sf''(s)\,ds\Bigg)^2 + o\bigg(h^2 + h^4 + \frac{h}{n}\bigg).
\end{aligned}
$$

□

The following corollaries can be obtained directly from the above Lemma.

Corollary 2.1. *The asymptotic mean integrated square error (AMISE) of* $\hat{V}_1(t)$ *is*

$$
\begin{aligned}
AMISE\Big(\widehat{V}_1(t)\Big) &= \frac{1}{n}\int \frac{\Gamma_2(t)}{R^2(t)}\,dt + \frac{h}{n}\,\alpha(k)\int\frac{\gamma_2(t)}{R^2(t)}\,dt \\
&\quad + \frac{h^4}{4}\,\mu_2^2(k)\int\frac{1}{R^2(t)}\Bigg(\int_t^\infty sf''(s)\,ds\Bigg)^2 dt\,.
\end{aligned}
\tag{2.7}
$$

Corollary 2.2. *The optimal bandwidth that minimizes the* $AMISE(\hat{V}_1(t))$ *is*

$$
\widehat{h}_{opt1} = n^{-\frac{1}{3}}\left(\frac{-\alpha(k)}{\mu_2^2(k)}\,\frac{\int \frac{\gamma_2(t)}{R^2(t)}\,dt}{\int\frac{1}{R^2(t)}\bigg(\int_t^\infty sf''(s)\,ds\bigg)^2 dt}\right)^{\frac{1}{3}}.
$$

Corollary 2.3. *The estimator $\widehat{V}_1(t)$ is a asymptotically consistent estimator of the vitality function $V(t)$. That is*

$$\widehat{V}_1(t) \xrightarrow{p} V(t) \ . \tag{2.8}$$

3. THE SECOND ESTIMATOR $\widehat{V}_2(t)$

In this section we use empirical estimate for the numerator and kernel estimate of the survival function in the denominator. Thus,

$$\widehat{V}_2(t) = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i 1_{(X_i>t)}}{R_n(t)} = \frac{M_e(t)}{R_n(t)} \ , \tag{3.1}$$

where $R_n(t)$ is the kernel reliability estimator

$$R_n(t) = \frac{1}{n}\sum_{i=1}^{n} W\left(\frac{t-X_i}{h}\right) \tag{3.2}$$

(see Nadaraya [12], Azzalini [2] and Swanepoel [17]), $k(x)$ is a class-2 symmetric kernel, $k_h(x) = \frac{1}{h}\,k(\frac{x}{h})$, $W(t) = \int_t^\infty k(s)\,ds$, and h is a bandwidth (or smoothing parameter) verifying $h \to 0$ and $nh \to \infty$ when $n \to \infty$; and

$$M_e(t) = \frac{1}{n}\sum_{i=1}^{n} X_i 1_{(X_i>t)} \ , \tag{3.3}$$

where

$$1_T = \begin{cases} 1 & \text{if T is true}\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

is the empirical estimate of the numerator in the definition of V.

In this case, the MRL estimator is

$$\widehat{L}_2(t) = \widehat{V}_2(t) - t \ . \tag{3.4}$$

3.1. Properties of $\widehat{V}_2(t)$

In this section we evaluate the MSE and the AMISE of $\widehat{V}_2(t)$. Also, we derive the optimal bandwidth and study the consistency of $\widehat{V}_2(t)$.

Proposition 3.1. *For any t with $R(t) > 0$,*

$$\widehat{V}_2(t) - V(t) = \frac{1}{nR(t)} \left[\sum_{i=1}^{n} \left(X_i 1_{(X_i > t)} - V(t) W\left(\frac{t - X_i}{h} \right) \right) \right] \Big(1 + o(1)\Big) \tag{3.5}$$

where $W(t) = \int_t^\infty k(s)\, ds$.

Proof:

$$\begin{aligned}
\widehat{V}_2(t) - V(t) &= \frac{M_e(t)}{R_n(t)} - V(t) \\
&= \left(\frac{M_e(t) - V(t) R_n(t)}{R(t)} \right) \left(1 + \frac{R(t) - R_n(t)}{R_n(t)} \right) \\
&= \frac{1}{nR(t)} \left[\sum_{i=1}^{n} \left(X_i 1_{(X_i > t)} - V(t) W\left(\frac{t - X_i}{h} \right) \right) \right] \Big(1 + o(1)\Big) . \quad \square
\end{aligned}$$

