Minimum-Distance Parametric Estimation Under Progressive Type-I Censoring

ARTICLE *in* IEEE TRANSACTIONS ON RELIABILITY · JULY 2010

Impact Factor: 1.93 · DOI: 10.1109/TR.2010.2044615 · Source: IEEE Xplore

3 AUTHORS, INCLUDING:

Narayanaswamy Balakrishnan
McMaster University
498 PUBLICATIONS **7,881** CITATIONS
SEE PROFILE

Xuejing Zhao
Lanzhou University
8 PUBLICATIONS **40** CITATIONS
SEE PROFILE

Minimum-Distance Parametric Estimation Under Progressive Type-I Censoring

Narayanaswamy Balakrishnan, *Member, IEEE*, Laurent Bordes, and Xuejing Zhao

***Abstract*—The objective of this paper is to provide a new estimation method for parametric models under progressive Type-I censoring. First, we propose a Kaplan-Meier nonparametric estimator of the reliability function taken at the censoring times. It is based on the observable number of failures, and the number of censored units occurring from the progressive censoring scheme at the censoring times. This estimator is then shown to asymptotically follow a normal distribution. Next, we propose a minimum-distance method to estimate the unknown Euclidean parameter of a given parametric model. This method leads to consistent, asymptotically normal estimators. The maximum likelihood estimation method based on group-censored samples is discussed next, and the efficiencies of these two methods are compared numerically. Then, based on the established results, we derive a method to obtain the optimal Type-I progressive censoring scheme, Finally we illustrate all these results through a Monte Carlo simulation study, and an illustrative example.**

***Index Terms*—Asymptotic distribution, Kaplan-Meier estimator, martingale, maximum likelihood estimator, minimum-distance estimator, minimum variance linear estimator, Nelson-Aalen estimator, optimal progressive censoring scheme, progressive Type-I censoring scheme.**

ACRONYMS

MLE	maximum likelihood estimate(or)
MDE	minimum distance estimate(or)
WSBE	weighted sum of the best estimator
OWSBE	sum of the best estimator with optimal weights
CDF	cumulative distribution function
CHF	cumulative hazard function

NOTATION

$\mathbb{R}^d$	real d-dimensional Euclidean space
n	total number of units placed on test
m	number of censored stage
Θ	parameter set
θ	unknown Euclidean parameter
T_i	prefixed censoring times, $1 \leq i \leq m$
N_i	number of units failed in $(T_{i-1}, T_i]$
R_i	number of units censored at time T_i
α_i	proportion of the censored units, $1 \leq i \leq m-1$
α_i^+, α_i^-	number of items at risk just after, and before T_i
F_i, $\overline{F}_i$	CDF, and reliability function at T_i
F, $\bar{F}$	$F = (F_1, \ldots, F_m)$ and $\bar{F} = (\bar{F}_1, \ldots, \bar{F}_m)$
λ_i, Λ_i	hazard rate, and CHF values of discrete distribution
λ, Λ	$\lambda = (\lambda_1, \ldots, \lambda_m)$ and $\Lambda = (\Lambda_1, \ldots, \Lambda_m)$
$\mathbb{F}$	filtration
$L(.)$	likelihood function
$I(.)$	Fisher's information matrix
$\xrightarrow{P}$	converge in probability
$\rightsquigarrow$	converge in distribution

Manuscript received September 21, 2008; revised July 17, 2009 and August 27, 2009; accepted November 30, 2009. First published April 26, 2010; current version published June 03, 2010. Xuejing Zhao is supported by the Fundamental Research Funds for the Central Universities, lzujbky-2009-120. Associate Editor: R. H. Yeh.

N. Balakrishnan is with the Department of Mathematics and Statistics, McMaster University, Canada (e-mail: bala@univmail.cis.mcmaster.ca).

L. Bordes is with the Department of Mathematics, Université de Pau et des Pays de l'Adour, UMR CNRS 5142, France (e-mail: laurent.bordes@univ-pau.fr).

X. Zhao is with the School of Mathematics and Statistics, Lanzhou University, China (e-mail: mathxjzhao@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TR.2010.2044615

I. INTRODUCTION

IN PRACTICAL life-testing experiments, one often encounters incomplete data (such as censored data, and truncated data) for which many inferential methods have been developed; see, for example, [1]–[4] for elaborate discussions in this direction. When it is necessary to reduce the cost and/or the duration of a life-testing experiment, one may choose to terminate the experiment early, which results in the so-called censored sampling plan, or censored sampling scheme. Many types of censoring have been discussed in the literature, with the most common censoring schemes being Type-I right censoring, and Type-II right censoring. Generalizations of these censoring schemes to progressive Type-I, and Type-II right censoring have also been discussed [5]. Progressively censored samples are observed when, at various stages of an experiment, some of the surviving units are removed from further observation. The remaining units are then continued on test under observation, either until failure, or until a subsequent stage of censoring. Progressive censoring schemes have been found to be useful in reliability analysis, product testing, and animal carcinogenicity experiments.

Considerable attention has been paid in recent years to parametric, semi-parametric, and nonparametric estimation under progressive Type-II censoring [6]–[10]; whereas for progressive Type-I censoring, relatively little work has been done.

From a non-parametric estimation viewpoint, [11] has studied the asymptotic behavior of the estimator of the reliability function under two types of progressive Type-I censoring using both martingale, and empirical processes theory. References [12]–[14] discussed the problem of estimation, and asymptotics for progressively Type-I right censored step-stress experiments, under an exponential cumulative exposure model. References [15], and [16] similarly discussed the same step-stress problem in the case of Type-I, and Type-II hybrid censored samples. Reference [17] considered some problems relating to the maximum likelihood estimation for the exponential distribution under progressive Type-I censoring, and changing failure rates.

In this paper, we consider a progressively Type-I censored sample defined by the progressive censoring scheme $R_1,\ldots,R_{m-1}$ (or $\alpha_1,\ldots,\alpha_{m-1}$ in [0, 1); see [11]), and pre-fixed censoring times $T_1,\ldots,T_m$. Suppose n s-independent units are placed simultaneously on a life-test at time 0. Let N_i $(i = 1,\ldots,m)$ denote the number of observed failures in the time interval $(T_{i-1},T_i]$, with $T_0 \equiv 0$. If at time T_i $(1 \le i \le m-1)$ the number of surviving units is more than R_i, then R_i is the number of surviving units that are selected at random, and removed (censored) from the life-test at time T_i. Otherwise, all the surviving units are removed from the test. The life-test ends at time T_m (at the latest), which means that all surviving units at time T_m are all censored at that time point.

In the case of the exponential distribution $\mathcal{E}(\theta)$, where $\theta > 0$ is the hazard rate, under a single-stage Type-I censored sample (i.e., case $m = 1$), [18] obtained an estimator of θ as

$$\hat{\theta}^{(1)} = -\frac{1}{T_1}\log\left(\frac{n-N_1}{n}\right), \tag{1}$$

and showed the asymptotic property that

$$\sqrt{n}\left(\hat{\theta}^{(1)}-\theta\right) \rightsquigarrow \mathcal{N}\left(0, \frac{1-\bar{F}(T_1,\theta)}{T_1^2\bar{F}(T_1,\theta)\theta^4}\right),$$

where $\bar{F}$ denotes the reliability function, and $\rightsquigarrow$ denotes the weak convergence.

When the failure times $X_{1:n},\ldots,X_{N_1:n}$ are observed, where N_1 is random, the maximum likelihood estimator of θ is

$$\hat{\theta}_{MLE} = \frac{N_1}{\sum_{i=1}^{N_1} X_{i:n} + (n-N_1)T_1}, \tag{2}$$

and further

$$\sqrt{n}(\hat{\theta}_{MLE}-\theta) \rightsquigarrow \mathcal{N}\left(0, \frac{\theta^2}{(1-\bar{F}(T_1,\theta))}\right).$$

Fig. 1 gives a comparison of the variances of $\hat{\theta}^{(1)}$ and $\hat{\theta}_{MLE}$ as a function of T_1 when $\theta = 1$. It can be seen that the optimal censoring time for the MLE in (2) is $T_1 = \infty$, while for Bartholomew's estimator in (1) the optimal censoring time is finite. This result implies that, for the estimator in (1), an optimal censoring time T_1 can be determined by minimizing the variance of the estimator.

Because order statistics arising from a multi-stage progressive censoring scheme are both left, and right truncated, it is difficult to handle the likelihood function; and for this reason, most of the inferential works are numerical in nature [9]. What we propose here is to use only partial information from such a progressively Type-I censored sample, to use the information on the number of failures occurring in each interval $(T_{i-1},T_i]$, and the number of censored units at each time T_i. The main idea here is to develop a non-parametric estimate of the reliability function at points T_i [2], [4]. Then, for a given parametric model, the unknown Euclidean parameter θ is obtained by the value $\hat{\theta}$ which minimizes a distance between the parametric reliability function and its nonparametric reliability estimate. In this sense, the work here extends the work of [18] because $\hat{\theta}^{(1)}$ in (1) minimizes $\{\exp(-\theta T_1) - (n-N_1)/n\}^2$.

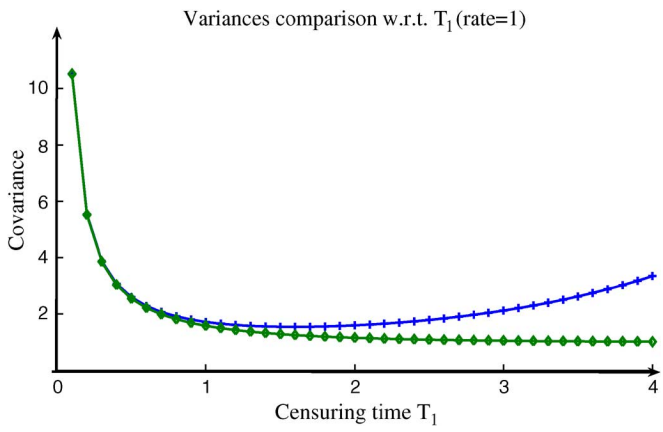

Fig. 1. Comparison of variances for $0 \le T_1 \le 4$: Bartholomew's estimator $(+)$ versus the MLE $(\diamondsuit)$.

