Journal of Statistical Planning and Inference 196 (2018) 115–131

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

The MRL function inference through empirical likelihood in length-biased sampling

Vahid Fakoor *, Ali Shariati, Majid Sarmad

Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Iran

ARTICLE INFO

Article history:
Received 5 September 2016
Received in revised form 10 October 2017
Accepted 7 November 2017
Available online 15 November 2017

Keywords:
Confidence band
Confidence interval
Empirical likelihood
Length-biased data
Mean residual life function

ABSTRACT

In survival analysis or reliability studies, the mean residual life (MRL) function is the other important function to characterize a lifetime alongside the distribution function. In this paper, an empirical likelihood (EL) procedure based on length-biased data is proposed for inference on the MRL function and the asymptotic distribution of the empirical log-likelihood ratio for the MRL function is derived. We use limiting distribution to obtain EL ratio confidence intervals for the MRL function. Moreover, it is shown that the empirical log-likelihood ratio converges weakly to a mean zero Gaussian process. We apply this result to the construction of a Gaussian process approximation based confidence band for the MRL function. Also, a confidence interval for the MRL function is driven by using the normal approximation (NA) method in a length-biased setting. Simulation results are obtained to reveal the better efficiency and accuracy of the empirical likelihood-based confidence intervals in comparison to the proposed normal approximation-based method. A real data application is presented for better illustration.

1. Introduction and preliminaries

Let F be an unknown distribution function (d.f.) of a real-valued random variable (X). The classical nonparametric maximum likelihood estimator of F is simply the empirical d.f. based on independent and identically distributed random variables sampled from F. In biased sampling, instead of the random sample of F, the random sample $Y_1, \dots, Y_n$ is observed from distribution G, which is a result of the biased sampling of F according to some known biasing (or weight) function w. For the nonnegative function $w(\cdot)$, the "w-biased" d.f. G is

$$G(t) = \frac{1}{W}\int_{-\infty}^{t} w(y)dF(y), \quad t \in \mathbb{R}, \tag{1.1}$$

where

$$W = \int w(y)dF(y). \tag{1.2}$$

The biased sampling problem is to estimate F from G on the basis of the random sample $Y_1, \dots, Y_n$. Suppose that F is a d.f. on $\mathbb{R}^+ = [0, \infty)$ with a finite mean μ and let $w(x) = x$ for $x \geq 0$. Hence, $W = \mu$ and according to (1.1), the following equation is obtained:

$$G(t) = \mu^{-1}\int_{0}^{t} ydF(y), \quad t \geq 0. \tag{1.3}$$

* Corresponding author.
E-mail addresses: fakoor@math.um.ac.ir (V. Fakoor), ali-shariati@mail.um.ac.ir (A. Shariati), sarmad@um.ac.ir (M. Sarmad).

https://doi.org/10.1016/j.jspi.2017.11.001
0378-3758/

This is the length-biased distribution corresponding to F and is known as the limiting distribution of "total life" in renewal theory (Feller, 1966, page 371). Moreover, it is easily deduced from (1.3) that if $F(0) = 0$,

$$F(t) = \mu \int_0^t y^{-1} dG(y), \quad t \geq 0,$$

which is the relationship used by Cox (1969) to discuss estimation of F.

Considering the empirical d.f. estimator of G, we have

$$G_n(t) = \frac{1}{n} \sum_{i=1}^{n} I(Y_i \leq t).$$

Cox (1969) proposed the empirical estimator of F by

$$\begin{aligned} F_n(t) &= \mu_n \int_0^t y^{-1} dG_n(y) \\ &= \frac{\mu_n}{n} \sum_{i=1}^{n} \frac{1}{Y_i} I(Y_i \leq t), \end{aligned} \tag{1.4}$$

in which

$$\begin{aligned} \mu_n^{-1} &= \int_0^\infty y^{-1} dG_n(y) \\ &= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{Y_i}, \end{aligned}$$

and where $I(A)$ denotes the indicator of event A.

The phenomenon of biased sampling was initially discovered and recognized by Wicksell (1925) in the field of anatomy. At that time, it was named the corpuscle problem, which later came to be known as length-biased sampling. In many practical experiences, such as survival analysis, renewal processes, epidemiology, econometrics and physics, length-biased data arise, which are obtained from the length-biased distribution. In such cases, the probability of selecting individuals from a target population is proportional to the length of time from diagnosis to the failure. Thus, the observations imply length-biased distribution, $G(\cdot)$, instead of the target population distribution, $F(\cdot)$.

The problem of length-biased data was later studied by Cox (1969), in the context of distribution estimation of fiber lengths in fabric. As Patil and Rao (1978) have represented in their comprehensive article, biased data arise frequently in applications regarding wildlife and human populations and their activities. An interesting overview of nonparametric contributions to the literature on estimation problems where the observations are taken from weighted distributions can be found in Cristóbal and Alcalá (2001).

In life testing situations, the expected additional lifetime of a component, given that it has survived until time t, is a function of t, called the mean residual life. For any d.f. L, τ_L denotes the right endpoint of its support when $\tau_L := \inf\{x : L(x) = 1\}$. Considering the definition, assuming that $\tau_F = \tau_G = \tau < \infty$, the MRL function M_F at $x \geq 0$ is defined by

$$M(x) := E(X - x | X > x) = \frac{I_{[0,\tau)}(x)}{1 - F(x)} \int_x^\infty (1 - F(t))\, dt. \tag{1.5}$$

The MRL function has been studied by various authors. Yang (1978) studied the estimation of $M(\cdot)$ on a fixed interval $0 \leq t \leq \tau < \infty$, and proved that the estimator is strongly uniformly consistent on $[0, \tau]$. Moreover, during this study, it is proven that when the estimator is properly centrally normalized, it converges weakly to a certain Gaussian process on $[0, \tau]$. Furthermore, Hall and Wellner (1979) extended Yang's results over $\mathbb{R}^+ = [0, \infty)$ by introducing suitable metrics. By comparison, Csörgő and Zitikis (1996) studied the statistical inferences for the MRL function over the whole positive half-line, $\mathbb{R}^+ = [0, \infty)$. They described a class of weighted functions which assisted them in establishing and illustrating the strong uniform consistency for MRL processes over the positive half-line as well as weak approximation of MRL processes. Moreover, they studied the constructions of confidence intervals for MRL functions, which are regulated by weights relying upon available information about the underlying distribution function. Proschan and Serfling (1974) and Guess and Proschan (1988) have represented comprehensive reviews for the MRL function.

However, individuals may be subject to biased sampling in various applications including epidemiology, medical follow-up and reliability studies. This bias could be caused by different reasons such as non-random sampling of subjects or confronting with incomplete data. For right-censored data, the consistent estimator of the MRL function along with its asymptotic normal distribution has been presented by Yang (1977) and Kumazawa (1987). Additionally, a more recent work on estimation of MRL with left-truncated and right-censored data has been presented by Zhao et al. (2013). Furthermore, an efficient estimation of the MRL function with length-biased and right-censored data was presented by Wu and Luan (2014).

As a specific case of biased sampling, a consistent estimator of the MRL function, $M_n(\cdot)$, for length-biased data has been introduced, through placing $F_n(\cdot)$ in $M(\cdot)$, and its asymptotic behaviors have been studied by Fakoor (2015). Therefore, by considering $Y_{(n)} = \max_{1\le i\le n} Y_i$ instead of τ, the empirical counterpart of $M(\cdot)$ is defined by

$$M_n(x) := \frac{I_{[0,Y_{(n)})}(x)}{1-F_n(x)} \int_x^\infty (1-F_n(t))\,dt. \tag{1.6}$$

Accordingly, a confidence interval for MRL at time t can be constructed via the NA-method using Fakoor's finding (2015), which will be represented in Section 3.

Owen (1988, 1990) proposed the method of empirical likelihood ratio as an alternative to the bootstrap for constructing confidence regions in nonparametric problems. In addition, DiCiccio et al. (1991) showed that in a very general setting, the EL method has a key advantage over the bootstrap method. It has been proven that the empirical likelihood approach is more accurate than the normal approximation in many situations, particularly when the underlying distribution is non-normal and the variance estimate is unstable. The EL method has numerous applications in various areas of statistics. For example, Zhou and Jeong (2011) applied a general empirical likelihood ratio method to test the mean or median residual life function for right-censored survival data, showing that their method preserves the advantages of the test, when it is inverted to reach confidence region or interval. However, they have not studied the data that are subject to biased sampling. In biased sample problems, the empirical likelihood method has been studied by Qin (1993) and Ning et al. (2013). Zhao and Qin (2006) and Qin and Zhao (2007) developed an empirical likelihood procedure for the inference on confidence intervals for MRL in general and for random right censored data respectively.

