Journal of Statistical Planning and Inference 75 (1999) 223–236

journal of statistical planning and inference

On smooth estimation of mean residual life

Yogendra P. Chaubey [a,*], Pranab K. Sen [b]

[a] Department of Mathematics & Statistics, Concordia University, Montreal, Quebec, Canada

[b] University of North Carolina, Chapel Hill, NC, USA

Abstract

The methodology developed in Chaubey and Sen (1996), (*Statistics and Decision*, **14**, 1–22) is adopted here for smooth estimation of mean residual life. It is seen that Hille (1948), (*Functional Analysis and Semigroups*, AMS, New York) theorem, which has been vital in the development of smooth estimators of the distribution, density, hazard and cumulative hazard functions, does not work well in the current context. For this reason a modified weighting scheme is proposed for estimation of the mean residual life. Asymptotic properties of the resulting estimator is investigated along with its aging aspects. © 1999 Elsevier Science B.V. All rights reserved.

AMS classifications: 62G07; 62G20

Keywords: Exponential smoothing; Hille's theorem; Strong consistency; Survival function; Tail-behavior

1. Introduction

Let T be a non-negative random variable with *distribution function* (d.f.) $F(\cdot)$, and *density function* $f(\cdot)$. Denote the associated *survival function* (s.f.) and *hazard function* by $S(t)=1-F(t)$ and $h(t)=f(t)/S(t)=-(\mathrm{d}/\mathrm{d}t)\log S(t)$, respectively. Further, $H(t)$, the *cumulative hazard function* (c.h.f.), is defined by $H(t)=\int_0^t h(y)\,\mathrm{d}y=-\log S(t),\ t\geqslant 0$. All these concepts have been extensively used in *reliability and survival analysis*. Another function, known as *life expectancy at age t* in life table analysis, and *mean residual life* (MRL) in survival analysis is defined as

$$\begin{aligned} m(t) &= E(T-t|T>t) \\ &= \frac{1}{S(t)}\int_t^\infty S(y)\,\mathrm{d}y. \end{aligned} \tag{1.1}$$

The reader may be referred to Guess and Proschan (1988) for an extensive review of this function. In this paper we are concerned with estimation of $m(t)$ when $f(t)$ is

* Correspondence address: Department of Mathematics & Statistics, Concordia University, Montreal, Quebec, Canada. E-mail: chaubey@vax2.concordia.ca.

0378-3758/99/$ – see front matter
PII: S0378-3758(98)00144-X

unspecified. We here note that Yang (1978) proposed estimating $m(t)$ by replacing S and F by their corresponding estimators based on the empirical d.f. F_n (to be defined later on) which results in the estimator $\hat{m}(t)$ given below

$$\hat{m}(t)=\begin{cases} \frac{1}{n-k}\sum_{i=k+1}^{n}(T_{n:i}-t) & \text{for } T_{n:k}\leqslant t<T_{n:k+1},\ k<n, \\ 0 & \text{for } t\geqslant T_{n:n}, \end{cases} \tag{1.2}$$

where $(0=T_{n:0}\leqslant)T_{n:1}\leqslant\cdots\leqslant T_{n:n}(<T_{n:(n+1)}=+\infty)$ stand for the ordered random variables associated with the observable random variables $T_1,T_2,\ldots,T_n$. Note the limitation of the above estimator in not being able to estimate $m(t)$ beyond the largest order statistic.

Despite the nice properties of F_n as an estimator of F, there has been considerable interest in its smooth estimation from the point of view of nonparametric inference. This lends itself to nonparametric estimation of derived functionals such as $f(\cdot),h(\cdot),H(\cdot)$ and $m(\cdot)$. We refer to Devroye (1989) and Wertz (1978) for density estimation, Nadaraya (1964,1965) for the estimation of distribution function, Watson and Leadbetter (1963,1964), Rice and Rosenblatt (1976), Singpurwalla and Wong (1983a, b) and Patil (1993) for that of hazard and related functionals.

Chaubey and Sen (1996) formulated an alternative approach based on the classical Hille (1948) theorem (on uniform smoothing in real analysis), and obtained smooth estimators of $S(\cdot)$ and $f(\cdot)$ which may have some advantage over their counterparts based on the usual *kernel* method of smoothing. They further studied (Chaubey and Sen 1997) the use of the resulting estimator in smooth estimation of the hazard and cumulative hazard functions.

The purpose of the present study is to investigate the properties of a modification of the above estimator for $m(t)$ obtained by replacing S in Eq. (1.1) by appropriate smooth versions. This is provided in Section 2, and Section 3 presents the asymptotic properties of the resulting estimator. We will see that our estimator of $m(t)$ converges to zero for large values of t. Therefore, the asymptotic properties are restricted for $t\in\mathscr{C}$, $\mathscr{C}$ being a compact interval in $\mathbb{R}^+$. Section 4 investigates the behavior of the proposed estimator for values of t near the largest order statistic.