Properties of consistency and asymptotic normality of the proposed estimator are considered. Moreover, given m estimates of θ, a minimum variance linear combination of these m estimates can be obtained from an m-stage sampling scheme. Furthermore, the maximum likelihood estimator for group-censored data is used. The efficiencies of these estimators are compared. Finally, we propose a method of determining an optimal progressive censoring scheme, and illustrate all the results developed here through a Monte Carlo simulation study, and an illustrative example provided in the following section.

The organization of the rest of this paper is as follows. First, we present a motivating example dealing with warranty analysis in Section II. We then discuss in Section III the construction of a nonparametric estimator of the reliability function, and its asymptotic behavior. Then in Section IV, for regular parametric models, we propose a minimum-distance method of estimation, and discuss its asymptotic properties. Next, in Section V, we show how the maximum likelihood estimator under group-censoring works, and then provide some numerical results in Section VI. We further discuss a method of determining an optimal progressive censoring scheme in Section VII. The motivating example presented earlier in Section II is the basis we use in Section VIII to illustrate all the inferential results developed in the preceding sections. Finally, we present some concluding remarks in Section IX. All proofs are given in the Appendix.

TABLE I
WARRANTY DATA ON $n = 1000$ UNITS WITH $m = 3$ WARRANTY TYPES

T_i	3	5	7
N_i	29	24	18
R_i	679	178	72

II. A MOTIVATING EXAMPLE

A consignment of $n = 1000$ units are sold by a dealer offering a basic warranty of 3 years for each unit, with options to purchase additional warranty for 2 or 4 more years. Of those 1000 units, 700 went to customers who just chose the basic warranty of 3 years, while 200 went to customers who chose an additional warranty for 2 more years, and the remaining 100 went to customers who chose an additional warranty for 4 more years.

In the first 3 years, 29 of the 1000 units failed out of which 21 were from customers with a basic warranty of 3 years, 6 were from customers with additional warranty of 2 years, and 2 were from customers with additional warranty of 4 years. The remaining 679 units with a basic warranty of 3 years get censored, which means $R_1 = 679$. Of the remaining 292 surviving units (past 3 years) with additional warranty, 24 units failed within the next 2 years, of which 16 were from customers with an additional warranty only for 2 years, so the remaining 178 units with additional warranty of only 2 years get censored, which means $R_2 = 178$. Finally, of the remaining 90 surviving units (past 5 years) with additional warranty for 2 more years, 18 units failed during this warranty period, which means $R_3 = 72$.

We thus arrive in this case at a 3-stage progressively Type-I censored data set, as presented in Table I. Based on these warranty data, we may wish to estimate parameters such as the mean lifetime of a unit, or the reliability function of the unit. For these purposes, we may develop either parametric or nonparametric methods of estimation based on such progressively Type-I censored data, and these form the subject matter of the following sections. The results for the warranty data in Table I are detailed in Section VIII.

III. NONPARAMETRIC ESTIMATION OF THE RELIABILITY

In this section, we construct a nonparametric estimator of the reliability function, and then discuss its asymptotic properties based on discrete martingale theory. We suppose that the true model is defined by a univariate unknown parameter $\theta \in \Theta \subset \mathbb{R}$ (the case $\Theta \subset \mathbb{R}^p$ with $p > 1$ will be discussed later in Section IV).

A. The Nonparametric Estimator

We suppose that the n s-independent units in the life-test all have the same lifetime distribution function F (and reliability function $\bar{F}$), $F \equiv F(\cdot;\theta)$ with $\theta \in \Theta \subset \mathbb{R}$, and that for $1 \leq i \leq m$ the function $\theta \mapsto \bar{F}(T_i;\theta)$ is one-to-one. This assumption means that, when $r = \bar{F}(T_i,\theta)$, there exists a unique $\theta \in \Theta$ such that

$$r = \bar{F}(T_i,\theta) \Leftrightarrow \theta = G(T_i,r). \tag{3}$$

Remark 3.1: For example, in the exponential case, we have $\bar{F}(T,\theta) = \exp(-\theta T) = r$, and so $\theta = -\log(r)/T$ for $T > 0$.

Now, given a nonparametric estimate of the reliability level r, we can obtain an estimator of the unknown Euclidean parameter θ. For $m = 1$ corresponding to a single-stage Type-I censored sampling plan, it is easy to see that the variable N_1 has a binomial distribution with parameters $(n, F(T_1,\theta))$. Then, $(n - N_1)/n$ is an estimator of $\bar{F}(T_1,\theta)$, and so θ may be estimated by $\hat{\theta}^{(1)} = G(T_1,(n - N_1)/n)$.

In the case of $m = 2$, we can still estimate θ by $\theta^{(1)}$. Moreover, the random variable N_2, if not null, contains information that can be accounted for. Indeed, conditional on $(N_1,R_1) = (n_1,r_1)$, the lifetimes of the remaining $n - n_1 - r_1$ units under test follow a left-truncated distribution with density

$$f_{LT}(t) = \frac{f(t\theta)}{\bar{F}(T_1,\theta)} 1(t \geq T_1);$$

see [1]. Then, conditional on $(N_1,R_1) = (n_1,r_1)$, and $n_1 + r_1 < n$, the random variable N_2 has a binomial distribution with parameters $n - n_1 - r_1$, and $(F(T_2,\theta) - F(T_1,\theta))/\bar{F}(T_1,\theta)$. Consequently, $1 - p = \bar{F}(T_2,\theta)/\bar{F}(T_1,\theta)$ is approximated by $(n - n_1 - r_1 - n_2)/(n - n_1 - r_1)$. And because $\bar{F}(T_1,\theta)$ is approximated by $(n - n_1)/n$, we can approximate $\bar{F}(T_2,\theta)$ by

$$\bar{F}(T_2,\theta) \approx \frac{(n - n_1)(n - n_1 - r_1 - n_2)}{n(n - n_1 - r_1)},$$

which results in $\hat{\theta}^{(2)} = G(T_2,(n-n_1)(n-n_1-r_1-n_2)/(n(n-n_1-r_1)))$ as a second estimate of θ. If we now estimate θ by the value that is as close as possible to both $\hat{\theta}_1$ and $\hat{\theta}_2$, we obtain the estimator $\hat{\theta} = (1/2)\sum_{i=1}^{2} \hat{\theta}^{(i)}$ for θ.

Remark 3.2: For the exponential case, we obtain

$$\hat{\theta} = \frac{1}{2T_1}\log\left(\frac{n}{n-n_1}\right) + \frac{1}{2T_2}\log\left(\frac{n(n-n_1-r_1)}{(n-n_1)(n-n_1-r_1-n_2)}\right).$$

In the case of $m \geq 2$, by means of induction, we have for $j = 1,\ldots,m$ that

$$\alpha_j^- = \#\{i \in \{1,\cdots,n\} : X_i \geq T_j\},$$
$$\alpha_j^+ = \#\{i \in \{1,\cdots,n\} : X_i > T_j\}.$$

We can show that $\bar{F}(T_i,\theta)$ is approximated by

$$\bar{F}(T_i,\theta) \approx \prod_{j=1}^{i} \frac{\alpha_j^-}{\alpha_{j-1}^+},$$

with the convention that $0/0 = 0$. Using this approximation, we can estimate θ by

$$\hat{\theta}^{(i)} = G\left(T_i, \prod_{j=1}^{i} \frac{\alpha_j^-}{\alpha_{j-1}^+}\right), \tag{4}$$

and then, using the above m estimators of θ, we can propose a linear estimator of θ as

$$\hat{\theta} = \sum_{i=1}^{m} \pi_i \hat{\theta}^{(i)}, \tag{5}$$

where $\pi = (\pi_1, \ldots, \pi_m)$ is a $\mathbb{R}^m$-valued vector such that $\sum_{i=1}^m \pi_i = 1$.

We will show later in Section III-D that π may be chosen so that $\hat{\theta}$ is the best linear combination of the $\hat{\theta}^{(i)}$ in the sense of the minimum variance criterion.

Remark 3.3: The estimators in (4), and (5) take into account only partial information from the progressively censored sample. Indeed, to estimate θ, we only need to know the N_i, and the R_i, because

$$\alpha_j^- = n - \sum_{k=1}^{j} N_k - \sum_{k=1}^{j-1} R_k, \quad \text{and} \quad \alpha_j^+ = n - \sum_{k=1}^{j} N_k - \sum_{k=1}^{j} R_k.$$

B. Nelson-Aalen, and Kaplan-Meier Type Estimators

Let us introduce the discrete time filtration $\mathbb{F} = \{\mathcal{F}_k; k = 0, \ldots, m\}$, where $\mathcal{F}_k = \sigma\{(N_j, R_j); 0 \le j \le k\}$ with $N_0 = R_0 = 0$. Let us consider the counting processes

$$K_i = \sum_{j=0}^{i} N_i, \quad \text{and} \quad Y_i = n - \sum_{j=0}^{i} (N_j + R_j)$$

for $i = 0, 1, \ldots, m$. Let us write $\lambda_i = (\bar{F}_{i-1} - \bar{F}_i)/\bar{F}_{i-1}$ for $i = 1, \ldots, m$, with $\bar{F}_i = \bar{F}(T_i, \theta)$, and $T_0 = 0$. Then, we have the following result.