Although biased inferences derived from bias sampling have been widely identified in the literature, there does not seem to be any adequate solution for the problem of constructing confidence band for the MRL function, which is why we have investigated this article. Mentioned previously, although Ning et al. (2013) has not presented the result of their confidence interval for MRL function, they studied the problem of drawing inferences for EL-based confidence intervals based on length-biased data in the presence of right censor. By comparison, in addition to using different procedures, while a NA-based method is introduced in our article, Ning et al. (2013) compared their EL-based confidence intervals to bootstrap intervals. From another point of view, we may regard Y as the form of $Y = A+R$ in terms of renewal theory, in which A is the observed current age (backward recurrence time), and R is the residual lifetime (forward recurrence time). Based on the variables A and R, Liang et al. (2016) have defined MRL for a subject observed at time t_0 by

$$M(t_0) = E(R|A > t_0). \tag{1.7}$$

Using (1.7), they have obtained an estimating equation based on unbiased distribution function F, and made inferences on MRL based on EL procedure. Under the condition $E(X^4) < \infty$, Liang et al. (2016) obtained confidence interval for the MRL function and compared it with their proposed NA confidence interval obtained by the delta method. It should be mentioned that EL-based confidence band for the MRL function, has not been considered in Ning et al. (2013) and Liang et al. (2016).

The rest of this paper is organized as follows: The theory of constructing the EL ratio is given in Section 2. In Section 3, confidence intervals for the MRL function are derived, including pointwise confidence intervals for $M(\cdot)$ via the EL approach as well as confidence interval for the MRL function without considering auxiliary information through the NA approach. During this section, the empirical likelihood ratio based confidence band is introduced based on length-biased data. In Section 4, a simulation study is conducted to compare the performance of proposed EL method to the introduced NA method for length-biased data in terms of coverage probability and length of interval. Moreover, the performance of proposed EL method is conducted to a real data set in Section 5. Finally, the proof of the main results has been deferred to the last section, Proofs.

2. Empirical likelihood ratio for MRL

In order to apply the empirical likelihood approach for length-biased data, it is necessary to restrict our attention to the estimation of the biased distribution function $G(\cdot)$ instead of the unbiased distribution $F(\cdot)$. On the other hand, the fundamental principle of the EL method is based on obtaining the empirical likelihood ratio statistics via the Lagrange multiplier method under the specified restrictions. Hence, we need to find the special corresponding constraint to evaluate the empirical likelihood ratio. Therefore, considering Eq. (1.5) and the following equation,

$$\int_t^\infty (1-F(u))du = \mu \int_t^\infty (1-\frac{t}{u})dG(u), \tag{2.1}$$

the empirical likelihood approach for calculating MRL can be based on the following equation as the special restriction,

$$E\left[\left(\frac{M(t)+t}{Y}-1\right) I_{(t,\infty)}(Y)\right] = 0. \tag{2.2}$$

Hence, the following estimation equation at a fixed time $0 \le t < \tau$ is proposed:

$$U(M(t)) := \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{t+M(t)}{Y_i}-1\Big)I(Y_i \ge t) = 0. \tag{2.3}$$

Apparently, the solution of this equation is $M_n(t)$.

Given Eq. (2.3), the EL ratio for the MRL function, $M(t)$, could be defined as follows. Suppose that $M_0(t)$ is the true value of $M(t)$ at time t. Using Eq. (2.3), it is easy to check that $E[U(M_0(t))] = 0$. Accordingly, let $\mathbf{p} = (p_1, \ldots, p_n)$ be a probability vector for which $\sum_{i=1}^{n} p_i = 1$ and $p_i \geq 0$ for each i. Define

$$V_i(t) := \Big(\frac{t + M_0(t)}{Y_i} - 1\Big) I(Y_i \geq t), \tag{2.4}$$

for $1 \leq i \leq n$ at a fix time t. Then, the evaluated EL at $M_0(t)$, is defined as

$$L(M_0(t)) := \sup\Big\{\prod_{i=1}^{n} p_i : \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i V_i(t) = 0\Big\}.$$

By using the Lagrange multiplier method, we have

$$p_i = \{n(1 + \alpha(t)V_i(t))\}^{-1}, \quad i = 1, \ldots, n,$$

where $\alpha(t)$ is the solution of

$$\frac{1}{n}\sum_{i=1}^{n} \frac{V_i(t)}{1 + \alpha(t)V_i(t)} = 0. \tag{2.5}$$

It should be mentioned that $\alpha(t)$ and $V_i(t)$ are functions of t. Therefore, they are considered and calculated at a fixed but arbitrary time like t_0, such that $0 \leq t_0 < \tau$. However, for the simplicity of notations we use $\alpha(t)$ and $V_i(t)$ instead of $\alpha(t_0)$ and $V_i(t_0)$.

Bear in mind that $\prod_{i=1}^{n} p_i$, subjecting to the condition $\sum_{i=1}^{n} p_i = 1$, attains its maximum which is n^{-n} at $p_i = n^{-1}$. Thus, the EL ratio for $M_0(t)$ can be defined as

$$R(M_0(t)) := \prod_{i=1}^{n} (np_i) = \prod_{i=1}^{n} \{1 + \alpha(t)V_i(t)\}^{-1}, \tag{2.6}$$

and the corresponding empirical log-likelihood ratio is defined as

$$l(M_0(t)) := -2 \log R(M_0(t)) = 2\sum_{i=1}^{n} \log\{1 + \alpha(t)V_i(t)\}, \tag{2.7}$$

where $\alpha(t)$ is the solution of (2.5).

3. Main results

In this section, the main results commence with inferencing EL-based pointwise confidence interval for MRL and will be continued by NA confidence interval and empirical likelihood ratio confidence band.

Theorem 1. *Assume that $E\left[\frac{1}{Y^2}\right] < \infty$. Then, for all $t \in [0, \tau)$, the limiting distribution of $l(M_0(t))$ is a chi-square distribution with 1 degree of freedom. That is,*

$$l(M_0(t)) \xrightarrow{\mathcal{D}} \chi^2_{(1)},$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution.

Proof. See the Proofs. □

Theorem 1 can be used to present the following confidence interval for $M(t)$ at a fixed time t, for $t \in [0, \tau)$. Therefore, an asymptotic $100(1 - \alpha)\%$ confidence interval for $M(t)$ can be obtained from the following equation:

$$C_1(t) = \left\{M(t) : l(M(t)) \leq \chi^2_{1,\alpha}\right\}, \tag{3.1}$$

where $\chi^2_{1,\alpha}$ is the upper α-quantile of the distribution of χ^2_1.

In order to calculate confidence intervals for the MRL function, one of the old fashioned ideas is to find confidence intervals straightaway without considering auxiliary information. By weak convergence results in Fakoor (2015), NA-based confidence intervals can be constructed for the MRL function.

Now, suppose that the condition $E\left[Y^{-r}\right] < \infty$, for some $r > 2$, is satisfied, then for each $t \in [0, \tau)$ we have,

$$\sqrt{n}\,(M_n(t) - M_0(t)) \xrightarrow{\mathcal{D}} N(0, \sigma^2(t)).$$

Although it is of theoretic interest to derive the asymptotic variance, $\sigma^2(t)$, explicitly, it takes a complicated form here and may not be easy to calculate directly. Hence, let $\sigma_n^2(t)$ be a consistent estimator of $\sigma^2(t)$. Then, an asymptotic $100(1-\alpha)\%$ confidence interval for $M(t)$ at time $t \in [0, \tau)$ is given by

$$C_2(t) = \left\{M(t) : |\sqrt{n}\,(M_n(t) - M(t))| \leq \sigma_n(t) z_{\alpha/2}\right\}, \tag{3.2}$$

where $z_{\alpha/2}$ is the upper $\frac{\alpha}{2}$-quantile of the normal distribution.