Lemma 3.1. *Let $\widehat{V}_2(t)$, $\Gamma_i(t)$, $\gamma_i(t)$, $\mu_2(k)$, and $\alpha(k)$ be as defined earlier, then*

$$\begin{aligned}
E\Big(\widehat{V}_2(t)\Big) &= V(t) + \frac{1}{2} h^2 \frac{V(t)}{R(t)} f'(t)\, \mu_2(k) + o(h^2) \ , \\
Var\Big(\widehat{V}_2(t)\Big) &= \frac{1}{nR^2(t)} \Gamma_2(t) + \frac{h}{nR^2(t)} V^2(t)\, f(t)\, \alpha(k) + o(h) + o\left(\frac{h}{n}\right) , \\
MSE\Big(\widehat{V}_2(t)\Big) &= \frac{1}{4} h^4 \frac{V^2(t)}{R^2(t)} \Big(f'(t)\Big)^2 \mu_2^2(k) + \frac{1}{nR^2(t)} \Gamma_2(t) + \frac{h}{nR^2(t)} V^2(t)\, f(t)\, \alpha(k) \\
&\quad + o\left(h + \frac{h}{n} \right) .
\end{aligned}$$

Proof: Using the result in Proposition (3.1)

$$\begin{aligned}
E\Big(\widehat{V}_2(t) - V(t)\Big) &= \frac{1}{R(t)} \left(E\big(y I_{(y>t)}\big) - V(t)\, E\left[W\left(\frac{t - X_i}{h} \right) \right] \right) \\
&= \frac{h^2 V(t)}{2R(t)} f'(t)\, \mu_2(k) + o(h^2) \ .
\end{aligned}$$

Thus,

$$E\Big(\widehat{V}_2(t)\Big) = V(t) + \frac{1}{2} h^2 \frac{V(t)}{R(t)} f'(t)\, \mu_2(k) + o(h^2) \ . \tag{3.6}$$

Now, we want to evaluate the variance.

$$\begin{aligned}
Var\Big(\widehat{V}_2(t) - V(t)\Big) &= \frac{1}{nR^2(t)}\left(Var\big(yI_{(y>t)}\big) - V(t)\,W\left(\frac{t-y}{h}\right)\right) \\
&= \frac{1}{nR^2(t)}\left(E\big(yI_{(y>t)}\big)^2 - 2V(t)\,E\left(yI_{(y>t)} - W\left(\frac{t-y}{h}\right)\right)\right. \\
&\quad \left. + V^2(t)\,E\left(W^2\left(\frac{t-y}{h}\right)\right)\right) + o\left(\frac{h}{n}\right) \\
&= \frac{1}{nR^2(t)}\left(\Gamma_2(t) - \frac{1}{2}h^2 V(t)\,\mu_2(k)\int_t^\infty s f''(s)\,ds + hV^2(t)\,f(t)\,\alpha(k)\right) \\
&\quad + o\left(h + \frac{h}{n}\right).
\end{aligned}$$

Thus,

$$Var\Big(\widehat{V}_2(t)\Big) = \frac{1}{nR^2(t)}\,\Gamma_2(t) + \frac{h}{nR^2(t)}\,V^2(t)\,f(t)\,\alpha(k) + o\left(h + \frac{h}{n}\right).$$

Therefore,

$$\begin{aligned}
MSE\Big(\widehat{V}_2(t)\Big) &= \frac{1}{4}\,h^4\frac{V^2(t)}{R^2(t)}\,\big(f'(t)\big)^2\mu_2^2(k) + \frac{1}{nR^2(t)}\,\Gamma_2(t) + \frac{h}{nR^2(t)}V^2(t)\,f(t)\,\alpha(k) \\
&\quad + o\left(h + \frac{h}{n}\right).
\end{aligned}$$

□

Corollary 3.1.

$$\begin{aligned}
AMISE\Big(\widehat{V}_2(t)\Big) &= \frac{1}{4}\,h^4\mu_2^2(k)\int \frac{V^2(t)\,\big(f'(t)\big)^2}{R^2(t)}\,dt + \frac{1}{n}\int\frac{\Gamma_2(t)}{R^2(t)}\,dt \\
&\quad + \frac{h\,\alpha(k)}{n}\int\frac{V^2(t)\,f(t)}{R^2(t)}\,dt\,.
\end{aligned} \tag{3.7}$$

Corollary 3.2. *The optimal bandwidth that minimizes the* $AMISE(\hat{V}_2(t))$ *is*

$$\widehat{h}_{opt2} = n^{-\frac{1}{3}}\left(\frac{-\alpha(k)}{\mu_2^2(k)}\,\frac{\int\frac{V^2(t)\,f(t)}{R^2(t)}\,dt}{\int\frac{V^2(t)\,(f'(t))^2}{R^2(t)}\,dt}\right)^{\frac{1}{3}}.$$