2. A smooth estimator of mean residual life

For a random sample $(T_1,\ldots,T_n)$ of size n from the d.f. F, the empirical d.f. F_n and empirical s.f. S_n are defined by

$$S_n(t)=1-F_n(t)=n^{-1}\sum_{i=1}^{n}I\{T_i>t\},\quad t\in\mathbb{R}^+.$$

so that

$$S_n(t)=(n-k)/n\quad \text{for } T_{n(k)}\leqslant t<T_{n(k+1)},\ k=0,1,\ldots,n,$$

where $T_{n:i}$, $0 \leqslant i \leqslant (n+1)$ are the order statistics defined earlier following Eq. (1.3). For every $y \in \mathbb{R}^+$ and $n \geqslant 1$, define a set of nonnegative weights

$$w_{nk}(y) = \left\{ \frac{y^k}{k!} \right\} \Big/ \left\{ \sum_{j=0}^{n} \frac{y^j}{j!} \right\}, \quad 0 \leqslant k \leqslant n\,.$$

Then a smooth estimator of $S(\cdot)$, proposed in Chaubey and Sen (1996), is defined by

$$\tilde{S}_n(t) = \sum_{k=0}^{n} w_{nk}(t\lambda_n) S_n(k/\lambda_n), \quad t \in \mathbb{R}^+, \tag{2.1}$$

where $\lambda_n (>0)$ is so chosen that $\lambda_n \to \infty$ but $n^{-1}\lambda_n \to 0$ as $n \to \infty$. In general, λ_n is chosen data-dependent, which makes the weighting function in Eq. (2.1) stochastic.

Note that by definition $\tilde{S}_n(t)$ is smooth, bounded and continuously differentiable with respect to $t \in \mathbb{R}^+$. Therefore, we have derived estimators for the density, cumulative hazard and hazard functions given, respectively, as

$$\begin{aligned}
\tilde{f}_n(t) &= -(\mathrm{d}/\mathrm{d}t)\tilde{S}_n(t) \\
&= \lambda_n \left\{ \tilde{S}_n(t)[1 - w_{nn}(t\lambda_n)] - \sum_{k=0}^{n} w_{nk}(t\lambda_n) S_n((k+1)/\lambda_n) \right\}, \\
\tilde{H}_n(t) &= -\log \tilde{S}_n(t), \\
\tilde{h}_n(t) &= -(\mathrm{d}/\mathrm{d}t)\tilde{H}_n(t) = \tilde{f}_n(t)/\tilde{S}_n(t), \quad t \in \mathbb{R}^+.
\end{aligned}$$

The choice of λ_n is important in the study of the asymptotic properties of the estimators $\tilde{S}_n$, $\tilde{f}_n$, $\tilde{h}_n$, and $\tilde{H}_n$. If these estimators are sought only for a compact interval $\mathscr{C} = [0, \eta]$, $\eta < \infty$, then $n^{-1}\lambda_n \to 0$ and $\lambda_n \to \infty$ as $n \to \infty$ suffice. On the other hand, in most applications, $S(\cdot)$ may not have a compact support (so that $T_{n:n} \overset{\text{a.s}}{\to} \infty$ as $n \to \infty$), and to gather information on the aging aspects of the MRL, we would like to study the behavior of these estimators when t is 'large'. For this purpose, we need to impose some regularity assumptions. A similar situation arises if we want to study asymptotic normality or related properties of $\tilde{f}_n$, $\tilde{h}_n$, etc. For $\tilde{S}_n$, Chaubey and Sen (1996) considered $\lambda_n = n/T_{n:n}$ (stochastic), so that, $S_n(n/\lambda_n) = S_n(T_{n:n}) = 0$ when T has the support $\mathbb{R}^+$ and $E(T) < \infty$. To cover the general case of a compact support and/or the existence of some αth-order moment ($\alpha > 0$), they also recommended

$$\lambda_n = n(\log\log n)^{-1} T_{n:(n-r_n+1)}^{-1},$$

where

$$r_n = \mathrm{o}(\log\log n).$$

In this setup, we have $\lambda_n = \mathrm{o}(n)$, so that defining

$$n^* = \min\{n, [\lambda_n T_{n:n}]\},$$

we get

$$S_n(k/\lambda_n)=0 \quad \forall k \geqslant n^*.$$

(Note that n^* may be stochastic, but $n^* \leqslant n$, with probability 1.)

Remark 2.1. In practice, we may choose $\lambda_n = c_n n/T_{n:n}$, with c_n obtained from the sample data so that some measure of discrepancy between the smooth estimator and the empirical survival function is minimized, e.g., similar to the least square analysis. We found this to be reasonable as observed through smooth fitting on some randomly generated data sets. A larger numerical study is under investigation.

A plug-in estimator of $m(t)$, replacing $S(t)$ by $\tilde{S}_n(t)$ in Eq. (1.1), thus may be considered as

$$\tilde{m}_n(t)=\frac{1}{\tilde{S}_n(t)}\sum_{k=1}^{n} S_n\left(\frac{n-k}{\lambda_n}\right)\int_t^\infty w_{n(n-k)}(\lambda_n y)\,\mathrm{d}y. \tag{2.2}$$

Below, we show that the $\tilde{S}_n(t)$ is not appropriate for smooth estimation of mean residual life as in the above expression by showing that it diverges. First, note that the numerator in Eq. (2.2)

$$\begin{aligned}
&=\sum_{k=0}^{n-1} S_n\left(\frac{k}{\lambda_n}\right)\left[\int_t^{T_{n:n}} w_{n(n-k)}(y\lambda_n)\,\mathrm{d}y+\int_{T_{n:n}}^\infty w_{n(n-k)}(y\lambda_n)\,\mathrm{d}y\right]\\
&\geqslant S_n\left(\frac{n-1}{\lambda_n}\right)\int_{T_{n:n}}^\infty w_{n(n-1)}(y\lambda_n)\,\mathrm{d}y
\end{aligned}$$

and since, $\tilde{S}_n(t)<1$,

$$\tilde{m}_n(t)\geqslant S_n\left(\frac{n-1}{\lambda_n}\right)\int_{T_{n:n}}^\infty w_{n(n-1)}(y\lambda_n)\,\mathrm{d}y.$$