Proposition 3.4: The discrete time process $L = \{L_i; 1 \le i \le m\}$, defined by $L_i = K_i - \sum_{\ell=1}^{i} Y_{\ell-1}\lambda_\ell = K_i - A_i$, is an $\mathbb{F}$-martingale.

Let $\Lambda = (\Lambda_1, \ldots, \Lambda_m)$ be the cumulative hazard rate function, where for $1 \le j \le m$,

$$\Lambda_j = \sum_{i=1}^{j} \lambda_i = \sum_{i=1}^{j} (\bar{F}_{i-1} - \bar{F}_i)/\bar{F}_{i-1}.$$

Note that we have interest in these quantities because, for discrete distributions, the Λ_j correspond to the cumulative hazard rate calculated at each time point T_j. Moreover, the reliability $\bar{F}_j$ is naturally linked to the cumulative hazard rate function by the product integral, which means that the relation between the Λ_j and the $\bar{F}_j$ is given by

$$\bar{F}_j = \prod_{i=1}^{j} (1 - \Delta\Lambda_i) = \prod_{i=1}^{j} (1 - \Lambda_i + \Lambda_{i-1}) = \prod_{i=1}^{j} (1 - \lambda_i).$$

It yields the following estimators for the Λ_j, and the $\bar{F}_j$. Because, for a F-distributed random variable X, we have

$$\lambda_j = P\left(X \in (T_{j-1}, T_j] | X > T_j\right),$$

it is natural to estimate λ_j, and Λ_j by

$$\hat{\lambda}_j = \frac{N_j}{Y_{j-1}}, \quad \text{and} \quad \hat{\Lambda}_j = \sum_{i=1}^{j} \frac{N_i}{Y_{i-1}},$$

respectively. Finally, the reliability function is estimated by

$$\hat{\bar{F}}_j = \prod_{i=1}^{j} (1 - \Delta\hat{\Lambda}_i) = \prod_{i=1}^{j} (1 - \hat{\lambda}_i) = \prod_{i=1}^{j} \left(1 - \frac{N_i}{Y_{i-1}}\right) = \prod_{i=1}^{j} \frac{Y_{i-1} - N_i}{Y_{i-1}},$$

as given earlier in Section III-A.

C. Asymptotic Properties

Let us introduce the process $M = (M_1, \ldots, M_m)$, where $M_i = L_i - L_{i-1} = N_i - Y_{i-1}\lambda_i$ for $i = 1, \ldots, m$.

Proposition 3.5: The discrete time process $H = \{H_i; 1 \le i \le m\}$, where

$$H_i = \sum_{j=1}^{i} \left(M_j^2 - Y_{j-1}\lambda_j(1 - \lambda_j)\right),$$

is an $\mathbb{F}$-martingale.

Let us introduce, for $1 \le i \le m$, the quantities (provided they exist)

$$\gamma_i = \lim_{n\to\infty} \frac{N_i}{n}, \quad \text{and} \quad \beta_i = \lim_{n\to\infty} \frac{Y_i}{n},$$

with $\beta_0 = 1$, and limits with respect to the convergence in probability.

Proposition 3.6: For $1 \le i \le m$, the limits β_i, and γ_i exist, and satisfy

$$\beta_i = \bar{F}_i \prod_{j=1}^{i} (1 - \alpha_j), \quad \text{and} \quad \gamma_i = (\bar{F}_{i-1} - \bar{F}_i) \prod_{j=1}^{i-1} (1 - \alpha_j).$$

Corollary 3.7: For $1 \le i \le m$, we have, as $n \to +\infty$,

(I) $\hat{\lambda}_i \xrightarrow{P} \lambda_i$,

(II) $\hat{\Lambda}_i \xrightarrow{P} \Lambda_i$, and

(III) $\hat{\bar{F}}_i \xrightarrow{P} \bar{F}_i$.

By means of induction,

$$\lambda = (\lambda_1, \cdots, \lambda_m), \ \Lambda = (\Lambda_1, \cdots, \Lambda_m), \text{ and } \bar{F} = (\bar{F}_1, \cdots, \bar{F}_m)$$

and

$$\hat{\lambda} = (\hat{\lambda}_1, \cdots, \hat{\lambda}_m), \ \hat{\Lambda} = (\hat{\Lambda}_1, \cdots, \hat{\Lambda}_m), \text{ and } \hat{\bar{F}} = (\hat{\bar{F}}_1, \cdots, \hat{\bar{F}}_m).$$

Theorem 3.8: If $\bar{F}_m > 0$, then as $n \to \infty$, we have $\sqrt{n}(\hat{\lambda} - \lambda) \rightsquigarrow \mathcal{N}(0, \Sigma)$, where Σ is a $m \times m$ matrix whose (i, j)-th entry is given by

$$\sigma_{ij} = \begin{cases} \frac{\lambda_i(1-\lambda_i)}{\beta_{i-1}}, & i = j \\ 0, & i \ne j. \end{cases}$$

Moreover, the $m \times m$ matrix $\hat{\Sigma} = (\hat{\sigma}_{ij})_{1 \le i,j \le m}$ defined by

$$\hat{\sigma}_{ij} = \begin{cases} \frac{\hat{\lambda}_i(1-\hat{\lambda}_i)}{\hat{\beta}_{i-1}}, & i = j \\ 0, & i \ne j, \end{cases}$$

is a consistent estimator of Σ where

$$\hat{\beta}_i = \hat{\bar{F}}_i \prod_{j=1}^{i}(1-\alpha_j).$$

Now, let us introduce three $m \times m$ matrices:

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 1 & 1 & 1 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix},$$

$$B = -\begin{pmatrix} \frac{\bar{F}_1}{1-\lambda_1} & 0 & 0 & \cdots & 0 \\ \frac{\bar{F}_2}{1-\lambda_1} & \frac{\bar{F}_2}{1-\lambda_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \frac{\bar{F}_{m-1}}{1-\lambda_1} & \frac{\bar{F}_{m-1}}{1-\lambda_2} & \frac{\bar{F}_{m-1}}{1-\lambda_3} & \cdots & 0 \\ \frac{\bar{F}_m}{1-\lambda_1} & \frac{\bar{F}_m}{1-\lambda_2} & \frac{\bar{F}_m}{1-\lambda_3} & \cdots & \frac{\bar{F}_m}{1-\lambda_m} \end{pmatrix},$$

and

$$\hat{B} = -\begin{pmatrix} \frac{\hat{\bar{F}}_1}{1-\hat{\lambda}_1} & 0 & 0 & \cdots & 0 \\ \frac{\hat{\bar{F}}_2}{1-\hat{\lambda}_1} & \frac{\hat{\bar{F}}_2}{1-\hat{\lambda}_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \frac{\hat{\bar{F}}_{m-1}}{1-\hat{\lambda}_1} & \frac{\hat{\bar{F}}_{m-1}}{1-\hat{\lambda}_2} & \frac{\hat{\bar{F}}_{m-1}}{1-\hat{\lambda}_3} & \cdots & 0 \\ \frac{\hat{\bar{F}}_m}{1-\hat{\lambda}_1} & \frac{\hat{\bar{F}}_m}{1-\hat{\lambda}_2} & \frac{\hat{\bar{F}}_m}{1-\hat{\lambda}_3} & \cdots & \frac{\hat{\bar{F}}_m}{1-\hat{\lambda}_m} \end{pmatrix}.$$

Applying the δ-method, we obtain the following corollary.

Corollary 3.9: If $\bar{F}_m > 0$, then as $n \to \infty$, we have $\sqrt{n}(\hat{\Lambda} - \Lambda) \rightsquigarrow \mathcal{N}(0, \Gamma)$, and $\sqrt{n}(\hat{\bar{F}} - \bar{F}) \rightsquigarrow \mathcal{N}(0, \Upsilon)$, where $\Gamma = A\Sigma A^T$, and $\Upsilon = B\Sigma B^T$. Moreover, $\hat{\Gamma} = A\hat{\Sigma}A^T$, and $\hat{\Upsilon} = \hat{B}\hat{\Sigma}\hat{B}^T$ are consistent estimators of Γ, and Υ, respectively.

D. The Minimum Variance Linear Estimator

Given m estimators $\hat{\theta}^{(i)}$ $(1 \le i \le m)$ of θ (obtained in Section III-A), let $\tilde{\theta} = (\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(m)})$, and $\tilde{\theta}_0 = (\theta_0, \ldots, \theta_0)$. Suppose that $r \mapsto G(T_i, r)$ is a continuously differentiable function (or, equivalently, that $\theta \mapsto \bar{F}(T_i, \theta)$ is a continuously differentiable function with $\partial \bar{F}(T_i, \theta_0)/\partial\theta \neq 0$).

Then, define two new $m \times m$ diagonal matrices:

$$H = \operatorname{diag}\left(\left.\frac{\partial G(T_i, x)}{\partial x}\right|_{x=\bar{F}_i}; i = 1, \cdots, m \right),$$

and

$$\hat{H} = \operatorname{diag}\left(\left.\frac{\partial G(T_i, x)}{\partial x}\right|_{x=\hat{\bar{F}}_i}; i = 1, \cdots, m \right).$$

Corollary 3.10: If $\bar{F}_m > 0$, then as $n \to \infty$, we have $\sqrt{n}(\tilde{\theta} - \tilde{\theta}_0) \rightsquigarrow \mathcal{N}(0, \Omega)$, where $\Omega = H\Upsilon H^T$. Moreover, $\hat{\Omega} = \hat{H}\hat{\Upsilon}\hat{H}^T$ is a consistent estimator of Ω.