Corollary 3.3. *The estimator* $\widehat{V}_2(t)$ *is a asymptotically consistent estimator of the vitality function* $V(t)$*. That is*

$$\widehat{V}_2(t) \xrightarrow[n\to\infty]{p} V(t)\,. \tag{3.8}$$

From Corollaries 2.2 and 3.2, we can conclude that the optimal bandwidths decrease at rate $O(n^{-\frac{1}{3}})$, which is the same of rate of convergence for bandwidth for the kernel distribution function estimator.

4. BANDWIDTH SELECTIONS

4.1. Likelihood Cross-Validation

The original cross-validation criterion, proposed by Habbema *et al.* [7] and Duin [4] to select the bandwidth h by minimizing the score function

$$LCV(h) = -\frac{1}{n}\sum_{i=1}^{n}\log \hat{f}_{-i}(X_i)$$

over possible values of h. $\hat{f}_{-i}(X_i)$ is the "leave-one-out" kernel density estimator defined using the data with X_i removed. That is

$$\hat{f}_{-i}(X_i) = \frac{1}{(n-1)\,h}\sum_{j\neq i}^{n} k\left(\frac{X_i - X_j}{h}\right).$$

The method of likelihood cross-validation is a natural development of the idea of using likelihood to judge the adequacy of fit of a statistical model. It is of general applicability beyond choosing h in kernel density estimation, having been used for both parameter estimation and model selection (e.g. Geisser [5]).

Analogous to this we propose this kind of technique to our estimators. That is we will minimize the following function:

$$LCV(h) = -\frac{1}{n}\sum_{i=1}^{n}\log \hat{V}_{j,-i}(X_i) \tag{4.1}$$

where $\hat{V}_{j,-i}(X_i)$, $j = 1, 2$, is the "leave-one-out" vitality function estimators defined using the data with X_i removed. That is

$$\hat{V}_{1,-i}(X_i) = \frac{\frac{1}{n-1}\left[\sum\limits_{j\neq i}^{n} X_i \cdot W\left(\frac{X_i-X_j}{h}\right) + h\sum\limits_{j\neq i}^{n} N_k\left(\frac{X_i-X_j}{h}\right)\right]}{\frac{1}{n-1}\sum\limits_{j\neq i}^{n} 1_{X_j>X_i}} \tag{4.2}$$

and

$$\hat{V}_{2,-i}(X_i) = \frac{\frac{1}{n-1}\sum\limits_{j\neq i}^{n} X_i \cdot 1_{X_j>X_i}}{\frac{1}{n-1}\sum\limits_{j\neq i}^{n} W\left(\frac{X_i-X_j}{h}\right)}. \tag{4.3}$$

4.2. Simulation

We have conducted a numerical study to asses the performance of the estimators that introduced earlier. We simulated repeated samples of size $n = 20, 40, 60, 80,$ and 100 from exponential distribution with different means. Thus, the true value being estimated is $L(t)$, when the $f(t)$ is the exponential distribution. The results obtained are based on 1000 repetitions at the sample sizes. We used Epanechnikov kernel, and the likelihood cross-validation for bandwidth selections. The Bias, Variance, and MSE are calculated by repeating the samples 1000 times for each case. Epanechnikov kernel is used for the estimators $\hat{L}_1, \hat{L}_2$, and $\hat{L}_3$. Note that $\hat{L}_0(t) = \frac{M_e(t)}{R_e(t)} - t$ and $\hat{L}_3(t) = \frac{M_n(t)}{R_n(t)} - t$.

Table 1: Simulation from exponential distribution of different means and sample size 20.

n	Mean	Estimators	Bias	Variance	MSE
20	0.5	$\hat{L}_0$	-0.2900	0.3222	0.4063
20	0.5	$\hat{L}_1$	0.3644	0.3034	0.4362
20	0.5	$\hat{L}_2$	0.2643	0.2792	0.3490
20	0.5	$\hat{L}_3$	0.6002	0.2358	0.5960
20	1	$\hat{L}_0$	0.1253	0.0842	0.0999
20	1	$\hat{L}_1$	-0.7525	0.0379	0.6083
20	1	$\hat{L}_2$	-0.2110	0.0575	0.1020
20	1	$\hat{L}_3$	0.0473	0.0542	0.0564
20	5	$\hat{L}_0$	0.1939	0.0478	0.0854
20	5	$\hat{L}_1$	-0.0125	0.0024	0.0026
20	5	$\hat{L}_2$	0.0634	0.0150	0.0191
20	5	$\hat{L}_3$	-0.0714	0.0040	0.0091