In order for establishing the weak convergence of the empirical likelihood ratio based stochastic process

$$\{l\,(M(t)) = -2\log R\,(M(t)) : t \in [0, a];\, a \leq \tau\},$$

it is essential to define $\xi(t)$ as follows.

$$\xi(t) := E\Big[\Big(\frac{t + M(t)}{Y_i} - 1\Big)^2 I(Y_i \geq t)\Big].$$

Theorem 2. *Assume that $E\left[\frac{1}{Y^2}\right] < \infty$. Then, there exists a mean zero Gaussian process $\{\zeta(t), 0 \leq t \leq a\}$ such that*

$$l(M(\cdot)) \xrightarrow{\mathcal{L}} \frac{\zeta^2(\cdot)}{\xi(\cdot)}, \tag{3.3}$$

in $D[0, a]$, the space of cadlag functions on $[0, a]$, and where $\xrightarrow{\mathcal{L}}$ denotes weakly convergence. Gaussian process $\zeta(\cdot)$ is given by

$$\zeta(t) = \int_t^\infty \left(\frac{M(t) + t}{u} - 1\right) dB(G(u)),$$

with the covariance function of

$$Cov(\zeta(t), \zeta(s)) = \int_{t\vee s}^\infty \left(\frac{M(t) + t}{u} - 1\right)\left(\frac{M(s) + s}{u} - 1\right) dG(u).$$

Proof. See the Proofs. □

Using Theorem 2 and the continuous mapping theorem, we obtain

$$\sup_{0\leq t\leq a}\{l(M(t))\} \xrightarrow{\mathcal{D}} \sup_{0\leq t\leq a}\frac{\zeta^2(t)}{\xi(t)}. \tag{3.4}$$

Consequently, for $t \in [0, a]$; $a \leq \tau$, an asymptotic $100(1 - \alpha)\%$ confidence band for the MRL function is

$$C_3 = \{M(t) : l\,(M(t)) \leq q_\alpha, t \in [0, a]\},$$

where q_α is the upper α-quantile of the distribution of

$$\sup_{0\leq t\leq a}\left\{\zeta^2(t)/\xi(t)\right\}. \tag{3.5}$$

Since it is difficult to evaluate the limiting distribution of (3.5) analytically, this Gaussian process may be approximated by means of the following method. Given Theorem 2 assumption, the approximation consists in the fact that for the non-random integrand $\left(\frac{M(t)+t}{u} - 1\right)$ and for every $t \in [0, a]$ we have

$$\int_t^\tau \left(\frac{M(t) + t}{u} - 1\right)^2 dG(u) < \infty.$$

Thus, the stochastic integral of $\zeta(t)$ remains a Gaussian process. Now, for a fixed t, $\zeta(t)$ as an Itô integral might be approximated through the following equation.

$$\widetilde{\zeta}(t) = \lim \sum_{i=0}^{n-1}\left(\frac{M(t) + t}{u_i{}^n} - 1\right)\left[B(G(u_{i+1}^n)) - B(G(u_i{}^n))\right], \tag{3.6}$$

in which for each n, u_i^n, when $i = 0, \ldots, n$, is a partition of $[t, \tau]$. Also, the lim is in quadratic mean, taken over all partitions with

$$\delta_n = \max_{1\leq i\leq n}(t_{i+1}^n - t_i^n) \longrightarrow 0,$$

as $n \longrightarrow \infty$.

Since the parameters $\xi(t)$, and also $M(t)$ in (3.5) and (3.6), respectively, are unknown, they are needed to be estimated via the corresponding empirical counterparts, which must be then replaced in the unknown quantities. For example, $M(t)$ and $\xi(t)$ may be estimated by their consistent estimators, i.e., $M_n(t)$ and

$$\xi_n(t) := \sum_{i=1}^{n}\Big[\Big(\frac{t+M_n(t)}{Y_i}-1\Big)^2 I(Y_i \geq t)\Big],$$

respectively. Now, we could obtain, for example N sample path of $\widetilde{\zeta}(t)$, denoted by $\{\widetilde{\zeta}_k(t);\, t\in[0.\tau]\}_{k=1}^{N}$ and, calculating

$$q_k = \max_{1\leq j\leq n} \frac{\widetilde{\zeta}_k^2(t_j)}{\xi_n(t_j)},$$

for a partition of the interval $[0,\tau]$. Ultimately, q_α can be approximated by the empirical percentile of $\{q_1,\ldots,q_N\}$.

4. Simulation study

In this section, simulation studies are conducted to inspect and illustrate the performance of the likelihood ratio confidence interval in comparison with the NA-based confidence interval under various sample sizes of length-biased data. Therefore, the performance of the proposed EL pointwise confidence interval, $C_1(t)$, has been compared with the introduced NA method, $C_2(t)$, in terms of coverage probability and lengths of confidence intervals simultaneously. Having presented these results, the uniform EL-based confidence band (C_3) will be illustrated.

To illustrate confidence intervals simulation results, in each iteration of simulation, we generated an i.i.d sample of $Y_1,\ldots,Y_n$ with the proposed length-biased distributions. The sample sizes of 50, 100 and 200 are considered, representing small, moderate and large sample sizes respectively. Accordingly, the confidence intervals for the MRL functions of the target populations based on the observations from the corresponding length-biased distributions have been assessed. The results and performances of confidence intervals have been calculated based on 5000 iterations for each scenario. Thus, the coverage probability (CP) have been evaluated as the proportion of the number of confidence intervals containing the real value of parameter out of 5000 simulated data sets. Moreover, as the other essential character of confidence interval, lengths of the confidence intervals (Δ) have been represented as well. Furthermore, all these scenarios have been evaluated for two separate nominal levels $(1-\alpha)$, 0.9 and 0.95.

We generated length-biased data following the above scenarios. Given the simulated data, the EL-based pointwise confidence intervals were calculated through the relation (3.1). Turning to the NA-based confidence interval, the resampling procedure can be used to estimate the variance of the stochastic process $\sqrt{n}(M_n(\cdot)-M_0(\cdot))$ in NA-based confidence interval (3.2) alternatively. We use bootstrap method to construct an estimator for $\sigma^2(t)$ as follows.

Let $Y_1^*,\ldots,Y_n^*$ be i.i.d random variables with the distribution function $G_n(\cdot)$, where $Y_1,\ldots,Y_n$ are fixed. Moreover, let $G_n^*(\cdot)$ be the empirical d.f. of the random sample $Y_1^*,\ldots,Y_n^*$. Then we define

$$M_n^*(x) := \frac{I_{[0,Y_{(n)}^*)}(x)}{1-F_n^*(x)}\int_x^\infty \big(1-F_n^*(t)\big)\,dt,$$

where

$$F_n^*(t) := \mu_n^*\int_0^t y^{-1}dG_n^*(y),$$

and

$$\mu_n^{*-1} := \int_0^\infty y^{-1}dG_n^*(y).$$

Defining $\theta^*(\cdot) := \sqrt{n}(M_n^*(\cdot)-M_n(\cdot))$, the above resampling procedure should be replicated B times, then the bootstrap estimate for $\sigma^2(t)$ would be the sample variance of $\theta_1^*(t),\ldots,\theta_B^*(t)$, that is

$$\nu_{boot}^2(t) := \frac{1}{B-1}\sum_{i=1}^{B}\left(\theta_i^*(t)-\frac{1}{B}\sum_{j=1}^{B}\theta_j^*(t)\right)^2.$$

It is worth mentioning that, in our resampling procedure, $B=500$ replication was considered for each iteration of the mentioned scenarios to estimate the asymptotic variance of the empirical process $\sqrt{n}(M_n(\cdot)-M_0(\cdot))$.

Table 1 summarizes the empirical 90% and 95% coverage probabilities of confidence intervals for MRL of Uniform(1,4) distribution based on length-biased data, which satisfies Theorem 1 condition. Generally, as expected, the lengths of EL-based confidence intervals have narrowed significantly by increasing in the sample sizes, and a similar trend is observed for the NA method. In comparison to the lengths of NA-based intervals, those of the EL ratio are almost narrower particularly for the small sample size.

Considering the CPs, the EL ratio confidence intervals roughly preserve the nominal level of significance. Moreover, although NA method tends to save the nominal level, the EL approach represents much better result in terms of CP despite

Table 1
90% and 95% confidence intervals for MRL of Uniform(1,4).