Remark 3.11: For the exponential case, the matrix H is simply

$$H = \operatorname{diag}\left((T_i \bar{F}_i)^{-1}; i = 1, \cdots, m\right).$$

Now, we seek the best linear estimate of θ as follows. Let $\pi = (\pi_1, \ldots, \pi_m)^T$ be an $\mathbb{R}^m$ real vector with components summing to 1, and $\hat{\theta} = \sum_{i=1}^m \pi_i \hat{\theta}^{(i)}$ be an estimator of θ. Then, we seek π such that

$$\hat{\pi} = \arg\min_{\pi \in \mathbb{R}^m} \operatorname{var}(\hat{\theta}) = \arg\min_{\pi \in \mathbb{R}^m} \pi^T \Omega \pi$$

subject to the condition $\pi^T C = 1$, where $C = (1, 1, \ldots, 1)^T$. Then, it is easy to obtain by the Lagrangian multiplier method that

$$\tilde{\pi} = (C^T \Omega^{-1} C)^{-1} \Omega^{-1} C.$$

which is consistently estimated by

$$\tilde{\pi} = (\tilde{\pi}_1, \cdots, \tilde{\pi}_m) = (C^T \hat{\Omega}^{-1} C)^{-1} \hat{\Omega}^{-1} C.$$

With this result, we finally obtain the minimum variance linear estimator of θ as

$$\hat{\theta} = \sum_{i=1}^m \tilde{\pi}_i \hat{\theta}^{(i)}. \tag{6}$$

IV. Minimum-Distance Estimator

In this section, we consider a parametric model defined by the Euclidean parameter $\theta \in \Theta \subset \mathbb{R}^p$. We then propose a simple estimator of θ, and discuss its asymptotic properties as well.

A. Estimation of the Parameter

We propose to estimate θ by minimizing the square of the Euclidean distance between $(\bar{F}_1, \ldots, \bar{F}_m)$ and its non-parametric estimate. The estimator is therefore defined by

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \sum_{i=1}^m \left(\bar{F}(T_i, \theta) - \prod_{j=1}^{i} \frac{\alpha_j^-}{\alpha_{j-1}^+} \right)^2. \tag{7}$$

Remark 4.1: Some other distance function can also be used for this purpose; for example, for the Weibull distribution, the $x \mapsto \log(-\log(x))$ transformation leads to the solution of a simple linear regression problem.

B. Asymptotic Behavior of the Estimator

Let us denote by $\theta_0 = (\theta_1^0, \ldots, \theta_p^0)$ the unknown parameter, and $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_p)$ its corresponding estimator. Define

$$\hat{\phi}_n(\theta) = \sum_{i=1}^m \left(\bar{F}(T_i, \theta) - \hat{\bar{F}}_i \right)^2,$$

and

$$\phi(\theta) = \sum_{i=1}^m \left(\bar{F}(T_i, \theta) - \bar{F}(T_i, \theta_0) \right)^2.$$

We then have

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \hat{\phi}_n(\theta), \quad \text{and} \quad \phi(\theta_0) = 0.$$

The asymptotic behavior of the estimator $\hat{\theta}$ is obtained under three assumptions:

(1) the space Θ is compact,
(2) the mapping $\theta \mapsto \bar{F}(T_i, \theta)$ belong to $C^2(\Theta)$ for $1 \leq i \leq m$, and
(3) the matrix $I_0 = \partial^2\phi(\theta_0)/\partial\theta\partial\theta^T$ is positive-definite.

Then, the functions $\hat{\phi}_n$, and ϕ satisfy the following results.

Lemma 4.2: Under Assumptions (A1)–(A3), we have

(I) $\sup_{\theta\in\Theta} |\hat{\phi}_n(\theta) - \phi(\theta)| = O_P(n^{-1/2})$,
(II) $\sup_{\theta\in\Theta} \|(\partial\hat{\phi}_n(\theta)/\partial\theta) - (\partial\phi(\theta)/\partial\theta)\| = O_P(n^{-1/2})$, and
(III) $\sup_{\theta\in\Theta} \|(\partial^2\hat{\phi}_n(\theta)/\partial\theta\partial\theta^T) - (\partial^2\phi(\theta)/\partial\theta\partial\theta^T)\| = O_P(n^{-1/2})$.

We can now state the following consistency result.

Theorem 4.3: Let $\hat{\theta}$ be the estimator of θ defined by (7). If ϕ is a contrast function, i.e., $\phi(\theta) = 0$ iff $\theta = \theta_0$, then we have $\hat{\theta} \xrightarrow{P} \theta_0$ as $n \to \infty$.

Moreover, the estimator $\hat{\theta}$ satisfies the following version of the central limit theorem. Let the $p \times m$ matrix $Q(\theta)$ be with entries $q_{ij}(\theta) = \partial\bar{F}(T_j, \theta)/\partial\theta_i$.

Theorem 4.4: Let $\hat{\theta}$ be the estimator of θ defined by (7). If ϕ is a contrast function, then we have

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightsquigarrow \mathcal{N}(0, M\Upsilon M^T), \tag{8}$$

where Υ is as defined in Corollary 3.10, and $M \equiv M(\theta_0)$ defined by

$$M(\theta) = -2\left(\frac{\partial^2\phi(\theta)}{\partial\theta\partial\theta^T}\right)^{-1} \times Q(\theta)$$

is consistently estimated by $\hat{M} = M(\hat{\theta})$.

V. Maximum Likelihood Estimator Under Group-Censoring

In addition to the above nonparametric estimation method, we can also use the maximum likelihood principle for group-censored data to estimate the unknown Euclidean parameter θ. For the special case of the exponential distribution, we can obtain the estimator as well as its asymptotic variance in a closed form.

From the partial information from $(N_i, R_i)_{1\leq i\leq m}$, we can estimate θ by the maximum likelihood estimation method (see [17] for an application to the exponential model with changing failure rates, and [19] for an application to the exponential distribution), based on the likelihood function

$$L(\theta) = \prod_{i=1}^{m}\left(\left(\int_{T_{i-1}}^{T_i} f(x,\theta)dx\right)^{N_i}\left(\bar{F}(T_i,\theta)\right)^{R_i}\right).$$

We can therefore estimate θ by maximizing the log-likelihood function as

$$\begin{aligned}\hat{\theta} &= \arg\max_{\theta\in\mathbb{R}^p} \log L(\theta)\\ &= \arg\max_{\theta\in\mathbb{R}^p} \sum_{i=1}^{m}\big(N_i \log\left(\bar{F}(T_{i-1},\theta) - \bar{F}(T_i,\theta)\right)\\ &\qquad + R_i \log\left(\bar{F}(T_i,\theta)\right)\big).\end{aligned} \tag{9}$$

Then the score function is

$$\begin{aligned}U(\theta) &= \frac{\partial \log L(\theta)}{\partial\theta}\\ &= \sum_{i=1}^{m}\left(N_i\frac{\partial\bar{F}(T_{i-1},\theta)/\partial\theta - \partial\bar{F}(T_i,\theta)/\partial\theta}{\bar{F}(T_{i-1},\theta) - \bar{F}(T_i,\theta)}\right.,\\ &\qquad \left. + R_i\frac{\partial\bar{F}(T_i,\theta)/\partial\theta}{\bar{F}(T_i,\theta)}\right)\end{aligned}$$

and the estimator $\hat{\theta}$ can be derived by solving $U(\theta) = 0$. For a regular parametric model, we can show that

$$\sqrt{n}(\hat{\theta} - \theta_0) \rightsquigarrow \mathcal{N}\left(0, I^{-1}(\theta_0)\right), \tag{10}$$

where $I(\theta_0)$ is the Fisher information matrix given by

$$I(\theta_0) = \sum_{i=1}^{m}\left(\frac{A_i\prod_{j=1}^{i-1}(1-\alpha_j)}{\bar{F}(T_{i-1},\theta_0) - \bar{F}(T_i,\theta_0)} - \frac{B_i\alpha_i\prod_{j=1}^{i-1}(1-\alpha_j)}{\bar{F}(T_i,\theta_0)}\right)$$

with

$$\begin{aligned}A_i(\theta) &= \left(\frac{\partial\bar{F}(T_{i-1},\theta)}{\partial\theta} - \frac{\partial\bar{F}(T_i,\theta)}{\partial\theta}\right)\\ &\quad\times\left(\frac{\partial\bar{F}(T_{i-1},\theta)}{\partial\theta} - \frac{\partial\bar{F}(T_i,\theta)}{\partial\theta}\right)^T\\ &\quad - \left(\frac{\partial^2\bar{F}(T_{i-1},\theta)}{\partial\theta\partial\theta^T} - \frac{\partial^2\bar{F}(T_i,\theta)}{\partial\theta\partial\theta^T}\right)\\ &\quad\times\left(\bar{F}(T_{i-1,\theta}) - \bar{F}(T_i,\theta)\right),\end{aligned}$$

and

$$B_i(\theta) = \left(\frac{\partial\bar{F}(T_i,\theta)}{\partial\theta}\right)\left(\frac{\partial\bar{F}(T_i,\theta)}{\partial\theta}\right)^T - \left(\frac{\partial^2\bar{F}(T_i,\theta)}{\partial\theta\partial\theta^T}\right)\bar{F}(T_i,\theta).$$

For the exponential case, the above information matrix reduces to

$$I(\theta) = \sum_{i=1}^{m}\frac{(T_i - T_{i-1})^2\bar{F}(T_i,\theta)\bar{F}(T_{i-1},\theta)}{\bar{F}(T_{i-1},\theta) - \bar{F}(T_i,\theta)}\prod_{j=1}^{i-1}(1-\alpha_j).$$

VI. Numerical Study and Examples

In this section, we evaluate the behavior of all the estimators using Monte Carlo simulations.