Furthermore, since $w_{n(n-1)}(y\lambda_n)=(n/y\lambda_n)w_{nn}(y\lambda_n)$, choosing $\lambda_n=n/T_{n:n}$ and making a transformation from y to $v=y/T_{n:n}$ we have

$$\int_{T_{n:n}}^\infty w_{n(n-1)}(y\lambda_n)\,\mathrm{d}y=T_{n:n}\int_1^\infty \frac{1}{v}\left\{1+\sum_{k=1}^{n}\frac{n^{[k]}}{n^k v^k}\right\}^{-1}\mathrm{d}v, \tag{2.3}$$

where $n^{[k]}=n!/(n-k)!$. Now, since $\sup_{v\geqslant 1}\sum_{k=1}^n n^{[k]}/n^k v^k$ is equal to $\sum_{k=1}^n n^{[k]}/n^k$, the integral in Eq. (2.3) diverges, hence the numerator of the estimator in Eq. (2.2) is not well defined. In what follows, we propose to modify $\tilde{S}_n(t)$ for the purpose of smooth estimation of $m(t)$ in one of the following two ways.

(i) Define

$$\tilde{S}_n^0(t)=\sum_{k\geqslant 0} S_n\left(\frac{k}{\lambda_n}\right) w_{nk}^0(t\lambda_n),$$

where

$$w^0_{nk}(t\lambda_n)=\mathrm{e}^{-t\lambda_n}\frac{(t\lambda_n)^k}{k!},\quad k=0,1,2,\ldots,\infty. \tag{2.4}$$

This provides a smooth estimator of $m(t)$ given by

$$\tilde{m}^0_n(t)=\frac{1}{\lambda_n}\frac{\sum_{k=0}^{n}\sum_{r=0}^{k}((t\lambda_n)^{k-r}/(k-r)!)S_n(k/\lambda_n)}{\sum_{k=0}^{n}((t\lambda_n)^k/k!)S_n(k/\lambda_n)}, \tag{2.5}$$

while preserving the strong convergence property of $\tilde{S}^0_n(t)$ (see the appendix for asymptotic equivalence of $\tilde{S}^0_n(t)$ and $\tilde{S}_n(t)$). The formula given in Eq. (2.5) can be established as follows. Note that the numerator of Eq. (2.2) with w_{nk} replaced by w^0_{nk} is equal to

$$\sum_{k=0}^{n}\frac{1}{k!}S_n\left(\frac{k}{\lambda_n}\right)\int_t^{\infty}\mathrm{e}^{-\lambda_n y}(\lambda_n y)^k\,\mathrm{d}y. \tag{2.6}$$

Using the transformation $y\to v: y=t(1+v)$, we can write the integral in the above expression as

$$\int_t^{\infty}\mathrm{e}^{-\lambda_n y}(\lambda_n y)^k\,\mathrm{d}y=t(t\lambda_n)^k\mathrm{e}^{-t\lambda_n}\int_0^{\infty}\mathrm{e}^{-(t\lambda_n)v}(1+v)^k\,\mathrm{d}v.$$

Next, the binomial expansion of $(1+v)^k$ and the standard Gamma integral simplify the above equation as

$$\int_t^{\infty}\mathrm{e}^{-\lambda_n y}(\lambda_n y)^k\,\mathrm{d}y=t(t\lambda_n)^k\mathrm{e}^{-t\lambda_n}\sum_{r=0}^{k}\frac{k^{[r]}}{(t\lambda_n)^{r+1}}. \tag{2.7}$$

Using Eq. (2.7) in Eq. (2.6) gives Eq. (2.5).

Another computational formula for $\tilde{m}^0_n(t)$ similar to one established in Chaubey and Sen (1997) for $\tilde{S}_n(t)$ as an alternative to Eq. (2.5) may be given as follows.

$$\tilde{m}^0_n(t)=\frac{(1/\lambda_n)\sum_{k=n-n^*+1}^{n}n^{[k]}/(t\lambda_n)^k\sum_{j=n-n^*+1}^{k}S_n((n-j)/\lambda_n)}{\sum_{k=n-n^*+1}^{n}(n^{[k]}/(t\lambda_n)^k)S_n((n-k)/\lambda_n)}$$

where $n^*\leqslant n$ is an integer such that $S_n(k/\lambda_n)=0$ for $k\geqslant n^*$.