Time	Sample	$1-\alpha=90\%$				$1-\alpha=95\%$			
t	n	EL		NA		EL		NA	
		Δ	C.P.	Δ	C.P.	Δ	C.P.	Δ	C.P.
1.2	50	0.412	0.895	0.413	0.881	0.489	0.940	0.492	0.928
	100	0.295	0.897	0.295	0.894	0.351	0.947	0.352	0.942
	200	0.210	0.903	0.210	0.898	0.250	0.950	0.250	0.948
1.7	50	0.342	0.886	0.343	0.879	0.406	0.948	0.409	0.939
	100	0.244	0.906	0.245	0.900	0.290	0.949	0.291	0.943
	200	0.173	0.896	0.173	0.892	0.206	0.950	0.207	0.947
2.2	50	0.282	0.894	0.285	0.886	0.335	0.947	0.338	0.938
	100	0.201	0.908	0.202	0.905	0.240	0.945	0.241	0.940
	200	0.143	0.899	0.143	0.897	0.170	0.955	0.171	0.952
2.7	50	0.227	0.898	0.229	0.888	0.269	0.942	0.273	0.931
	100	0.162	0.894	0.163	0.888	0.193	0.946	0.194	0.939
	200	0.115	0.902	0.116	0.898	0.137	0.948	0.138	0.946
3.2	50	0.169	0.892	0.174	0.882	0.200	0.940	0.207	0.930
	100	0.121	0.897	0.123	0.889	0.144	0.947	0.146	0.940
	200	0.086	0.894	0.087	0.890	0.103	0.949	0.104	0.944
3.7	50	0.092	0.837	0.270	0.881	0.107	0.885	0.316	0.929
	100	0.070	0.887	0.075	0.879	0.083	0.934	0.090	0.925
	200	0.050	0.897	0.051	0.892	0.060	0.949	0.062	0.942

the fact that it has formed shorter intervals (see Remark 1). Accordingly, the length of intervals simultaneous with the CPs indicates the better efficiency of EL-based confidence intervals. It is worth mentioning that, despite fluctuations, there is no significant and meaningful differences between two groups of the nominal levels, 90% and 95%.

The following remark is presented to clarify why it is claimed that the EL confidence intervals have possessed much better results in comparison to the NA method. The remark is represented here, but it is satisfied for all simulation results given in this article.

Remark 1.

1. Compared with the EL method, constructing the introduced NA confidence intervals is much more time-consuming and very computationally intensive.
2. We have considered results for the α-quantiles when the amounts of αs range from 0.1 to 0.9, and we went even further in the presented figures. So, it was observed that, for small samples, the NA method do not present any result for large ts, unless the sample contained at least one observation larger than those ts. In other words, the NA method only represents the confidence interval for ts smaller than the maximum value of the sample. A similar situation happens for evaluating the bootstrap estimator. Thus, to earn $B = 500$ valid bootstrap samples to calculate bootstrap variance for the NA intervals, we were forced to generate data much more than 500 times for large αs and small samples (here for $n = 50$). Therefore, this issue leads the variance to surge illusively for large ts, increasing dramatically the length of interval, and therefore raising the coverage probabilities minimally. Also, it makes the NA method even worse in terms of time and computation. By contrast, the EL based confidence interval exhibits proper results for all ts separate from the values of the observations. It is worth mentioning that although there are many issues about Liang et al. (2016), some of which has been discussed previously, while Liang et al. (2016) declared that their EL and NA based intervals do not represent proper results for large ts, our proposed EL (and even NA method) preserves its accurate results even for ts larger than 90%-quantile.

Table 2 compares the performance of the proposed methods for MRL of Gamma(4,2) distribution based on length-biased observations, that satisfies Gamma(5,2) distribution and so the Theorem 1 condition. Moreover, the survey is conducted in two different categories of the nominal level of significance, 90% and 95%. In comparison with the EL method, it is revealed that the length of NA-based confidence intervals is always smaller by a moderate margin, except for ($N = 50$, $t = 3.5$). However, the coverage probabilities of the latter are noticeably lower than that of the former, which indicates the better performance of EL method (see Remark 1).

Figs. 1 and 2 reveal the average pointwise 95% confidence intervals for the MRL function of the respective target populations Uniform(1,4) and Gamma(4,2) through the EL ratio method. The diagrams consist in 5000 iterations for the large sample scenario, $N = 200$, of length-biased observations. As had been expected, the corresponding empirical 95% CPs, which are not shown due to space limitations, were so close to the nominal level for all ts.

Moreover, whereas the lengths of intervals in Fig. 2 have widened markedly over time, those of intervals in Fig. 1 seem to have decreased tremendously, which happens due to dramatic dive in the variance of $Y_i I(Y_i > t)$ over time in this case.

Table 2
90% and 95% confidence intervals for MRL of Gamma(4,2).

Time	Sample	$1-\alpha=90\%$				$1-\alpha=95\%$			
t	n	EL		NA		EL		NA	
		Δ	C.P.	Δ	C.P.	Δ	C.P.	Δ	C.P.
1.0	50	0.425	0.902	0.426	0.897	0.505	0.944	0.506	0.940
	100	0.301	0.900	0.301	0.895	0.359	0.951	0.359	0.946
	200	0.213	0.901	0.213	0.898	0.254	0.949	0.254	0.947
1.5	50	0.406	0.896	0.407	0.891	0.482	0.946	0.483	0.942
	100	0.287	0.900	0.288	0.896	0.342	0.948	0.343	0.947
	200	0.203	0.898	0.204	0.894	0.243	0.951	0.243	0.950
2.0	50	0.426	0.897	0.430	0.892	0.508	0.947	0.510	0.940
	100	0.300	0.898	0.301	0.895	0.360	0.948	0.361	0.944
	200	0.213	0.903	0.213	0.901	0.254	0.952	0.254	0.949
2.5	50	0.479	0.896	0.486	0.889	0.573	0.934	0.578	0.929
	100	0.339	0.902	0.341	0.895	0.406	0.950	0.406	0.945
	200	0.240	0.893	0.240	0.890	0.286	0.954	0.286	0.952
3.0	50	0.564	0.872	0.586	0.875	0.679	0.927	0.703	0.923
	100	0.404	0.886	0.408	0.882	0.481	0.944	0.483	0.934
	200	0.286	0.898	0.287	0.896	0.342	0.949	0.342	0.942
3.5	50	0.675	0.827	0.815	0.872	0.788	0.889	0.936	0.915
	100	0.500	0.873	0.515	0.873	0.600	0.923	0.613	0.918
	200	0.355	0.890	0.358	0.887	0.426	0.944	0.427	0.936

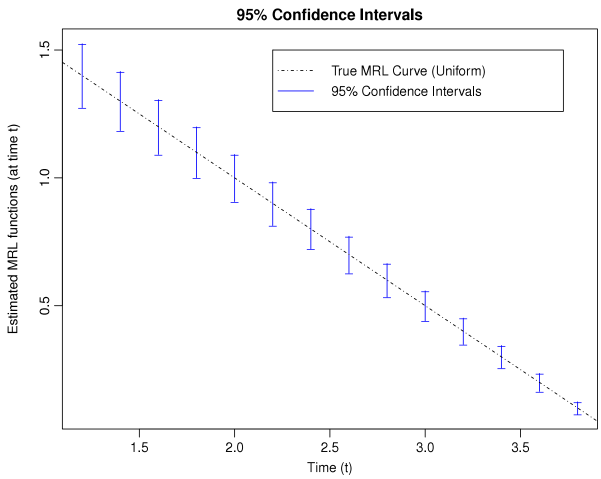

Fig. 1. Uniform MRL function with 95% confidence intervals.

According to the above diagrams and tables, when the condition $E\left[Y^{-r}\right]<\infty$ for some $r>2$ is held, which is necessary for the confidence interval (3.2) and it is stronger than our condition in (3.1) $\left(E\left[\frac{1}{Y^2}\right]<\infty\right)$, the confidence intervals represent appropriate results. Moreover, it seems that the EL method exhibits superiority over the NA approach. Now, an interesting question is that if the NA and EL methods are robust with respect to these assumptions.

In order to discern whether or not the confidence intervals are robust, we applied the confidence intervals for two target populations, namely Uniform $(0, 3)$ and Exponential(1) (Exp(1)). It is easy to check that the corresponding length-biased distributions do not satisfy the conditions of the confidence intervals (3.2) and (3.1).

Tables 3 and 4 compare simultaneously the empirical 90% and 95% CPs and the lengths of EL-based confidence intervals to those of NA-based intervals for the MRL function of Uniform(0,3) and Exp(1) distributions respectively. Broadly, the lengths of both EL and NA confidence intervals have narrowed substantially as the sample sizes climbed. Additionally, it can be observed that, in comparison to the NA method, the lengths of EL-based confidence intervals are mostly smaller by a margin.

Turning to the CP, although NA method tended to achieve the nominal level, the EL method represented much better results (see Remark 1). Accordingly, the lengths of intervals in parallel to the CPs indicate much better efficiency of EL method. Moreover, the vital issue is that these results are absolutely comparable with Tables 1 and 2. Therefore, the EL and NA approaches are both entirely robust with respect to the mentioned conditions.