Table 2: Simulation from exponential distribution of different means and sample size 40.

n	Mean	Estimators	Bias	Variance	MSE
40	0.5	$\hat{L}_0$	-0.0546	0.3402	0.3432
40	0.5	$\hat{L}_1$	0.1494	0.0834	0.1057
40	0.5	$\hat{L}_2$	-0.9094	0.2316	1.0586
40	0.5	$\hat{L}_3$	0.0260	0.1623	0.1630
40	1	$\hat{L}_0$	0.2041	0.1025	0.1442
40	1	$\hat{L}_1$	0.2277	0.0857	0.1375
40	1	$\hat{L}_2$	-0.1786	0.0370	0.0688
40	1	$\hat{L}_3$	-0.0282	0.0311	0.0319
40	5	$\hat{L}_0$	-0.0406	0.0014	0.0031
40	5	$\hat{L}_1$	0.0632	0.0016	0.0056
40	5	$\hat{L}_2$	-0.0115	0.0036	0.0037
40	5	$\hat{L}_3$	0.0073	0.0019	0.0020

The original purpose of this study was to provide kernel based estimators of mean residual life function. We have found kernel estimation to be a useful tool for nonparametric estimations of reliability functions such as MRL. However, the use of this tool in practice can be hampered by the lack of a suitable bandwidth selection procedure. The likelihood cross-validation proposed in this paper is a suitable technique to select the bandwidth but we can not say it is the optimal one. Also, we can not conclude in the MRL estimators that the smoothing technique is better than the non smoothing technique.

Table 3: Simulation from exponential distribution of different means and sample size 60.

n	Mean	Estimators	Bias	Variance	MSE
60	0.5	$\hat{L}_0$	0.2186	0.1971	0.2449
60	0.5	$\hat{L}_1$	0.1788	0.1156	0.1476
60	0.5	$\hat{L}_2$	0.5544	0.1545	0.4619
60	0.5	$\hat{L}_3$	−0.2967	0.3261	0.4141
60	1	$\hat{L}_0$	−0.0105	0.0309	0.0310
60	1	$\hat{L}_1$	−0.0600	0.0535	0.0571
60	1	$\hat{L}_2$	−0.1795	0.0174	0.0497
60	1	$\hat{L}_3$	−0.0730	0.0450	0.0512
60	5	$\hat{L}_0$	−0.0195	0.0013	0.0017
60	5	$\hat{L}_1$	0.0443	0.0010	0.0030
60	5	$\hat{L}_2$	−0.0215	0.0006	0.0011
60	5	$\hat{L}_3$	−0.0309	0.0022	0.0032

Table 4: Simulation from exponential distribution of different means and sample size 80.

n	Mean	Estimators	Bias	Variance	MSE
80	0.5	$\hat{L}_0$	0.0833	0.0600	0.0669
80	0.5	$\hat{L}_1$	0.0319	0.0597	0.0607
80	0.5	$\hat{L}_2$	0.1759	0.2009	0.2317
80	0.5	$\hat{L}_3$	−0.0866	0.1025	0.1100
80	1	$\hat{L}_0$	0.0392	0.0606	0.0622
80	1	$\hat{L}_1$	0.0536	0.0478	0.0507
80	1	$\hat{L}_2$	0.2112	0.0465	0.0910
80	1	$\hat{L}_3$	−0.1948	0.0242	0.0622
80	5	$\hat{L}_0$	0.0013	0.0011	0.0012
80	5	$\hat{L}_1$	−0.0130	0.0011	0.0013
80	5	$\hat{L}_2$	−0.0516	0.0016	0.0043
80	5	$\hat{L}_3$	−0.0677	0.0018	0.0064

Table 5: Simulation from exponential distribution of different means and sample size 100.

n	Mean	Estimators	Bias	Variance	MSE
100	0.5	$\hat{L}_0$	0.2901	0.0980	0.1822
100	0.5	$\hat{L}_1$	−0.0888	0.2403	0.2482
100	0.5	$\hat{L}_2$	−0.1093	0.1291	0.1411
100	0.5	$\hat{L}_3$	−0.1420	0.1204	0.1405
100	1	$\hat{L}_0$	−0.0415	0.0293	0.0310
100	1	$\hat{L}_1$	−0.0794	0.0119	0.0182
100	1	$\hat{L}_2$	−0.0278	0.0246	0.0253
100	1	$\hat{L}_3$	0.0736	0.0165	0.0220
100	5	$\hat{L}_0$	0.0131	0.0005	0.0006
100	5	$\hat{L}_1$	−0.0217	0.0020	0.0025
100	5	$\hat{L}_2$	−0.0324	0.0012	0.0022
100	5	$\hat{L}_3$	0.0147	0.0005	0.0007