(ii) Define

$$\tilde{S}^1_n(t)=\sum_{k\geqslant 0}S_n\left(\frac{k}{\lambda_n}\right)w^1_{nk}(t\lambda_n)$$

where

$$w^1_{nk}(t\lambda_n)=\begin{cases} w_{nk}(t\lambda_n) & \text{if } t\leqslant T_{n:n},\\ 0 & \text{if } t>T_{n:n}\ \forall k\leqslant n-1.\end{cases}$$

It is easily seen that replacing the weights $w_{nk}(t\lambda_n)$ in Eq. (2.2) by $w^1_{nk}(t\lambda_n)$ forces $\tilde{S}^1_n(t)=0$, for $t>T_{n:n}$, similar to the case of the empirical distribution function and a jump discontinuity is introduced at the point of the largest order statistics. In this case the mean residual life is estimated only for $t<T_{n:n}$. Thus, the smooth estimator of $m(t)$ will be defined as

$$\tilde{m}_n^1(t)=\begin{cases} \frac{1}{\tilde{S}_n^1(t)}\sum_{k=1}^{n} S_n(\frac{k}{\lambda_n})\int_t^{T_{n:n}} w^1_{n(n-k)}(\lambda_n y)\,\mathrm{d}y & \text{if } t<T_{n:n}\\ 0 & \text{otherwise}\end{cases}$$

The choice of weights given by Eq. (2.4), however, can be used to estimate $m(t)$ which will yield a nonzero estimate of mean residual life even for values of $t\geqslant T_{n:n}$ as is clear from Eq. (2.5).

3. Asymptotic properties of $\tilde{m}_n^0(\cdot)$

First, we establish strong consistency of the estimators over a compact interval in the form of following theorem.

Theorem 3.1. *Let the pdf f be defined on $\mathbb{R}^+$ such that $E(T)<\infty$ and $m(t)<\infty$ for all $t\in\mathbb{R}^+$. Then for any compact interval $\mathscr{C}\subset\mathbb{R}^+$, such that $S(t)>0$ for $t\in\mathscr{C}$, we have*

$$\|\tilde{m}_n^0-m\|_{\mathscr{C}}=\sup_{t\in\mathscr{C}}|\tilde{m}_n^0(t)-m(t)|\rightarrow 0 \text{ a.s.} \quad \text{as } n\rightarrow\infty. \tag{3.1}$$

if $\lambda_n=\mathrm{o}(n)$.

Proof. If F has a compact support, i.e., $S(t)=0$ for all $t>t_0(<\infty)$, then $m(t)$ is not defined beyond t_0 and for all $t<t_0$, $m(t)$ is bounded. Therefore, $m(t)$ is a Hadamard-differentiable function of $S(t)$. Hence, from the proposition in the appendix we obtain the same convergence property for $\tilde{m}_n^0(t)$. Therefore, we confine ourself to the case of infinite support. First note that $\tilde{m}_n^0(t)$ can be expressed as

$$\tilde{m}_n^0(t)=\frac{(1/\lambda_n)\sum_{k=0}^{n} S_n(k/\lambda_n)W^0_{nk}(t\lambda_n)}{\sum_{k=0}^{n-1} S_n(k/\lambda_n)w^0_{nk}(t\lambda_n)} \tag{3.2}$$

where $W^0_{nk}(t\lambda_n)=\sum_{r=0}^{k} w^0_{nr}(t\lambda_n)$.

The numerator in Eq. (3.2) can be written as $\sum_{k=0}^{n-1}\tilde{G}_n^0(k/\lambda_n)w^0_{nk}(t\lambda_n)$ where $\tilde{G}_n^0(k/\lambda_n)=(1/\lambda_n)\sum_{j\leqslant k} S_n(j/\lambda_n)$. Note further that

$$\tilde{G}_n^0\left(\frac{k}{\lambda_n}\right)-\frac{1}{\lambda_n}S_n\left(\frac{k}{\lambda_n}\right)\leqslant G_n\left(\frac{k}{\lambda_n}\right)\leqslant \tilde{G}_n^0\left(\frac{k}{\lambda_n}\right),$$

where $G_n(t)=\int_t^\infty S_n(x)\,\mathrm{d}x$. Hence,

$$\sup_{t\in\mathscr{C}}\left|G_n\left(\frac{k}{\lambda_n}\right)-\tilde{G}_n^0\left(\frac{k}{\lambda_n}\right)\right|\rightarrow 0 \text{ a.s.} \quad \text{as } n\rightarrow\infty,$$

where $\mathscr{C}=[0,T]$. Further, since, $G_n(t)$ is non-decreasing and $G_n(0)$ is bounded a.s., using Hille's theorem as in Chaubey and Sen (1996), we get

$$\sum_{k\geqslant 0} w_{nk}^0(t\lambda_n)\tilde{G}_n^0(k/\lambda_n)\to G_n(t) \text{ a.s.} \quad \text{as } n\to\infty \ \forall t\in\mathscr{C}. \tag{3.3}$$

Now using Lemma A in Barlow et al. (1972), (p. 237) we have

$$G_n(t)\to \int_t^{\infty} S(x)\,\mathrm{d}x \text{ a.s.} \quad \text{as } n\to\infty. \tag{3.4}$$

Combining Eqs. (3.3) and (3.4) along with the fact that $\tilde{S}_n^0(t)^{-1}\to S(t)^{-1}$ (see Sen and Singer, 1993), uniformly in any compact interval we obtain Eq. (3.1).

We may also use Theorem 3.2 of Chaubey and Sen (1996) in establishing the asymptotic normality of $\tilde{m}_n^0(t)$ as given in the following theorem.