Figs. 3 and 4 illustrate the average 95% pointwise confidence intervals via the EL ratio method for the MRL function based on length-biased observations of Uniform(0, 3) and Exp(1) distributions respectively. The survey was conducted based on

Table 3
90% and 95% confidence intervals for MRL of Uniform(0,3).

Time	Sample	$1-\alpha=90\%$				$1-\alpha=95\%$			
t	n	EL		NA		EL		NA	
		Δ	C.P.	Δ	C.P.	Δ	C.P.	Δ	C.P.
0.2	50	0.494	0.843	0.498	0.838	0.585	0.904	0.594	0.892
	100	0.373	0.870	0.371	0.859	0.443	0.932	0.442	0.922
	200	0.270	0.885	0.268	0.877	0.322	0.943	0.319	0.934
0.7	50	0.354	0.889	0.356	0.882	0.422	0.943	0.424	0.932
	100	0.255	0.890	0.255	0.882	0.303	0.949	0.303	0.940
	200	0.182	0.894	0.182	0.889	0.216	0.955	0.217	0.950
1.2	50	0.274	0.893	0.275	0.881	0.325	0.947	0.328	0.940
	100	0.196	0.902	0.196	0.897	0.233	0.950	0.234	0.945
	200	0.139	0.897	0.139	0.893	0.166	0.954	0.166	0.951
1.7	50	0.213	0.902	0.215	0.892	0.252	0.945	0.255	0.934
	100	0.152	0.898	0.152	0.893	0.181	0.951	0.182	0.948
	200	0.108	0.894	0.108	0.891	0.129	0.955	0.129	0.952
2.2	50	0.155	0.893	0.159	0.881	0.184	0.941	0.188	0.931
	100	0.111	0.892	0.112	0.886	0.132	0.945	0.133	0.938
	200	0.079	0.894	0.079	0.890	0.094	0.946	0.094	0.941
2.7	50	0.085	0.858	0.135	0.875	0.100	0.910	0.159	0.921
	100	0.064	0.896	0.066	0.886	0.076	0.945	0.079	0.936
	200	0.046	0.895	0.046	0.889	0.054	0.949	0.055	0.944

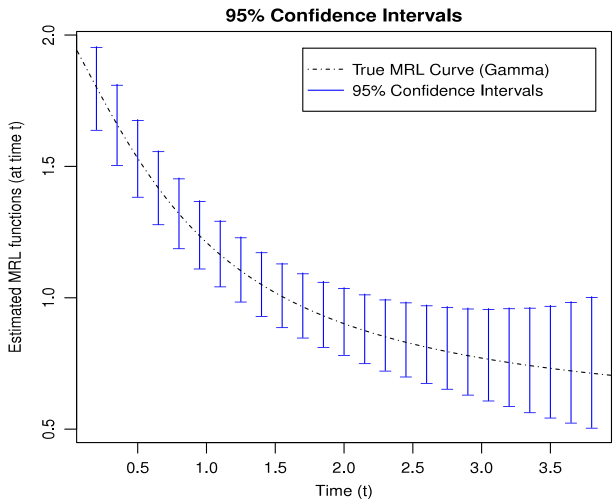

Fig. 2. Gamma MRL function with 95% confidence intervals.

5000 replications of the large sample scenario. Once again, the corresponding empirical 95% CPs, which are not shown due to space restrictions, were so close to the nominal value for all t. Additionally, it can be observed that the lengths of confidence intervals dipped swiftly as the time t climbed in Fig. 3 and the results are wholly comparable with Fig. 1. By contrast, in Fig. 4, despite an initial slight drop, the lengths of confidence intervals have risen dramatically over time.

We investigated the performance of the EL and NA confidence intervals (and EL band) when the condition $E\left[Y^{-r}\right]<\infty$ for some $r>2$ is violated, generating data from length-biased distributions of Uniform$(0,3)$ and Exponential(1) target populations. By comparison, it is revealed that the results in this occasion are absolutely comparable to the previous simulation results when the conditions were held. Therefore, although the validity of the confidence intervals is not analytically verified under more general sitting, given the results that are listed in Tables 3, 4, Figs. 3 and 4, it is interesting to declare that the proposed EL-based and the introduced NA-based confidence intervals are both quite robust. Furthermore, we compared the empirical coverage probabilities and the lengths of confidence intervals for the proposed EL method with those for the NA method in this occasion too. Considering Remark 1, it is concluded that the EL based confidence intervals exhibit infinitely preferable behavior.

Fig. 5 illustrates the empirical likelihood ratio 95% confidence band for the MRL function of Gamma(4,1) distribution. The diagram consists in a sample of size 400 generated from the corresponding length-biased distribution. This simulation shows that the proposed band is appropriate for practical use. Also, it is worth mentioning that we checked the confidence

Table 4
90% and 95% confidence intervals for MRL of Gamma(1,1).

Time	Sample	$1-\alpha=90\%$				$1-\alpha=95\%$			
t	n	EL		NA		EL		NA	
		Δ	C.P.	Δ	C.P.	Δ	C.P.	Δ	C.P.
	50	0.485	0.878	0.497	0.880	0.574	0.936	0.590	0.935
0.2	100	0.350	0.898	0.354	0.896	0.415	0.945	0.419	0.940
	200	0.249	0.894	0.250	0.891	0.297	0.950	0.299	0.947
	50	0.454	0.890	0.461	0.889	0.542	0.948	0.548	0.944
0.7	100	0.321	0.891	0.324	0.892	0.384	0.949	0.386	0.947
	200	0.227	0.900	0.228	0.899	0.271	0.948	0.272	0.945
	50	0.498	0.893	0.505	0.890	0.598	0.947	0.604	0.941
1.2	100	0.353	0.896	0.356	0.894	0.422	0.953	0.424	0.948
	200	0.249	0.899	0.250	0.899	0.344	0.954	0.345	0.952
	50	0.574	0.889	0.584	0.887	0.691	0.941	0.700	0.934
1.7	100	0.408	0.897	0.411	0.895	0.487	0.950	0.490	0.948
	200	0.288	0.906	0.288	0.905	0.344	0.954	0.345	0.952
	50	0.687	0.882	0.708	0.883	0.824	0.937	0.844	0.932
2.2	100	0.484	0.890	0.490	0.890	0.579	0.949	0.583	0.946
	200	0.341	0.901	0.343	0.899	0.408	0.947	0.408	0.944
	50	0.821	0.874	0.875	0.882	0.988	0.928	1.044	0.930
2.7	100	0.584	0.892	0.594	0.890	0.698	0.944	0.706	0.936
	200	0.410	0.894	0.413	0.898	0.489	0.951	0.491	0.945

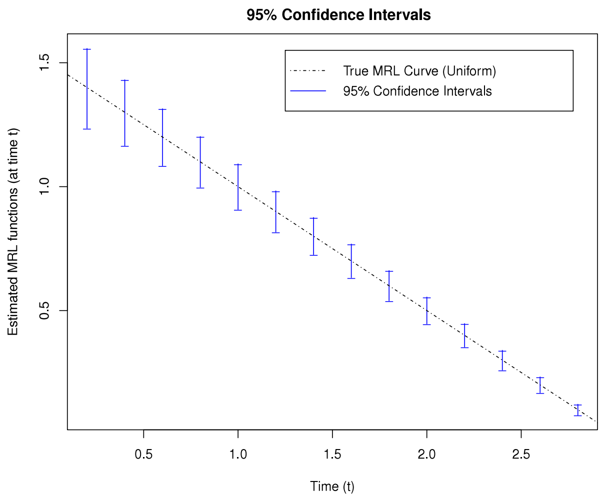

Fig. 3. Uniform MRL function with 95% confidence intervals.

band for some other distribution for which Theorem 2 condition ($E\left(\frac{1}{Y^2}\right) < \infty$) was violated, so the band seems to be robust with respect to this condition. However, we have not represented these results here.

5. Real data application

We shall apply the proposed EL-based confidence interval for MRL in a real application for better illustration. A setup similar to that is introduced in Introduction and Preliminaries arises in a study on the lifetime of automobile brake pads (Kalbfleisch and Lawless, 1992). In a real data given by Lawless (2002), the number of units, in kilometers or miles, driven until the brake pads were reduced to a specified minimum thickness was considered as the lifetime and therefore a nominal lifetime is assumed for the brake pads. To study the lifetime distribution, a manufacturer selected a random sample of vehicles that had been sold over proceeding 12 months at a specific group of dealers. It is worth mentioning that only cars that still had had the initial brake pads were selected. For each car the brake pad lifetime t_i could have then been observed by following the cars prospectively. Alternatively, to save time the current odometer reading u_i (in km) and the remaining pad thickness above the minimum were used in conjunction with the initial pad thickness to estimate the lifetime t_i (in km); this was treated as the actual lifetime in the analysis. Consequently, in any case the lifetime t_i is left truncated at u_i.