The MRL estimators proposed in this paper seem natural, reasonable, and intuitively appealing. It is shown that the MRL estimators are asymptotically unbiased and consistent. Note also that the simulation study seem to indicate that the MRL estimators have small variance and MSE. The optimal bandwidth using mean integrated squared error criterion in each MRL estimators is $h = cn^{-1/3}$. It is also proven that the choice of a kernel function is less sensitive to the MRL estimators.

ACKNOWLEDGMENTS

The authors thank the associate editor and the referee for their valuable comments on the earlier three versions of this article.

REFERENCES

[1] Abdous, B. and Berred, A. (2005). Mean residual life estimation, *Journal of Statistical Planning and Inference*, **132**, 3–19.

[2] Azzalini, A. (1981). A note on the estimation of a distribution function and quantiles by a kernel method, *Biometrika*, **68**(1), 326–328.

[3] Chaubey, Yogendra P. and Sen, Pranab K. (1999). On smooth estimation of mean residual life, *Journal of Statistical Planning and Inference*, **75**, 223–236.

[4] Duin, R.P.W. (1976). On the choice of smoothing parameters for Parzen estimators of probability density functions, *IEEE Trans. Comput.*, **C-25**, 1175–1179.

[5] Geisser, S. (1975). The predictive sample reuse method with applications, *J. Amer. Statist. Assoc.*, **70**, 320–328.

[6] Guillamon, A.; Navarro, J. and Ruiz, J.M. (1998). Nonparametric estimator for the mean residual life and vitality function, *Statistical Papers*, **39**, 263–276.

[7] Habbema, J.D.F.; Hermans, J. and van der Broek, K. (1974). *A stepwise discrimination program using density estimation.* In "Compstat" (G. Bruckman, Ed.), Physica Verlag, Vienna.

[8] Hollander, M. and Proschan, F. (1984). *Nonparametric concepts and methods in reliability.* In "Handbook of Statistics" (P.R. Krishnaiah and P.K. Sen, Eds.), **4**, 613-655.

[9] Lo, S.H.; Mack, Y.P. and Wang, J.L. (1989). Density and hazard rate estimation for censored data via strong representation of the Kaplan-Meier estimator, *Probab. Theory Rel. Fields*, **80**, 461–473.

[10] Marron, J.S. and Padgett, W.J. (1987): Asymptotically optimal bandwidth selection for kernel density estimators from randomly right censored samples, *Ann. Statist.*, **15**, 1520–1535.

[11] Mitra, M. and Basu, S. (1995). Change point estimation in non-monotonic aging models, *Ann. Inst. Statist. Math.*, **47**(3), 483–491.

[12] Nadaraya, E.A. (1964). Some new estimates for distribution functions, *Theory Prob. Appl.*, **9**, 497–500.

[13] Patil, P.N.; Wells, M.T. and Marron, J.S. (1993). Some heuristics of kernel based estimators of ratio functions, *J. Nonparametric Statistics*, **4**, 203–209.

[14] Ruiz, Jose M. and Guillamon, A. (1999). Nonparametric Recursive Estimator for Mean Residual Life and Vitality Function under dependence conditions, *Commun. Statist. – Theory and Methods*, **25**(9), 1997–2011.

[15] Ruiz, Jose M. and Navarro, J. (1994). Characterization of Distributions By Relationships Between Failure Rate and Mean Residual Life, *IEEE Transactions on Reliability*, **43**(4), 640–644.

[16] Sarda, P. and Vieu, P. (1991). Smoothing parameter selection in hazard estimation, *Statistics & Probability Letters*, **11**, 429–434.

[17] Swanepoel, Jan W.H. (1988). Mean integrated squared error properties and optimal kernels when estimating a distribution function, *Commun. Statist. – Theory Meth.*, **17**(11), 3785–3799.

[18] Watson, G.S. and Leadbetter, M.R. (1964). Hazard analysis I, *Biometrika*, **51**, 175–184.

[19] Watson, G.S. and Leadbetter, M.R. (1964). Hazard analysis II, *Sankhya Ser. A*, **26**, 101–116.