Theorem 3.2. *Let the pdf f be defined on $\mathbb{R}^+$ such that $E(T)<\infty$ and $m(t)<\infty$. Then for any compact interval $\mathscr{C}\subset\mathbb{R}^+$, such that $S(t)>0$ for $t\in\mathscr{C}$, and choosing $\lambda_n=\mathrm{o}(n)$, we have*

$$\lambda_n^{1/2}(\tilde{m}_n^0(t)-m(t))\to^{\mathscr{D}}\mathscr{N}\left(0,\frac{m(t)}{S(t)}\right). \tag{3.5}$$

Proof. Again as in Theorem 3.2, the proof goes through for any compact interval $\mathscr{C}=[0,\eta]$ where η is fixed but may be large, such that $S(t)>0$ for $t\in\mathscr{C}$. Define

$$K_n(t)=\sum_{k=0}^{n} W_{nk}^0(t\lambda_n)S_n\left(\frac{k}{\lambda_n}\right)$$

which may be written as the mean of a sequence of i.i.d. random variables $\{\xi_{ni}\colon i=1,\ldots,n\}$, where

$$\xi_{ni}=\sum_{k=0}^{n} W_{nk}^0(t\lambda_n)I_{[T_i>k/\lambda_n]}.$$

The mean and variance of ξ_{ni} may easily be calculated; they are respectively given as

$$E(\xi_{ni})=\sum_{k=0}^{n} W_{nk}^0(t\lambda_n)S\left(\frac{k}{\lambda_n}\right),$$

$$\mathrm{Var}(\xi_{ni})=\sum_{k=0}^{n} W_{nk}^0(t\lambda_n)S\left(\frac{k}{\lambda_n}\right)\left[\sum_{k=0}^{n} W_{nk}^0(t\lambda_n)\left\{1-S\left(\frac{k}{\lambda_n}\right)\right\}\right]. \tag{3.6}$$

Using Eq. (3.6), it is easy to see that

$$\begin{aligned}\mathrm{Var}\left(\frac{K_n(t)}{\lambda_n}\right)=\frac{\sum_{k=0}^{n} W_{nk}^0(t\lambda_n)S\left(\frac{k}{\lambda_n}\right)}{\lambda_n^2}&\left[\frac{n+1}{n}\sum_{k=0}^{n} w_{nk}^0\,(t\lambda_n)\right.\\ &\left.-\frac{1}{n}\sum_{k=0}^{n} kw_{nk}^0\,(t\lambda_n)-\frac{1}{n}\sum_{k=0}^{n} W_{nk}^0(t\lambda_n)S\left(\frac{k}{\lambda_n}\right)\right].\end{aligned} \tag{3.7}$$

Using the facts that
(i) $\lim_{n\to\infty}\sum_{k=0}^{n} w_{nk}^0(t\lambda_n)=1$,
(ii) $\lim_{n\to\infty}\left|\sum_{k=0}^{n} kw_{nk}^0(t\lambda_n)-t\lambda_n\right|=0$
along with Eqs. (3.3), (3.4) and (3.7) and almost sure convergence of $\tilde{S}_n^0(t)$ to $S(t)$ we conclude

$$\mathrm{Var}(\tilde{m}_n^0(t))\simeq\frac{m(t)}{S(t)}\left[\frac{n+1}{n\lambda_n}-\frac{t+m(t)S(t)}{n}\right]. \tag{3.8}$$

Further using Eqs. (3.6), (3.8), almost sure convergence of $\tilde{S}_n^0(t)$ to $S(t)$ and the properties $\lambda_n\to\infty$ and $n^{-1}\lambda_n\to 0$ proves Eq. (3.5).

Remark 3.1. In Theorems 3.1 and 3.2 we have implicitly assumed a nonstochastic choice of λ_n; however, the results hold true if we replace the condition $\lambda_n=\mathrm{o}(n)$ by $\lambda_n\overset{\text{a.s.}}{=}\mathrm{o}(n)$ as $n\to\infty$.

The asymptotic behavior of $\tilde{m}_n^0(t)$ and that of $\tilde{m}_n^1(t)$ can be easily shown to be the same. The estimator $\tilde{m}_n^0(t)$, however, has the flexibility of providing estimates beyond the largest order statistic. The almost sure convergence of $\tilde{m}_n^0(t)$ may not be uniform. Hence, it is of interest to examine the behavior of $\tilde{m}_n^0(t)$ for large values of t, specially near $T_{n:n}$. This is pursued in the following section.

4. Behavior of $\tilde{m}_n^0(t)$ for large values of t

The tail behavior of MRL depends very much on that of the underlying density. We may consider the two important classes of such densities, known as IMRL (increasing mean residual life) class and DMRL (decreasing mean residual life) class. As we shall subsequently see the proposed estimator $\tilde{m}_n^0(t)$ is nonincreasing in the tail and hence may not be appropriate for the IMRL class. Even the derived tail behavior may not be appropriate for the DMRL class. Therefore, some modifications are necessary in either case. We will establish a lemma relating the hazard function ($h(t)$) and MRL ($m(t)$) for the DMRL class which is useful in providing the modification. Chaubey and Sen (1997) studied the behavior of the hazard function and related functions for large t. This has to be reexamined in the light of modified weights $w_n^0(t\lambda_n)$. We establish the following behavior of $\tilde{S}_n^0(t)$ and $\tilde{h}_n^0(t)$ for large $t\geqslant n\lambda_n$. First, note that with $n^*=\min[n,\{\lambda_n T_{n:n}\}]$, so that $S_n(k/\lambda_n)=0$, $\forall k\geqslant n^*$, letting $n_0=n^*-1$, we have,

$$\tilde{S}_n^0(t)=\sum_{k\geqslant 0} w_{nk}^0(t\lambda_n)S_n\left(\frac{k}{\lambda_n}\right)$$

$$=\mathrm{e}^{-t\lambda_n}\sum_{k=0}^{n_0}\frac{(t\lambda_n)^k}{k!}S_n\frac{k}{\lambda_n}$$

$$= e^{-t\lambda_n}\left[S_n\frac{n_0}{\lambda_n}\frac{(t\lambda_n)^{n_0}}{n_0!}+S_n\frac{n_0-1}{\lambda_n}\frac{(t\lambda_n)^{n_0-1}}{(n_0-1)!}+\cdots\right]$$

$$= e^{-t\lambda_n}S_n\frac{n_0}{\lambda_n}\frac{(t\lambda_n)^{n_0}}{n_0!}\left[1+O\left(\frac{n_0}{t\lambda_n}\right)\right]. \tag{4.1}$$