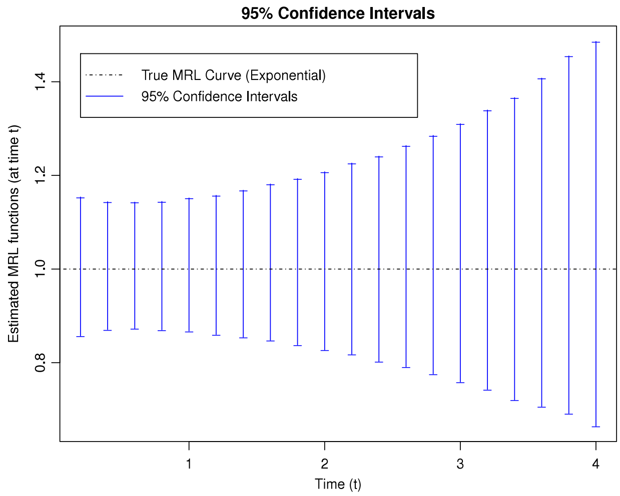

Fig. 4. Exponential MRL function with 95% confidence intervals.

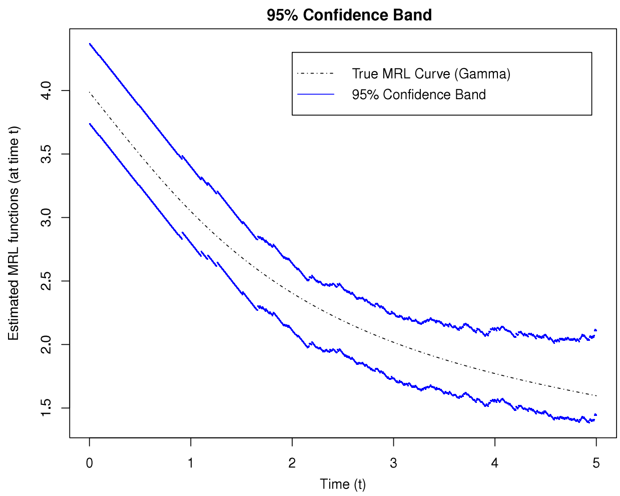

Fig. 5. Gamma (4,1) MRL function with 95% confidence band.

The brake pads data, given in Lawless (2002), include the brake pads lifetimes of 98 individual cars, ranging from 18 600 to 165 500 km. It can be deduced that the pads which possess longer lifetimes own greater chance of being observed in the selected sample. This bias happens owing to non-random sampling of components and therefore the presence of left-truncation variable. Thus, it can be derived that the distribution of the observed brake pads is Eq. (1.1), in which $F(\cdot)$ denotes the distribution of all brake pads lifetimes (observed and unobserved components). So, if the truncation time (u_i) and the remaining pad thickness ($t_i - u_i$) are identically distributed (see Cox, 1969) or equivalently a stationary Poisson process, the so-called stationarity assumption, satisfies, the lifetime of observed subjects own a special case of distribution (1.1) that is length-biased distribution (1.3) (see Asgharian et al. (2004)). Although we do not represent the proof here, we checked this condition and it could be verified that the observed lifetimes are length-biased.

Accordingly, we have applied the proposed EL ratio confidence intervals for the brake pads data to estimate the MRL function. So, we chose the empirical lower 10%, 20%, . . . , 90%-quantile, as the propped time to represent the confidence intervals in Table 5. Accordingly, it is revealed that the lower bounds of the confidence intervals have dipped substantially by increasing in the time of study. However, the upper bounds of the confidence intervals have decreased significantly before climbing swiftly back to its initial level.

For better illustration, Fig. 6 plots the pointwise 95% EL confidence intervals for the MRL function based on brake pads data. For this purpose, confidence intervals for each 0.5 unit (500 km) were calculated and have been plotted in the diagram. Accordingly, the confidence intervals have estimated that MRL of brake pads should drop steeply until roughly 69 units (about the lower 70%-quantile) when it experiences an unexpected surge. After that, despite violent fluctuations, it is

Table 5
95% confidence intervals for MRL function.

Time: (in 1000-km units)	38.77	44.96	50.87	56.20	65.05
Upper	30.614	29.266	28.419	28.443	24.593
Lower	22.263	20.789	19.803	19.642	15.370
Time: (in 1000-km units)	70.12	78.64	87.24	102.41	
Upper	25.743	23.940	25.949	29.013	
Lower	15.598	12.300	11.467	08.620	

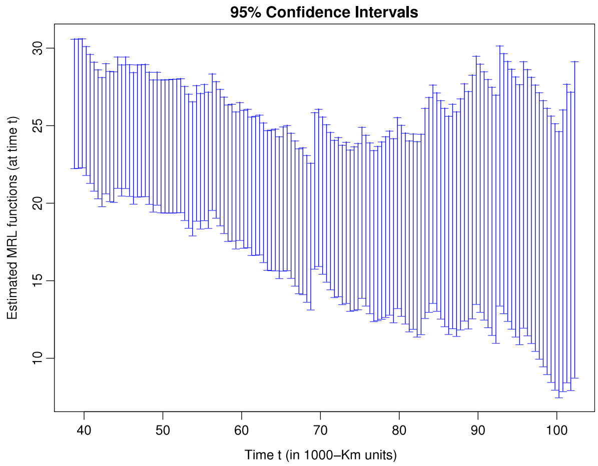

Fig. 6. 95% EL-based confidence intervals for MRL function.

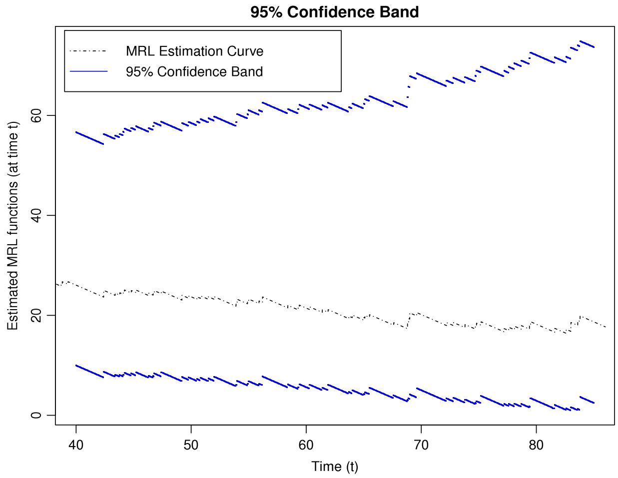

Fig. 7. 95% EL-based confidence band for MRL function.

deduced that MRL ought to decline gradually. Moreover, it seems broadly that the lengths of intervals have surged during the period of survey.

To illustrate how the confidence band C_3 performs, Fig. 7 reveals the empirical likelihood ratio 95% confidence band for the MRL function of the automobile brake pads data. Moreover, the MRL function is evaluated and plotted using the consistent estimator (1.6). According to this diagram, it is estimated that MRL of the brake pads data should decline steadily and considerably over the period of the survey. Also, as the age of this kind of brake pads increases, more fluctuation for MRL is estimated.

6. Proofs

We need the following lemma in order to prove Theorem 1.

Lemma 1. *Under the same condition as in Theorem 1, for all $t \in [0, \tau)$, we have*

$$E[V_1^2(t)] < \infty,$$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} V_i(t) \xrightarrow{\mathcal{D}} N(0, \xi(t)), \tag{6.1}$$

and

$$\frac{1}{n} \sum_{i=1}^{n} V_i^2(t) \xrightarrow{\mathcal{P}} \xi(t) \tag{6.2}$$

where $\xrightarrow{\mathcal{P}}$ denotes convergence in probability, $V_i(t)$ is given in (2.4), and

$$\xi(t) = E\Big[\Big(\frac{t + M_0(t)}{Y_i} - 1\Big)^2 I(Y_i \geq t)\Big] < \infty.$$

Proof. Let $t \in [0, \tau)$ be a fix point. Under assumption of the lemma, we have

$$\begin{aligned} E[V_i^2(t)] &= E\left[\left(\frac{t + M_0(t)}{Y_i} - 1\right) I(Y_i \geq t)\right]^2 \\ &\leq \left(E\left[\frac{t + M_0(t)}{Y_i} I(Y_i \geq t)\right]^2 + E[I(Y_i \geq t)]^2\right) \\ &\leq \left((t + M_0(t))^2 E\left[\frac{1}{Y_i}\right]^2 + 1\right) \\ &< \infty. \end{aligned} \tag{6.3}$$

Since

$$E\Big[\Big(\frac{t + M_0(t)}{Y} - 1\Big) I(Y \geq t)\Big] = 0,$$

given the central limit theorem for i.i.d. random variables $V_1(t), V_2(t), \ldots, V_n(t)$, we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} V_i(t) \xrightarrow{\mathcal{D}} N(0, \xi(t)),$$

so (6.1) is obtained. Turning to the second part of the lemma, using the law of large numbers, it follows that

$$\frac{1}{n} \sum_{i=1}^{n} V_i^2(t) \xrightarrow{\mathcal{P}} E\left[V_1^2(t)\right].$$

Thus, (6.2) is proved. □

The proof of Theorem 1 is mainly inspired by Zhao and Qin (2006).