A. Numerical Examples for the Exponential Distribution

We consider the m-stage procedure with $m = 4$. The estimator $\theta^{(i)}$ is defined by (4). We chose $T = (T_1, \ldots, T_4) = (1, 2, 4, 5)$, and $\alpha_i = 0.2$ for $i = 1, 2, 3$. Then, at time $T_i (i = 1, \ldots, 3)$, 20% of the units still functioning are removed from the sample, which means that $R_i = \lfloor \alpha_i\alpha_i^{-} \rfloor$. The experiment terminates at time T_4, i.e., the lifetimes of all surviving units at time T_4 are censored.

TABLE II
ESTIMATION OF THE HAZARD RATE 0.5 OF AN EXPONENTIAL DISTRIBUTION FOR THE 4-STAGE PROGRESSIVE CENSORING: MEAN, AND STANDARD DEVIATION (WITHIN PARENTHESES) OF $N = 1000$ BARTHOLOMEW'S ($\theta^{(i)}$), WSBE, AND MLE ESTIMATES

sample size n	$\theta^{(1)}$	$\theta^{(2)}$	$\theta^{(3)}$	$\theta^{(4)}$	WSBE	MLE
20	0.5145 (0.1800)	0.5263 (0.1734)	0.4459 (0.2014)	0.3236 (0.2186)	0.4526 (0.1395)	0.5228 (0.1135)
25	0.5127 (0.1553)	0.5209 (0.1534)	0.4839 (0.1746)	0.3845 (0.2065)	0.4755 (0.1067)	0.5143 (0.1178)
30	0.5078 (0.1532)	0.5163 (0.1333)	0.5025 (0.1681)	0.4203 (0.2049)	0.4863 (0.1025)	0.5154 (0.1096)
40	0.5054 (0.1285)	0.5118 (0.1159)	0.5169 (0.1487)	0.4573 (0.1883)	0.4969 (0.0889)	0.5123 (0.0958)
50	0.5077 (0.0701)	0.5097 (0.0702)	0.4996 (0.0566)	0.4972 (0.0387)	0.5036 (0.1204)	0.5097 (0.0656)
100	0.4961 (0.0490)	0.4975 (0.0529)	0.4972 (0.0424)	0.4976 (0.0282)	0.4971 (0.0883)	0.5067 (0.0568)
200	0.4978 (0.0331)	0.4974 (0.0346)	0.4973 (0.0282)	0.4963 (0.0208)	0.4972 (0.0592)	0.5042 (0.0447)
500	0.5014 (0.0223)	0.5001 (0.0245)	0.4996 (0.0173)	0.4993 (0.0141)	0.5011 (0.0400)	0.5015 (0.0265)
1000	0.4987 (0.0141)	0.4994 (0.0173)	0.4997 (0.0141)	0.4992 (0.0102)	0.4993 (0.0283)	0.5026 (0.0181)

TABLE III
MINIMUM DISTANCE ESTIMATES OF THE HAZARD RATE 0.5 OF AN EXPONENTIAL DISTRIBUTION FOR THE 4-STAGE PROGRESSIVE CENSORING: MEAN, AND STANDARD DEVIATION (WITHIN PARENTHESES) OF $N = 1000$, MDE, AND OWSBE ESTIMATES

sample size n	MDE	OWSBE
20	0.5288 (0.1491)	0.5077 (0.0874)
25	0.5148 (0.1199)	0.4932 (0.0838)
30	0.5132 (0.1114)	0.5061 (0.0823)
40	0.5115 (0.0958)	0.4988 (0.0714)
50	0.5102 (0.0846)	0.5045 (0.0641)
100	0.5043 (0.0582)	0.4970 (0.0471)
200	0.5014 (0.0411)	0.4972 (0.0313)
500	0.5009 (0.0257)	0.5002 (0.0214)
1000	0.4998 (0.0186)	0.4992 (0.0150)

Because the estimator $\hat{\theta}$ in (5) is based on the weighted sum of several Bartholomew's estimators, it is denoted by WSBE; while the estimator in (9), based on the maximum likelihood method, is denoted by MLE. The WSBE is obtained here with weights $\pi_i = 1/4\ (1 \leq i \leq 4)$. The performance of these two estimators is illustrated for various sample sizes in Table II, and for each sample size the mean and the standard deviation were obtained from $N = 1000$ simulated samples. We can see that the two estimators behave quite similarly, but for small sample size, the WSBE seems to outperform the MLE in terms of bias.

We compare the performance of the minimum distance estimator (MDE) in (7) to the WSBE with optimal weights (OWSBE) defined in (6). The numerical results presented in Table III show that these two estimators have good performance even with small sample sizes.

B. Multivariate Case: $\theta \in \Theta \subset \mathbb{R}^p$

Table IV presents estimation results of a Weibull distribution for 4-stage progressively Type-I censored samples with $T_1 = 2$, $T_2 = 4, T_3 = 6, T_4 = 8$, and $\alpha_i = 0.2$. The Weibull distribution has the reliability function

$$\bar{F}(t,\theta) = \bar{F}(t,\gamma,\sigma) = \exp\left(-(t/\sigma)^{\gamma}\right),$$

where $\gamma > 0$ is the shape parameter, σ is the scale parameter, and in this case $\theta = (\gamma, \sigma)$. Defining $g(x) = \log(-\log(x))$, $\beta_0 = \gamma$, and $\beta_1 = -\gamma\log(\sigma)$, we have for $1 \leq i \leq 2$

$$\begin{pmatrix}\log(T_i) & 1\end{pmatrix}\begin{pmatrix}\beta_0 \\ \beta_1\end{pmatrix} = g\left(\bar{F}(T_i,\theta)\right).$$

Replacing in the above equation the unknown quantities $\bar{F}(T_i,\theta)$ by their corresponding non-parametric estimates $\hat{\bar{F}}_i$, and solving

$$\begin{pmatrix}\log(T_i) & 1 \\ \log(T_{i+1}) & 1\end{pmatrix}\begin{pmatrix}\hat{\beta}_0^{(i)} \\ \hat{\beta}_1^{(i)}\end{pmatrix} = \begin{pmatrix}g(\hat{\bar{F}}_i) \\ g(\hat{\bar{F}}_{i+1})\end{pmatrix},$$

we estimate the regression parameters $\gamma^{(i)}$, and $\sigma^{(i)}$ by $\hat{\gamma}^{(i)} = \hat{\beta}_0^{(i)}$, and $\hat{\sigma}^{(i)} = \exp(-\hat{\beta}_0^{(i)}/\hat{\beta}_1^{(i)})$, respectively.

Results on $\theta^{(i)}$ in Table IV show that this estimator does not behave well. We observe especially large bias, and standard deviation for small to moderate sample sizes, which show that this estimator converges quite slowly. Alternatively, we can estimate θ by 1) the minimum-distance method leading to the estimator $\hat{\theta}$ defined by (7), or 2) the WSBE with optimal weights (OWSBE), and compare the behavior of the minimum-distance estimator (MDE) with the OWSBE. The corresponding numerical results are presented in Table IV. From these results, see that the OWSBE behaves poorly, while the MDE shows good convergence properties.

In Table V, we have presented some numerical results for the MDE with the log-logistic distribution with $(\gamma, \sigma) = (2, 4)$. Yet again, we see that, even for small to moderate sample sizes, this method yields estimators with a good performance.

VII. SEQUENTIAL PROGRESSIVE TYPE-I CENSORING PLANS

In this section, we discuss how an optimal progressive Type-I censoring plan can be determined by using MDE, and MLE. For each of these estimators, we give the asymptotic variance-covariance matrix as a function of the unknown parameter θ. This matrix is consistently estimated by providing estimates into the expression whenever we have a consistent estimator of θ, from which we determine the optimal censoring times by using the determinant criterion.

A. Univariate Case $\theta \in \Theta \subset \mathbb{R}$

Suppose that θ belongs to $\mathbb{R}$, and assume that $m > 1, T_1 > 0$, and $\alpha_1, \ldots, \alpha_{m-1}$ are given. We now present a step-by-step method that enables the determination of an m-stage optimal progressive censoring plan. We denote by $\hat{\theta}^{(i)}$ the MLE or MDE of θ that uses the available information on $[0, T_i]\ (1 \leq i \leq m)$.

Step 1. Given the first censoring time T_1 (and using N_1), obtain $\hat{\theta}^{(1)}$.