From Eq. (4.1), we conclude that for t such that when $t\lambda_n \gg n$,

$$\tilde{S}_n^0(t) \approx e^{-t\lambda_n}S_n\left(\frac{n_0}{\lambda_n}\right)\frac{(t\lambda_n)^{n_0}}{n_0!}.$$

Using the relation $H(t) = -\log_e S(t)$ and $h(t) = dH(t)/dt$, we obtain, approximately, for large t,

$$\tilde{h}_n^0(t) \approx \lambda_n - \frac{n_0}{t\lambda_n}.$$

This shows that $\tilde{h}_n^0(t)$ increases with t for large t and hence we conclude that it may be appropriate in case of IFR family of distributions. In passing we observe that since

$$m(t) = \int_0^\infty e^{-[H(t+u)-H(t)]}\,du \tag{4.2}$$

and for the IFR family of distributions, $m(t)$ is a decreasing function of t, which makes the IFR class a subclass of the DMRL family. Analogously, a DFR family of distributions is a subclass of an IMRL family. These two classes will be investigated for appropriateness of our estimator for large t. Using Eq. (4.2), we get for large t

$$\begin{aligned}\tilde{m}_n^0(t) &= \frac{1}{\tilde{S}_n^0(t)}\int_0^\infty \tilde{S}_n^0(t+u)\,du\\ &\approx \int_0^\infty e^{-u\lambda_n}\left(1+\frac{u}{t}\right)^{n_0}du\\ &= \sum_{j=0}^{n_0}\binom{n_0}{j}\frac{j!}{\lambda_n(t\lambda_n)^j}\\ &\approx \frac{1}{\lambda_n}\left[1+\frac{n_0}{t\lambda_n}\right].\end{aligned}$$

This shows that our estimator may be appropriate for DMRL family for large t. However, near $T_{n:n}$, it may not be appropriate for either the IMRL or DMRL classes. We propose a modification based on the behavior of $\tilde{h}_n^*(t)$ in an interval $T_{n:\tilde{n}} \leqslant t \leqslant T_{n:n}$ where $\tilde{n} = [\alpha n]$ for some α less than but close to 1 (e.g. $\alpha = 0.9$), i.e. where the estimator $\tilde{h}_n^*(t)$ of $h(t)$ is based on a modified estimator of $S(t)$

$$S_n^*(t) = \sum_{k=\tilde{n}} w_{n(k-\tilde{n})}^0(t\lambda_n)S_n\left(\frac{k}{\lambda_n}\right).$$

The estimator of mean residual life is then obtained using the following expression for values of t beyond $T_{n:n}$:

$$m_n^*(t) = \frac{1}{h_n^*(t)} - \frac{h_n^{*\prime}(t)}{\{h_n^*(t)\}^3}$$

which is suggested by the following lemma.

Lemma 4.1. *Let the probability density function $f(t)$ belong to the family of IFR with the hazard function being continuously differentiable. Then for large values of t, we may approximate $m(t)$ as*

$$m(t) \approx \frac{1}{h(t)} - \frac{h'(t)}{(h(t))^3}. \tag{4.3}$$

Proof. Writing

$$H(t+u) - H(t) = uh(t) + \tfrac{1}{2}u^2 h'(t + \xi u)$$

where $0 < \xi < 1$, and using Eq. (4.4) we can write

$$m(t) = \int_0^\infty \mathrm{e}^{-uh(t)}[1 - \tfrac{1}{2}u^2 h'(t+\xi u) + \tfrac{1}{8}u^4\{h'(t+\xi u)\}^2 + \cdots]\,\mathrm{d}u. \tag{4.4}$$

Expanding further $h'(t+\xi u)$ and using Gamma-integral in Eq. (4.4) we get

$$m(t) = \frac{1}{h(t)} - \frac{h'(t)}{(h(t))^3} + \mathrm{O}\left(\frac{1}{(h(t))^4}\right),$$

which proves Eq. (4.3).

Below, we verify the validity of this approximation for some standard distributions by explicitly expanding $m(t)$ and $h(t)$ for large t.