Proof of Theorem 1. Let $t \in [0, \tau)$. According to Lemma 1, since $E(V_1^2(t)) < \infty$, by using Lemma 3 of Owen (1990), we have

$$\max_{1 \leq i \leq n} |V_i(t)| = o_p(n^{1/2}), \tag{6.4}$$

and also,

$$\frac{1}{n} \sum_{i=1}^{n} \left|V_i(t)\right|^3 = o_p(n^{1/2}). \tag{6.5}$$

Thus, considering (6.4) and (6.5) and applying the same arguments used in Owen (1991), it can be proven that

$$\left|\alpha(t)\right| = O_p(n^{-1/2}). \tag{6.6}$$

Using the Taylor expansion for (2.7), it can be shown that

$$\begin{aligned} l(M(t)) &= 2 \sum_{i=1}^{n} \log\{1 + \alpha(t) V_i(t)\} \\ &= 2 \sum_{i=1}^{n} \Big(\alpha(t) V_i(t) - \frac{(\alpha(t) V_i(t))^2}{2}\Big) + R_n(t), \end{aligned} \tag{6.7}$$

where by applying Eqs. (6.5) and (6.6), it can be seen that

$$\begin{aligned}|R_n(t)| &\le C\sum_{i=1}^{n}|\alpha(t)V_i(t)|^3\\ &\le C|\alpha(t)|^3\sum_{i=1}^{n}|V_i(t)|^3\\ &= o_p(1).\end{aligned} \tag{6.8}$$

Turning to the first part of (6.7) , recalling Eq. (2.5), it follows that

$$\begin{aligned}0 &= \sum_{i=1}^{n}\frac{V_i(t)}{1+\alpha(t)V_i(t)}\\ &= \sum_{i=1}^{n}V_i(t)\Big[1-\alpha(t)V_i(t)+\frac{(\alpha(t)V_i(t))^2}{1+\alpha(t)V_i(t)}\Big]\\ &= \sum_{i=1}^{n}V_i(t)-\Big(\sum_{i=1}^{n}V_i^2(t)\Big)\alpha(t)+\sum_{i=1}^{n}\frac{V_i(t)(\alpha(t)V_i(t))^2}{1+\alpha(t)V_i(t)}.\end{aligned} \tag{6.9}$$

Now, having the (6.4) and (6.6) in the one hand and applying Lemma 1 on the other hand, it follows from Eq. (6.9) that

$$\alpha(t) = \Big(\sum_{i=1}^{n}V_i(t)^2\Big)^{-1}\sum_{i=1}^{n}V_i(t)+o_p(1). \tag{6.10}$$

Once more by recollecting (2.5),

$$\begin{aligned}0 &= \sum_{i=1}^{n}\frac{\alpha(t)V_i(t)}{1+\alpha(t)V_i(t)}\\ &= \sum_{i=1}^{n}(\alpha(t)V_i(t))-\sum_{i=1}^{n}(\alpha(t)V_i(t))^2+\sum_{i=1}^{n}\frac{(\alpha(t)V_i(t))^3}{1+\alpha(t)V_i(t)}.\end{aligned}$$

Additionally, having (6.4) and (6.6), we have

$$\sum_{i=1}^{n}\frac{(\alpha(t)V_i(t))^3}{1+\alpha(t)V_i(t)} = o_p(n^{-1/2}).$$

Therefore, given the former and the latter equations, it has been apparent that

$$\sum_{i=1}^{n}(\alpha(t)V_i(t))^2 = \sum_{i=1}^{n}\alpha(t)V_i(t)+o_p(1).$$

Eventually, it follows from Eqs. (6.7), (6.10) and Lemma 1 that

$$\begin{aligned}l(M_0(t)) &= \sum_{i=1}^{n}(\alpha(t)V_i(t))+o_p(1)\\ &= \frac{\left(\sum_{i=1}^{n}V_i(t)\right)^2}{\sum_{i=1}^{n}V_i(t)^2}+o_p(1)\\ &\xrightarrow{\mathcal{D}} \chi^2_{(1)}.\end{aligned}$$

The following two lemmas are needed to prove Theorem 2.

Lemma 2. *Assume that $E[\frac{1}{Y^2}]<\infty$, then we have*

$$\frac{1}{n}\sum_{i=1}^{n}V_i^2(\cdot)\xrightarrow{\mathcal{P}}\xi(\cdot), \tag{6.11}$$

uniformly over $t\in[0,a]$, where $a<\tau$.

The following remark is needed for proving Lemma 2.

Remark 2 (*Theorem 3 of Van der Vaart and Wellner (2000)*)**.** Let $G(\cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a continuous function. Moreover, suppose that $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ are P-Glivenko–Cantelli classes of functions. Then the class of functions $\mathcal{S} = G(\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n)$ is also P-Glivenko–Cantelli and it has an integrable envelope function. The class $\mathcal{S}$ is the collection of all functions $S(x)$ which are in the form of $S(x) = G(F_1(x), F_2(x), \ldots, F_n(x))$ where F_i is in $\mathcal{F}_i$.

Proof of Lemma 2. Let $\mathcal{F}_1$ be the class $\{I(y \geq t); t > 0\}$. Using the usual Glivenko–Cantelli theorem, it is apparent that $\mathcal{F}_1$ is a P-Glivenko–Cantelli class. Additionally, suppose $\mathcal{F}_2$ and $\mathcal{F}_3$ be the same class that is

$$\left\{\left(\frac{t + M_0(t)}{y}\right) - 1; t \in [0, a]\right\}.$$

Hence, through the usage of the strong law of large numbers it seems that $\mathcal{F}_2, \mathcal{F}_3$ are also P-Glivenko–Cantelli classes.

Now, let $G(x, y, z) = xyz$, which is obviously continuous. Recollecting the assumption of theorem and considering the fact that for any arbitrary t such that $t \in [0, a)$ there exist $C > 0$ such that

$$\left(\frac{t + M_0(t)}{y}\right) - 1 < \left(\frac{C}{y}\right),$$

the class of functions

$$\mathcal{S} : \left\{\left(\frac{t + M_0(t)}{y} - 1\right)^2 I(y \geq t); t \in [0, a]\right\}$$

definitely has an integrable envelope function (C/Y). Thus, considering the result in the recent argument,

$$\sup_{t\in[0,a]} \left|\frac{1}{n}\sum_{i=1}^{n} V_i^2(t) - E(V_i^2(t))\right| = o_p(1),$$

uniformly over $[0, a]$. Therefore, the proof of Lemma 2 is completed. □

Lemma 3. *Let $E[\frac{1}{Y^2}] < \infty$. Then $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_i(\cdot)$ converges weekly to a mean zero Gaussian process $\zeta(\cdot)$ in Skorokhod-space $D[0, a]$, where*

$$\zeta(t) := \int_t^\infty \left(\frac{M(t) + t}{u} - 1\right) dB\,(G(u))\,,$$

and

$$Cov(\zeta(t), \zeta(s)) = \int_{t\vee s}^\infty \left(\frac{M(t) + t}{u} - 1\right)\left(\frac{M(s) + s}{u} - 1\right) dG(u).$$

Here, $B(\cdot)$ is a Brownian Bridge on the unit interval.