Step 2. Calculate the asymptotic variance of $\hat{\theta}^{(2)}$ as a function of T_1, T_2, and θ, replacing θ by $\hat{\theta}^{(1)}$. Then, find the

TABLE IV
REGRESSION ESTIMATES OF THE WEIBULL PARAMETERS: MEAN, AND STANDARD DEVIATION (WITHIN PARENTHESES) BASED ON $N = 1000$ TYPE-I PROGRESSIVELY CENSORED SAMPLES ($m = 4, T_1 = 2, T_2 = 4, T_3 = 6, T_4 = 8$, AND $\alpha_i = 0.2$)

sample size n	θ	$\theta^{(1)}$	$\theta^{(2)}$	$\theta^{(3)}$	$\theta^{(4)}$	MDE	OWSBE
20	$\hat{\sigma}$	4.2055 (1.4097)	3.9044 (0.7324)	2.9000 (1.6122)	0.9489 (1.7468)	4.0341 (0.3968)	4.0264 (1.2735)
	$\hat{\gamma}$	2.2213 (1.6322)	2.4697 (1.2133)	2.3771 (2.6645)	0.4791 (0.9159)	2.0468 (0.4154)	2.2677 (1.5918)
25	$\hat{\sigma}$	4.1240 (0.8360)	3.9290 (0.6672)	3.1252 (1.3922)	1.0498 (1.7767)	4.0310 (0.3348)	3.9880 (0.8418)
	$\hat{\gamma}$	2.1420 (1.1423)	2.3130 (0.9284)	2.4578 (1.8560)	0.5842 (0.9982)	2.0320 (0.3608)	2.1829 (1.1362)
30	$\hat{\sigma}$	4.0967 (0.6755)	3.9563 (0.5958)	3.2759 (1.2789)	1.3098 (1.8724)	4.0302 (0.2859)	3.9840 (0.7083)
	$\hat{\gamma}$	2.0952 (0.6062)	2.2369 (0.6909)	2.3944 (1.0488)	0.6983 (1.0346)	2.0480 (0.3343)	2.1276 (0.6563)
40	$\hat{\sigma}$	4.0717 (0.5256)	3.9924 (0.4830)	3.4541 (0.9912)	1.5258 (1.9760)	4.0282 (0.2453)	3.9801 (0.5899)
	$\hat{\gamma}$	2.0740 (0.5027)	2.1393 (0.5289)	2.5097 (0.8650)	0.7702 (1.0174)	2.0239 (0.2582)	2.1015 (0.5360)
50	$\hat{\sigma}$	4.0701 (0.5081)	3.9761 (0.4296)	3.5391 (0.8691)	1.7331 (1.9975)	4.0139 (0.1843)	3.9825 (0.5824)
	$\hat{\gamma}$	2.0328 (0.4430)	2.0953 (0.4565)	2.4807 (0.7769)	0.8991 (1.0419)	2.0221 (0.2109)	2.0504 (0.4877)
100	$\hat{\sigma}$	4.0311 (0.3036)	3.9968 (0.2764)	3.8118 (0.5146)	2.7865 (1.9062)	4.0050 (0.1100)	4.0030 (0.4001)
	$\hat{\gamma}$	2.0274 (0.3043)	2.039 (0.2295)	2.3123 (0.5680)	1.3635 (0.9392)	2.0050 (0.1232)	2.0420 (0.3363)
200	$\hat{\sigma}$	4.0186 (0.2078)	3.9988 (0.2061)	3.9940 (0.7672)	3.5641 (1.3074)	4.0049 (0.0640)	4.0006 (0.3227)
	$\hat{\gamma}$	2.0114 (0.2142)	2.0308 (0.2902)	2.4712 (1.5633)	1.7649 (0.6554)	2.0052 (0.0728)	2.0285 (0.4502)
500	$\hat{\sigma}$	4.0005 (0.1280)	3.9993 (0.1232)	3.9850 (0.1606)	3.9834 (0.3337)	4.0010 (0.0400)	3.9991 (0.1355)
	$\hat{\gamma}$	2.0116 (0.1315)	2.0104 (0.1014)	2.0253 (0.1356)	1.9925 (0.1868)	2.0012 (0.0387)	2.0116 (0.1284)
1000	$\hat{\sigma}$	4.008 (0.0916)	3.9982 (0.0848)	3.9950 (0.1095)	4.001 (0.15490)	4.0009 (0.0331)	4.0050 (0.0931)
	$\hat{\gamma}$	1.998 (0.0932)	2.0040 (0.0714)	2.0078 (0.0812)	1.9969 (0.0911)	2.0018 (0.0331)	2.0000 (0.0892)

TABLE V
MINIMUM-DISTANCE ESTIMATES OF THE LOG-LOGISTIC DISTRIBUTION PARAMETERS: MEAN, AND STANDARD DEVIATION (WITHIN PARENTHESES) OBTAINED FROM $N = 1000$ PROGRESSIVELY TYPE-I CENSORED SAMPLES WITH $T = (2, 4, 6, 8, 10, 12)$, AND $\alpha = (0.5, 0.5, 0.5, 0.1, 0.5)$

sample size n	shape parameter γ	scale parameter σ
20	1.7094 (0.6966)	4.7516 (0.9529)
25	1.8391 (0.6638)	4.5078 (0.8200)
30	1.8740 (0.6120)	4.4073 (0.6764)
40	1.9207 (0.3104)	4.2413 (0.3557)
50	2.0120 (0.1556)	4.0137 (0.1682)
100	1.9976 (0.0700)	3.9952 (0.0186)
200	2.0043 (0.0346)	4.0055 (0.0480)
500	2.0003 (0.0141)	4.0010 (0.0245)
1000	2.0001 (0.0173)	4.0000 (0.0283)

value of T_2 that minimizes this asymptotic variance. Observe until time T_2, and calculate $\hat{\theta}^{(2)}$, the new MDE or MLE of θ.

$\vdots$

Step *i*. Calculate the asymptotic variance of $\hat{\theta}^{(i)}$ as a function of $T_1, \ldots, T_{i-1}, T_i$, and θ, replacing θ by $\hat{\theta}^{(i)}$. Then, find T_i by minimizing this asymptotic variance. Observe until time T_i to calculate $\hat{\theta}^{(i)}$, the new MDE or MLE of θ. Finally, stop when $i = m$.

1) Sequential Progressive Censoring Plans Using the MLE: We look for an optimal progressive censoring plan by using the maximum likelihood estimation method for group-censored data. Denote by σ^2 the variance of the i-stage estimator $\theta^{(i)}$ as given in (10).

$$\sigma^{-2}(T_1, \cdots, T_i, \theta) = \sum_{k=1}^{i} \left(\frac{A_k(\theta) \prod_{j=1}^{k-1}(1-\alpha_j)}{\bar{F}(T_{k-1}, \theta) - \bar{F}(T_k, \theta)} - \frac{B_k(\theta)\alpha_k \prod_{j=1}^{k-1}(1-\alpha_j)}{\bar{F}(T_k, \theta)} \right);$$

then for $M \geq i \geq 2$, we have

$$T_i = \arg\min_{T_i > T_{i-1}} \sigma^2\left(T_1, \cdots, T_{i-1}, T_i, \hat{\theta}^{(i-1)}\right) = \arg\max_{T_i > T_{i-1}} \sum_{k=1}^{i} \left(\frac{A_k\left(\hat{\theta}^{(i-1)}\right) \prod_{j=1}^{k-1}(1-\alpha_j)}{\bar{F}\left(T_{k-1}, \hat{\theta}^{(i-1)}\right) - \bar{F}\left(T_k, \hat{\theta}^{(i-1)}\right)} - \frac{B_k\left(\hat{\theta}^{(i-1)}\right) \alpha_i \prod_{j=1}^{k-1}(1-\alpha_j)}{\bar{F}\left(T_k, \hat{\theta}^{(i-1)}\right)} \right).$$

For the exponential case, upon using the s-independence between T_i and the given parameters $T_1, \ldots, T_{i-1}$, α_i, we obtain

$$T_i = \arg\min_{T_i > T_{i-1}} \frac{\bar{F}\left(T_{i-1}, \hat{\theta}^{(i-1)}\right) - \bar{F}\left(T_i, \hat{\theta}^{(i-1)}\right)}{(T_i - T_{i-1})^2 \bar{F}\left(T_i, \hat{\theta}^{(i-1)}\right)}, \quad (11)$$

for which it is easy to see that the solution is $T_i = T_{i-1} + 1.6/\hat{\theta}^{(i-1)}$. Thus, we retrieve the result of [17].

Optimal Progressive Censoring Plans Using the MDE: Because we have

$$\sqrt{n}(\hat{\lambda}_i - \lambda_i) \rightsquigarrow \mathcal{N}\left(0, \frac{\left(\bar{F}(T_{i-1}, \theta) - \bar{F}(T_i, \theta)\right) \bar{F}(T_i, \theta)}{\left(\bar{F}(T_{i-1}, \theta)\right)^3 \prod_{j=1}^{i-1}(1-\alpha_j)}\right),$$

by using the δ-method, we obtain

$$\sqrt{n}\left(\hat{\theta}^{(i)}-\theta\right) \rightsquigarrow \mathcal{N}\left(0, \frac{\left(\bar{F}(T_{i-1},\theta)-\bar{F}(T_i,\theta)\right)}{\left(\bar{F}(T_i,\theta)\frac{\partial \bar{F}(T_{i-1},\theta)}{\partial\theta}-\bar{F}(T_{i-1},\theta)\frac{\partial \bar{F}(T_i,\theta)}{\partial\theta}\right)^2} \times \frac{\bar{F}(T_i,\theta)\bar{F}(T_{i-1},\theta)}{\prod_{j=1}^{i-1}(1-\alpha_j)}\right).$$

The optimal value of the i-th censoring time is therefore equal to

$$T_i = \arg\min_{T_i>T_{i-1}} \left\{ \frac{\left(\bar{F}(T_{i-1},\theta)-\bar{F}(T_i,\theta)\right)}{\left(\bar{F}(T_i,\theta)\frac{\partial \bar{F}(T_{i-1},\theta)}{\partial\theta}-\bar{F}(T_{i-1},\theta)\frac{\partial \bar{F}(T_i,\theta)}{\partial\theta}\right)^2} \cdot \frac{\bar{F}(T_i,\theta)\bar{F}(T_{i-1},\theta)}{\prod_{j=1}^{i-1}(1-\alpha_j)} \right\}. \quad (12)$$

For an exponentially distributed sample, and for the same reason as explained above, we obtain

$$T_i = \arg\min_{T_i>T_{i-1}} \frac{\bar{F}(T_{i-1},\hat{\theta})-\bar{F}(T_i,\hat{\theta})}{(T_i-T_{i-1})^2\bar{F}(T_i,\hat{\theta})}, \quad (13)$$

which also leads to optimal sequential censoring times as $T_i = T_{i-1} + 1.6/\hat{\theta}^{(i-1)}$.