Examples

(i) *Folded normal distribution*: Denoting by $\phi(x)$ and $\Phi(x)$ the density and the distribution function of a standard normal distribution we have

$$f(t) = \sqrt{\frac{2}{\pi}}\mathrm{e}^{-(1/2)t^2} = 2\phi(x), \quad x \geqslant 0,$$
$$S(t) = 2(1 - \Phi(t)).$$

The mean residual life is given by

$$m(t) = -t + \frac{\phi(t)}{1 - \Phi(t)}. \tag{4.7}$$

Since

$$1 - \Phi(t) \simeq \phi(t)\frac{1}{t}\left[1 - \frac{1}{t^2} + \frac{3}{t^4} - \cdots\right], \tag{4.8}$$

for large t, using Eq. (4.7) we get

$$m(t) \simeq \frac{1}{t} - \frac{2}{t^3}. \tag{4.9}$$

Furthermore, using Eq. (4.8), we have, for large t,

$$\begin{aligned} h(t) &= t\left[1 - \frac{1}{t^2} + \frac{3}{t^4} - \cdots\right]^{-1} \\ &\simeq t + \frac{1}{t} - \frac{2}{t^3} \end{aligned} \tag{4.10}$$

and

$$h'(t) \simeq 1 - \frac{1}{t^2} + \frac{6}{t^4}. \tag{4.11}$$

Using Eqs. (4.10) and (4.11), we get

$$\frac{1}{h(t)} - \frac{h'(t)}{h^3(t)} \simeq \frac{1}{t} - \frac{2}{t^3},$$

which matches with Eq. (4.9).

(ii) *Pareto distribution*: Here

$$\begin{aligned} S(t) &= ct^{-v}, \quad t \geqslant 1, \ v > 1, \ c > 0, \\ H(t) &= v \log t - \log c, \end{aligned}$$

Hence $m(t)$ is given by

$$m(t) = \frac{t}{v-1}. \tag{4.12}$$

We note that since $h(t) = v/t$, this is a case of DFR family, and

$$\begin{aligned} \frac{1}{h(t)} - \frac{h'(t)}{h^3(t)} &= \frac{t}{v} + \frac{t}{v^2} \\ &\simeq \frac{t}{v-1}, \quad \text{if } v \text{ is large,} \end{aligned}$$

which matches with (4.12). This example shows that the approximation proposed in Lemma 4.1 may even be appropriate for some IMRL family.

(iii) *Gamma distribution*: In this case the density and the survival distributions are given by

$$\begin{aligned} f(t) &= \frac{1}{\Gamma(\alpha)} e^{-t} t^{\alpha-1}, \quad t \geqslant 0, \ \alpha > 1, \\ S(t) &= \frac{1}{\Gamma(\alpha)} \int_t^\infty e^{-u} u^{\alpha-1} \, du, \end{aligned}$$

Assuming, for some integer j, $\alpha - j > 0$ we can write

$$S(t) \equiv S_\alpha(t) = \sum_{k=0}^{\alpha-j+1} f_k(t) + S_{\alpha-j}(t),$$

which gives

$$h(t)=\frac{f_\alpha(t)}{\sum_{k=0}^{\alpha-j+1} f_k(t)+S_{\alpha-j}(t)}.$$

Hence, we conclude that $h(t)\leqslant 1$, and for $\alpha>j>1$

$$\begin{aligned} h(t) &= \left[1+\frac{f_{\alpha-1}(t)}{f_\alpha(t)}+\frac{f_{\alpha-2}(t)}{f_\alpha(t)}+\cdots\right]^{-1} \\ &= \left[1+\frac{\alpha-1}{t}+\frac{(\alpha-1)(\alpha-2)}{t^2}+\cdots\right]^{-1} \\ &\simeq 1-\frac{\alpha-1}{t}+\frac{\alpha-1}{t^2}. \end{aligned} \tag{4.13}$$

It can be shown in the present case

$$m(t)=\alpha+t(h(t)-1)). \tag{4.14}$$

From Eqs. (4.13) and (4.14) we get for large t,

$$m(t)\simeq 1+\frac{\alpha-1}{t}. \tag{4.15}$$

Using Eq. (4.13) again, up to order $(1/t)$ we get

$$\frac{1}{h(t)}-\frac{h'(t)}{h^3(t)}\simeq 1+\frac{\alpha-1}{t}$$

which matches with the expression in Eq. (4.15) confirming the validity of Lemma 4.1 for $\alpha>1$.

(iv) *Weibull distribution*: For this family of distributions, we have

$$S(t)=\exp[-t^\gamma], \quad \gamma>0,$$
$$h(t)=\gamma t^{\gamma-1}$$

and

$$h'(t)=\gamma(\gamma-1)t^{\gamma-2}.$$

For $\gamma>1$, the above family belongs to IFR class and therefore we can carry out calculations as in the proof of Lemma 4.1 and obtain

$$m(t)\simeq\frac{1}{\gamma t^{\gamma-1}}-\frac{\gamma-1}{\gamma^2 t^{2\gamma-1}}$$

which matches with the approximation

$$\frac{1}{h(t)}-\frac{h'(t)}{h^3(t)}\simeq\frac{1}{\gamma t^{\gamma-1}}-\frac{\gamma(\gamma-1)t^{\gamma-2}}{\gamma^3 t^{3\gamma-3}}.$$

For the case $0<\gamma<1$ we cannot use the expansion as in the proof of Lemma 4.1. However, we observe that for this case we can still use the following approximation for $\exp[-(H(t+u)-H(t))]$ for large t,

$$\begin{aligned}\exp[-(H(t+u)-H(t))] &= \mathrm{e}^{-u\gamma t^{\gamma-1}}\mathrm{e}^{(1/2)u^2(\gamma(1-\gamma)/(t+\xi u)^{(2-\gamma)})}\\ &\simeq \mathrm{e}^{-u\gamma t^{\gamma-1}}\left[1+\frac{1}{2}u^2\frac{\gamma(1-\gamma)}{t^{2-\gamma}(1+\frac{\xi u}{t})^{2-\gamma}}\right]\\ &\simeq \mathrm{e}^{-u\gamma t^{\gamma-1}}\left[1+\frac{1}{2}u^2\frac{\gamma(1-\gamma)}{t^{2-\gamma}}\right]\end{aligned}$$

and hence, again we get from Eq. (4.4), for large t

$$m(t)\simeq\frac{1}{\gamma t^{\gamma-1}}+\frac{(1-\gamma)}{\gamma^2 t^{2\gamma-1}},$$

which is the same as in Eq. (4.34). Thus we observe that for $0<\gamma<1$ also, the MRL function can be approximated using the expression given in Lemma 4.1.