Proof of Lemma 3. Considering Eq. (2.2), it follows that

$$\begin{aligned} M_n(t) - M(t) &= -\mu\frac{I_{[0,\tau)}(t)}{1 - F(t)}\left[\frac{1}{n}\sum_{i=1}^{n} V_i(t) - \int_t^\infty \frac{M(t)}{u} d\left(G_n(u) - G(u)\right)\right] \\ &\quad + \left(\mu_n\frac{I_{[0,Y_{(n)})}(t)}{1 - F_n(t)} - \mu\frac{I_{[0,\tau)}(t)}{1 - F(t)}\right)\int_t^\infty (1 - \frac{t}{u})dG_n(u)dt. \end{aligned}$$

On the other hand, turning the left side of above equation, $M_n(t) - M(t)$, straightaway

$$\begin{aligned} M_n(t) - M(t) &= \mu_n\frac{I_{[0,Y_{(n)})}(t)}{1 - F_n(t)}\left[\int_t^\infty (1 - \frac{t}{u})d\left(G_n(u) - dG(u)\right)\right] \\ &\quad + \left(\mu_n\frac{I_{[0,Y_{(n)})}(t)}{1 - F_n(t)} - \mu\frac{I_{[0,\tau)}(t)}{1 - F(t)}\right)\int_t^\infty (1 - \frac{t}{u})dG(u). \end{aligned}$$

Given this and the former equations, it can be observed that

$$\begin{aligned} \frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_i(t) &= \int_t^\infty \frac{M(t)}{u}\alpha_n(du) \\ &\quad - \mu_n\frac{I_{[0,Y_{(n)})}(t)}{1 - F_n(t)}\left(\mu\frac{I_{[0,\tau)}(t)}{1 - F(t)}\right)^{-1}\left[\int_t^\infty (1 - \frac{t}{u})\alpha_n(du)\right] \\ &\quad + \left(\mu\frac{I_{[0,\tau)}(t)}{1 - F(t)}\right)^{-1}\left(\mu_n\frac{I_{[0,Y_{(n)})}(t)}{1 - F_n(t)} - \mu\frac{I_{[0,\tau)}(t)}{1 - F(t)}\right) \times \int_t^\infty (1 - \frac{t}{u})\alpha_n(du), \end{aligned} \tag{6.12}$$

where $\alpha_n(u) = \sqrt{n}(G_n(u) - G(u))$. Now by classical invariance principle of Donsker (1952), which asserts that empirical process $\alpha_n(u)$, $u \geq 0$ in the Skorokhod space $D[0, \infty]$, weakly converges to $B(G(\cdot))$, and using the below convergence

$$\mu_n \frac{I_{[0,Y_{(n)})}(t)}{1 - F_n(t)} \xrightarrow{\mathcal{P}} \mu \frac{I_{[0,\tau)}(t)}{1 - F(t)},$$

the desired result follows from Eq. (6.12) and the continuous mapping theorem. □

Proof of Theorem 2. Since $E[\frac{1}{Y^2}] < \infty$, thus by the same argument as in the proof of Lemma 3 of Owen (1990), it can be observed that

$$\max_{1 \leq i \leq n} \left| \frac{1}{Y_i} \right| = o_p(n^{1/2}), \tag{6.13}$$

and

$$\frac{1}{n} \sum_{i=1}^{n} \left| \frac{1}{Y_i} \right|^3 = o_p(n^{1/2}). \tag{6.14}$$

By using (6.13) it can be seen that uniformly over $t \in [0, a]$,

$$\begin{aligned} \max_{1 \leq i \leq n} |V_i(t)| &= \max_{1 \leq i \leq n} \left| \Big(\frac{t + M_0(t)}{Y_i} - 1 \Big) I(Y_i \geq t) \right| \\ &\leq C \max_{1 \leq i \leq n} \left| \frac{1}{Y_i} \right| \\ &= o_p(n^{1/2}). \end{aligned} \tag{6.15}$$

And also,

$$\frac{1}{n} \sum_{i=1}^{n} \left| V_i(t) \right|^3 = o_p(n^{1/2}). \tag{6.16}$$

Given (6.15) and (6.16), and similar to the proof of Theorem 1, we have

$$\begin{aligned} l(M(t)) &= \sum_{i=1}^{n} (\alpha(t) V_i(t)) + o_p(1) \\ &= \frac{\left(\sum_{i=1}^{n} V_i(t) \right)^2}{\sum_{i=1}^{n} V_i^2(t)} + o_p(1), \end{aligned}$$

uniformly over $t \in [0, a]$. Therefore, by using Lemmas 2 and 3, we obtain (3.3). □

Acknowledgment

The authors would like to sincerely thank the anonymous referees for their valuable comments on the previous draft, which resulted in the present version of this paper. This research was supported by a grant from Ferdowsi University of Mashhad (No. 38716).

References

Asgharian, M., Wolfson, D.B., Zhang, X., 2004. A simple criterion for the stationarity of the incidence rate from prevalent cohort studies. Dept. of Mathematics and Statistics, McGill University, Technical report 2004-01.
Cox, D.R., 1969. In: Johnson, N.L., Smith, H. (Eds.), Some Sampling Problems in Technology. New Developments in Survey Sampling. Wiley, New York, pp. 506–527.
Cristóbal, J.A., Alcalá, J.T., 2001. An overview of nonparametric contributions to the problem of functional estimation from biased data, 10(2), 309–332.
Csörgő, M., Zitikis, R., 1996. Mean residual life processes. Ann. Statist. 24, 1717–1739.
DiCiccio, T., Hall, P., Romano, J., 1991. Empirical likelihood is Bartlett-Correctable. Ann. Statist. 19 (2), 1053–1061.
Donsker, M.D., 1952. Justification and extension of Doob's heuristic approach to the Kolmogorov- Smirnov theorems. Ann. Math. Statist. 23 (2), 277–281.
Fakoor, V., 2015. On the nonparametric mean residual life estimator in length-biased sampling. Comm. Statist. Theory Methods 44, 512–519.
Feller, W., 1966. An Introduction To Probability Theory and Its Applications, Vol. 2. Wiley, New York.
Guess, F., Proschan, F., 1988. Mean residual life, theory and applications. In: Handbook of Statistics: Relaiability and Quality Control, Vol. 7. pp. 215–224.
Hall, W.J., Wellner, J.A., 1979. Estimation of mean residual life. Tech. rep., University of Rochester.
Kalbfleisch, J.D., Lawless, J.F., 1992. Some useful methods for truncated data. J. Qual. Technol. 24–152.
Kumazawa, Y., 1987. A note on an estimator of life expectancy with random censorship. Biometrika 74, 655–658.
Lawless, J.F., 2002. Statistical Models and Methods for Lifetime Data, second ed. In: (Wiley Series in Probability and Statistics)-Wiley-Interscience.
Liang, W., Shen, J.S., He, S.Y., 2016. Likelihood ratio inference for mean residual life of length-biased random variable. Acta Math. Appl. Sin. Engl. Ser. 32 (2), 269–282.
Ning, J., Qin, J., Asgharian, M., Shen, Y., 2013. Empirical likelihood-based confidence intervals for length biased data. Stat. Med. 32, 2278–2291.

Owen, A.B., 1988. Empirical likelihood ratio confidence intervals for single functional. Biometrika 75, 249–273.
Owen, A.B., 1990. Empirical likelihood ratio confidence regions. Ann. Statist. 18, 90–120.
Owen, A.B., 1991. Empirical likelihood for linear models. Ann. Statist. 19, 1725–1747.
Patil, G.P., Rao, C.R., 1978. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. Biometrics 34, 179–184.
Proschan, F., Serfling, R.J., 1974. Reliability and biometry. In: Statistical Analysis of Life Length. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.
Qin, J., 1993. Empirical likelihood in biased sample problems. Ann. Statist. 21, 1182–1196.
Qin, G.S., Zhao, Y., 2007. Empirical likelihood inference for the mean residual life under random censorship. Statist. Probab. Lett. 77, 549–557.
Van der Vaart, A.W., Wellner, J., 2000. Weak Convergence and Empirical Processes: With Applications To Statistics. In: Springer Series in Statistics, Springer.
Wicksell, S.D., 1925. The corpuscle problem. Biometrika 17, 84–99.
Wu, H., Luan, Y., 2014. An efficient estimation of the mean residual life function with length-biased right-censored data. Math. Probl. Eng. 1–5.
Yang, G., 1977. Lifetime expectancy under random censorship. Stochastic Process. Appl. 6, 33–39.
Yang, G.L., 1978. Estimation of a biometric function. Ann. Statist. 6, 112–116.
Zhao, M., Jiang, H., Liu, X., 2013. A note on estimation of the mean residual life function with left-truncated and right-censored data. Statist. Probab. Lett. 83, 2332–2336.
Zhao, Y., Qin, G.S., 2006. Inference for the mean residual life function via empirical likelihood. Commun. Statist. 35, 1025–1036.
Zhou, M., Jeong, J.H., 2011. Empirical likelihood ratio test for median and mean residual lifetime. Stat. Med. 30, 152–159.