Remark 7.1: Thus, for an exponential sample, the MLE, and MDE methods result in the same optimal progressive censoring plans.}

B. Multivariate Case $\theta \in \Theta \subset \mathbb{R}^p$ With $p > 1$

Suppose $\theta \in \Theta \subset \mathbb{R}^p$. In this case, based on the variance-covariance matrix derived above, we computed the expected asymptotic variance-covariance matrix of the NP-estimators, and then determined an optimal progressive group-censoring plan based on the D-optimality criterion.

Given $T_1,\ldots,T_i$, assume that $\hat{\theta}^{(i)}$ is an estimator of θ satisfying

$$\sqrt{n}\left(\hat{\theta}^{(i)}-\theta\right) \rightsquigarrow \mathcal{N}\left(0, \Omega(T_1,\cdots,T_i,\theta)\right). \quad (14)$$

Because the volume of the asymptotic joint confidence region of θ is proportional to the determinant of the asymptotic variance-covariance matrix of $\hat{\theta}^{(i)}$, we can find T_{i+1} by choosing the value of T_{i+1} which minimizes the determinant $|\Omega(T_1,\ldots,T_i,T,\hat{\theta}^{(i)})|$. This criterion is the so-called D-optimality criterion. Therefore, we have

$$T_{i+1} = \arg\min_{T>T_i} \left|\Omega\left(T_1,\cdots,T_i,T,\hat{\theta}^{(i)}\right)\right|. \quad (15)$$

1) Optimal Progressive Group-Censoring Plan Using the Maximum Likelihood Estimator: We look for the optimal progressive censoring plan using the maximum likelihood estimator under group-censored samples. We denote by $\Omega(T_m,\theta)$ the asymptotic variance of the m-stage estimator $\theta^{(m)}$. Then, the m-th optimal censoring time can be determined as

$$T_m = \arg\max_{T_m>T_{m-1}} \sum_{i=1}^{m} \left(\frac{A_i\left(\hat{\theta}^{(m-1)}\right)\prod_{j=1}^{i-1}(1-\alpha_j)}{\bar{F}\left(T_{i-1},\hat{\theta}^{(m-1)}\right)-\bar{F}\left(T_i,\hat{\theta}^{(m-1)}\right)} + \frac{B_i\left(\hat{\theta}^{(m-1)}\right)\alpha_i\prod_{j=1}^{i-1}(1-\alpha_j)}{\bar{F}\left(T_i,\hat{\theta}^{(m-1)}\right)} \right). \quad (16)$$

2) Optimal Progressive Group-Censoring Plan Using the Minimum-Distance Estimator: In Theorem 4.4, it has been proved that the m-stage MDE estimator $\hat{\theta}^{(m)}$ is such that

$$\sqrt{n}\left(\hat{\theta}^{(m)}-\theta\right) \rightsquigarrow \mathcal{N}(0, MB\Sigma B^T M^T),$$

where the asymptotic variance-covariance matrix $MB\Sigma B^T M^T$, which depends on $(T_1,\ldots,T_m,\theta)$, is written as $\Omega(T_1,\ldots,T_m,\theta)$. So, given $T_1,\ldots,T_{m-1}$, and $\hat{\theta}^{(m-1)}$, we define the m-th optimized censoring time T_m as

$$T_m = \arg\min_{T>T_{m-1}} \left|\Omega\left(T_1,\cdots,T_{m-1},T,\hat{\theta}^{(m-1)}\right)\right|. \quad (17)$$

Therefore, an optimal multi-stage progressive Type-I censoring scheme can be determined in the following manner. Assume that the initial sample size is n, and set the censoring proportions to $\alpha_1,\ldots,\alpha_m$. The first censoring time is fixed to be T_1. We can then determine the optimal progressive censoring scheme by repeatedly using formula (17), or (16).

3) Monte Carlo Study: Suppose that the lifetimes of the units under test are Weibull distributed with scale parameter $\sigma = 4$, and shape parameter $\gamma = 2$. At the censoring time T_i, a proportion α_i of surviving units are randomly removed (censored) from the experiment. For each simulated sample, the censoring times $T_2,\ldots,T_m$ are determined by one of the above D-optimality criteria. We simulated $N = 1000$ samples, and examined the empirical behavior of the estimated T_i for various choices of the α_i, and sample sizes n.

For $m = 4$, we fixed $(\alpha_1,\alpha_2,\alpha_3)$ to be (0.1, 0.2, 0.4), and the sample size n to be $\{100, 200, 500, 1000\}$. Table VI summarizes the empirical behavior of the optimal sequential progressive censoring plan when the MDE method was used, while Table VII summarizes the empirical behavior of the optimal sequential progressive censoring plan when the MLE method was used. The "true" optimal sequential progressive censoring plans reported in these tables were calculated by using the known values of σ, and γ.

From the results presented in Tables VI and VII, we observe the following.

- The larger the proportion of censored units, the smaller are the censoring times. This relationship means that the whole

TABLE VI
OPTIMAL PROGRESSIVE CENSORING PLANS USING MINIMUM-DISTANCE ESTIMATOR FOR THE WEIBULL DISTRIBUTION ($T_1 = 4$): MEAN, AND STANDARD DEVIATION (WITHIN PARENTHESES)

sample size n	α	T_2	T_3	T_4
100	0.1	7.0639 (0.3866)	7.9694 (0.5535)	-
	0.2	7.0637 (0.3958)	-	-
	0.4	7.0322 (0.4038)	-	-
200	0.1	7.0986 (0.3593)	8.0265 (0.2353)	-
	0.2	7.0956 (0.3677)	8.0061 (0.2467)	-
	0.4	7.0949 (0.3710)	-	-
500	0.1	7.2165 (0.2817)	8.1177 (0.1767)	8.2140 (0.1415)
	0.2	7.2012 (0.2869)	8.0972 (0.1501)	8.1997 (0.1565)
	0.4	7.2009 (0.2931)	8.0855 (0.1685)	8.1956 (0.3110)
1000	0.1	7.2982 (0.2278)	8.1712 (0.0975)	8.2717 (0.2197)
	0.2	7.3010 (0.2282)	8.1703 (0.1043)	8.2709 (0.2260)
	0.4	7.2990 (0.2352)	8.1657 (0.1820)	8.2689 (0.2317)
"True" values	0.1	7.2736107	8.1507206	8.2631964
	0.2	7.2736109	8.1507179	8.2631834
	0.4	7.2736107	8.1507139	8.2631654

TABLE VII
OPTIMAL PROGRESSIVE CENSORING PLANS USING MAXIMUM LIKELIHOOD ESTIMATOR FOR THE WEIBULL DISTRIBUTION ($T_1 = 4$): MEAN, AND STANDARD DEVIATION (WITHIN PARENTHESES)

sample size n	α	T_2	T_3	T_4
100	0.1	7.3090 (0.4465)	-	-
	0.2	7.2718 (0.4298)	-	-
	0.4	7.2315 (0.4299)	-	-
200	0.1	7.3246 (0.3737)	9.1547 (0.4547)	10.5978 (0.2468)
	0.2	7.2820 (0.3552)	9.1256 (0.4604)	10.5831 (0.1732)
	0.4	7.2331 (0.3575)	9.0583 (0.5467)	10.5281 (0.1546)
500	0.1	7.4118 (0.3302)	9.2158 (0.2834)	10.6188 (0.2236)
	0.2	7.3988 (0.3197)	9.1946 (0.2846)	10.5994 (0.2060)
	0.4	7.3665 (0.3060)	9.1630 (0.2266)	10.5654 (0.1861)
1000	0.1	7.5127 (0.2386)	9.2790 (0.1872)	10.6789 (0.1582)
	0.2	7.4825 (0.2557)	9.2533 (0.1965)	10.6490 (0.1688)
	0.4	7.4423 (0.2320)	9.2205 (0.2279)	10.6164 (0.1981)
"True" values	0.1	7.4814307	9.2603923	10.667217
	0.2	7.4627656	9.2411046	10.648380
	0.4	7.4225328	9.1999080	10.608379

TABLE VIII
ESTIMATES OF THE WEIBULL PARAMETERS FOR THE WARRANTY DATA IN TABLE I

θ	$\theta^{(1)}$	$\theta^{(2)}$	$\theta^{(3)}$	OWSBE	MLE
$\hat{\sigma}$	9.958	11.228	9.928	10.075	10.332
$\hat{\gamma}$	2.939	2.671	3.110	2.917	2.879

experiment is shorter when the censoring proportions increase.

- Let us denote $\Delta T_i = T_i - T_{i-1}$. Then, we observe that $\Delta T_i > \Delta T_{i+1}$. This relationship means that the widths of the time intervals decrease as the number of censoring stages increase.
- For fixed α_i, a larger sample size corresponds to a larger censoring time T_i at the ith stage. This relationship shows that the test requires shorter time for a smaller sample size than for a larger sample size.
- Finally, note that for moderate sample sizes, if T_1 is too large, then it may happen that there is not sufficient data to estimate properly the optimal value of the next censoring time (this is indeed the case in Table VI when $n = 100$, or $n = 200$).

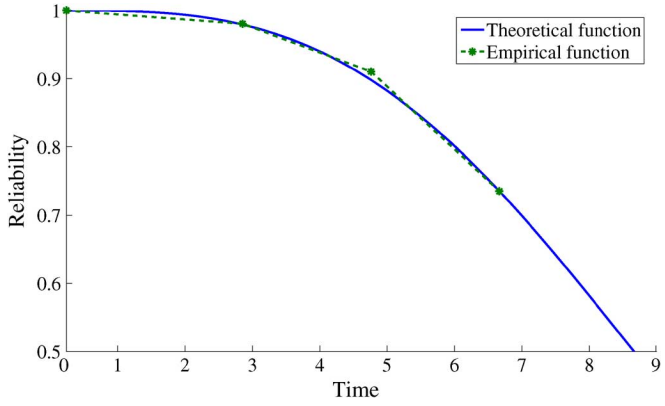

Fig. 2. Empirical (dotted) versus theoretical (plain) reliability function of the Weibull distribution.