Appendix

Proposition. *Let the pdf f be defined on $\mathbb{R}^+$ such that $E(T)<\infty$ and $H(T)<\infty$ for all $t\in\mathbb{R}^+$. Then for any compact interval $\mathscr{C}\subset\mathbb{R}^+$, we have*

$$\|\tilde{S}_n^0-S\|=\sup|\tilde{S}_n^0(t)-S(t)|\to 0 \ \textit{a.s.} \quad \textit{as } n\to\infty. \tag{A.1}$$

Proof. Note that we can write

$$\tilde{S}_n^0(t)=\tilde{S}_n(t)K_n^0(t\lambda_n), \tag{A.2}$$

where

$$K_n^0(t\lambda_n)=\sum_{k=0}^{n-1}w_{nk}^0(t\lambda_n). \tag{A.3}$$

From Eqs. (A.2) and (A.3) it follows that for all finite t, $\tilde{S}_n^0(t)<\tilde{S}_n(t)$ since $K_n^0(t\lambda_n)<1, \forall t<\infty$. Further, since $K_n^0(t\lambda_n)=\Pr[X_{t\lambda_n}<n]$, where X_θ denotes a Poisson random variable with mean θ, it follows that

$$K_n^0(t\lambda_n)\to 1 \quad \text{as } n\to\infty, \quad \text{for } t \text{ fixed}, \tag{A.4}$$

and

$$K_n^0(t\lambda_n)\to 0 \quad \text{as } t\to\infty, \quad \text{for } n \text{ fixed}. \tag{A.5}$$

Eqs. (A.4) and (A.5) along with (A.2) imply that

$$\|\tilde{S}_n^0-\tilde{S}_n\|=\sup|\tilde{S}_n^0(t)-\tilde{S}_n(t)|\to 0 \text{ a.s.} \quad \text{as } n\to\infty. \tag{A.6}$$

Eq. (A.6) along with Theorem 3.1 of Chaubey and Sen (1996), for the *strong consistency* of $\tilde{S}_n(t)$, prove Eq. (A.1).

Acknowledgements

The research of the first author (YPC) is partially supported from a research grant from NSERC of Canada. The authors are thankful to a referee for some constructive suggestions on the original manuscript.

References

Barlow, R.E., Bartholomew, D.J., Bremner, J.M., Brunk, H.D., 1972. Statistical Inference Under Order Restrictions. Wiley, New York.

Chaubey, Y.P., Sen, P.K., 1996. On smooth estimation of survival and density functions. Statist. Decisions 14, 1–22.

Chaubey, Y.P., Sen, P.K., 1997. On smooth estimation of hazard and cumulative hazard functions. In: Mukherjee, S.P. et al. (Eds.), Frontiers in Probability and Statistics. Narosa Publishing House, Delhi, India, pp. 92–100.

Devroye, L., 1989. A Course in Density Estimation. Birkhaüser, Boston.

Guess, F., Proschan, F., 1988. Mean residual life, theory and applications. In: Krishnaiah, P.R., Rao, C.R. (Eds.), Handbook of Statistics, vol. 7, Reliability and Quality Control, pp. 215–224.

Hille, E., 1948. Functional Analysis and Semigroups, Am. Math. Soc. Colloq. Pub., vol. 31. American Mathematical Society, New York.

Nadaraya, E.A., 1964. Some new estimates for distribution functions. Theor. Probab. Appl. 9, 497–500.

Nadaraya, E.A., 1965. On non-parametric estimates of density functions and regression curves. Theor. Probab. Appl. 10, 186–190.

Patil, P.N., 1993. Bandwidth choice for nonparametric hazard rate estimation. J. Statist. Plann. Inference 35, 15–30.

Rice, J., Rosenblatt, M., 1976. Estimation of the log survivor function and hazard function. Sankhya 38, 60–78.

Sen, P.K., Singer, J.M., 1993. Large Sample Methods in Statistics. Chapman & Hall, New York.

Singpurwalla, N.D., Wong, M., 1983a. Estimation of the failure rate – A survey of nonparametric methods, Part I: non-Bayesian methods. Commun. Statist. – Theor. Meth. 12 (5), 559–588.

Singpurwalla, N.D., Wong, M., 1983b. Kernel estimator of failure-rate function and density estimation: an analogy. J. Amer. Statist. Assoc. 78, 478–481.

Watson, G.S., Leadbetter, M.R., 1963. Hazard analysis I. Biometrika 51, 175–184.

Watson, G.S., Leadbetter, M.R., 1964. Hazard analysis II. Sankhya A 26, 110–116.

Wertz, W., 1978. Statistical Density Estimation, A Survey. Vandenhoeck & Ruprecht, Göttingen.

Yang, G.L., 1978. Estimation of a biometric function. Ann. Statist. 6, 112–116.