VIII. ILLUSTRATIVE EXAMPLE

Let us reconsider the warranty data in Table I. Table VIII summarizes the minimum distance estimates based on k-stage progressive censoring schemes for $k = 1, 2, 3$; and the pooled OWSBE, and the MLEs of the Weibull parameters. We see that both methods provide quite close estimates. Fig. 2 presents a plot of the Weibull reliability function with $\sigma = 10$, and $\gamma = 3$

(a Weibull model that is closely reflected by the estimated parameters in Table VIII) against the empirical distribution function obtained from the data in Table I.

IX. Conclusions

In this paper, we have discussed the parametric estimation problem under a multi-stage progressive Type-I censoring scheme. A nonparametric Kaplan-Meier type estimator that uses partial information from progressively censored samples enables the estimation of the reliability function at the censoring times. A minimum-distance method has been proposed to estimate the unknown Euclidean parameter of a parametric model. We have also proposed a minimum variance linear estimator for deriving an asymptotically optimal estimator based on m estimators of θ. We have then discussed the asymptotic properties of all these estimators, and have presented a simple algorithm for the determination of an optimal sequential progressive censoring plan. Finally, we have evaluated the performance of all these estimators by means of Monte Carlo simulations, and have illustrated the proposed methods applied to warranty data.

Appendix

Proof of Proposition 3.4: Let us calculate the compensator $A = (A_1, \ldots, A_m)$ by induction. First, calculate

$$\mathbb{E}(K_{i+1} - K_i|\mathcal{F}_i) = \mathbb{E}(N_{i+1}|\mathcal{F}_i) = Y_i\lambda_{i+1}.$$

Then, by using the fact that A_{i+1} is predictable, i.e., $\mathcal{F}_i$-measurable, write

$$\mathbb{E}(L_{i+1} - L_i|\mathcal{F}_i) = 0 = \mathbb{E}(N_{i+1}|\mathcal{F}_i) - A_{i+1} + A_i \quad \text{a.s.}$$

to obtain the compensator A of L by induction. □

Proof of Proposition 3.5: Because, conditional on $\mathcal{F}_{i-1}$, N_i is a binomial random variable with parameters Y_{i-1}, and λ_i, we have $\mathbb{E}((N_i - Y_{i-1}\lambda_i)^2|\mathcal{F}_{i-1}) = Y_{i-1}\lambda_i(1-\lambda_i)$ almost surely, and so we obtain for $1 \le i \le m$, almost surely,

$$\mathbb{E}(H_i|\mathcal{F}_{i-1}) = H_{i-1}.$$

The fact that $\mathbb{E}(H_i) < \infty$ for $1 \le i \le m$ completes the proof. □

Proof of Proposition 3.6: For $1 \le i \le m$, we have $\mathbb{E}(M_i) = 0$. Moreover, as $n \to \infty$,

$$\begin{aligned}\mathbb{E}\left[M_i^2\right] &= \mathbb{E}\left(\mathbb{E}\left(M_i^2|\mathcal{F}_{i-1}\right)\right)\\ &= \frac{1}{n^2}\sum_{j=1}^{i}\mathbb{E}(Y_{j-1})\lambda_j(1-\lambda_j)\\ &\le \frac{1}{n}\sum_{j=1}^{i}\lambda_j(1-\lambda_j) \to 0,\end{aligned}$$

and so we have, for $1 \le i \le m$,

$$\gamma_i = \beta_{i-1}\lambda_i. \tag{18}$$

Now, upon noting that

$$Y_i = Y_{i-1} - N_i - R_i = Y_{i-1} - N_i - \lfloor \alpha_i(Y_{i-1} - N_i)\rfloor,$$

dividing the above equality by n, and then letting n tend to ∞, we obtain

$$\begin{aligned}\beta_i &= (1-\alpha_i)(\beta_{i-1} - \gamma_i) = (1-\alpha_i)\beta_{i-1}(1-\lambda_i)\\ &= (1-\alpha_i)\beta_{i-1}\frac{\bar{F}_i}{\bar{F}_{i-1}}\end{aligned} \tag{19}$$

for $1 \le i \le m$, with $\alpha_0 = 1$. The required formula follows by using (18), (19), and the fact that $\beta_0 = 1$. □

Proof of Corollary 3.7: Writing $\hat{\lambda}_i = (N_i/n)/(Y_{i-1}/n)$, we have $\hat{\lambda}_i \to \gamma_i/\beta_{i-1} = \lambda_i$ in probability, which proves (i). The results in (ii) and (iii) then follow immediately by using the relationships between the estimators $\hat{\Lambda}_i$ and $\hat{\bar{F}}_i$, and the $\hat{\lambda}_i$. □

Proof of Theorem 3.8: Note that $\sqrt{n}(\hat{\lambda} - \lambda) = \mathrm{diag}(n/Y_{i-1})M^T/\sqrt{n} + o_P(1)$. Applying the central limit theorem for a triangular array of martingale difference (see, for example, [2]), we have $M^T/\sqrt{n}$ to be asymptotically normal with zero mean, and a covariance matrix with entries

$$\sigma_{ij} = \begin{cases}\lambda_i(1-\lambda_i)\beta_{i-1}, & i = j\\ 0, & i \ne j.\end{cases}$$

The diagonal matrix $\mathrm{diag}(n/Y_{i-1})$ converges in probability to $\mathrm{diag}(1/\beta_{i-1})$ by Proposition 3.6, and so by Slutsky's Lemma, we obtain the expected weak convergence result. The consistency of $\hat{\Sigma}$ is straightforward upon using Proposition 3.6, and Corollary 3.7. □

Proof of Lemma 4.2:

Proof of (i): Note that

$$\begin{aligned}&\left|\hat{\phi}_n(\theta) - \phi(\theta)\right|\\ &\quad= \left|\sum_{i=1}^{m}\left(\bar{F}(T_i,\theta) - \hat{\bar{F}}_i\right)^2 - \left(\bar{F}(T_i,\theta) - \bar{F}(T_i,\theta_0)\right)^2\right|\\ &\quad= \left|\sum_{i=1}^{m}\left(\bar{F}(T_i,\theta) - \hat{\bar{F}}_i + \bar{F}(T_i,\theta) - \bar{F}(T_i,\theta_0)\right)\right.\\ &\qquad\left.\times\left(\bar{F}(T_i,\theta_0) - \hat{\bar{F}}\right)\right|\\ &\quad\le 4\sum_{i=1}^{m}\left|\left(\bar{F}(T_i,\theta_0) - \hat{\bar{F}}_i\right)\right|.\end{aligned}$$

Moreover, by Corollary 3.9, we have $\sqrt{n}(\bar{F}(T_i,\theta_0) - \hat{\bar{F}}_i) \rightsquigarrow \mathcal{N}(0,\gamma_{ii})$, so that $\bar{F}(T_i,\theta_0) - \hat{\bar{F}} = O_P(n^{-1/2})$. Hence,

$$\sup_{\theta\in\Theta}\left|\hat{\phi}_n(\theta) - \phi(\theta)\right| = O_P(n^{-1/2}).$$

Proof of (ii): We have

$$\left\|\frac{\partial\hat{\phi}_n(\theta)}{\partial\theta}-\frac{\partial\phi(\theta)}{\partial\theta}\right\| \leq \max_{1\leq i\leq m}\left\|\bar{F}(T_i,\theta_0)-\hat{\bar{F}}\right\| \max_{1\leq i\leq m}\sup_{\theta\in\Theta}\left\|\frac{\partial\bar{F}(T_i,\theta)}{\partial\theta}\right\|.$$

Because $\theta\mapsto\bar{F}(T_i,\theta)$ belongs to $C^2(\Theta)$, and that Θ is compact, we have

$$\max_{1\leq i\leq m}\sup_{\theta\in\Theta}\left\|\frac{\partial\bar{F}(T_i,\theta)}{\partial\theta}\right\|<\infty$$

which, with Corollary 3.9, yields

$$\sup_{\theta\in\Theta}\left\|\frac{\partial\hat{\phi}_n(\theta)}{\partial\theta}-\frac{\partial\phi(\theta)}{\partial\theta}\right\|=O_P(n^{-1/2}).$$

Proof of (iii): Write

$$\frac{\partial^2\hat{\phi}_n(\theta)}{\partial\theta\partial\theta^T}-\frac{\partial^2\phi(\theta)}{\partial\theta\partial\theta^T}=2\sum_{i=1}^{m}\left(F(T_i,\theta_0)-\hat{\bar{F}}_i)\right)\times\frac{\partial^2\bar{F}(T_i,\theta)}{\partial\theta\partial\theta^T},$$

and then use the same line of proof as above for (ii). □

Proof of Theorem 4.3: By the first part of Lemma 4.2, we have $\sup_{\theta\in\Theta}|\hat{\phi}_n(\theta)-\phi(\theta)|=O_P(n^{-1/2})$. Let $\varepsilon>0$ be a real number, and $B_0=\{\theta\in\Theta:\|\theta-\theta_0\|<\varepsilon\}$. Then, we have

$$\{\hat{\theta}\notin B_0\}\subset\left\{\inf_{\theta\in\Theta\setminus B_0}\hat{\phi}_n(\theta)<\hat{\phi}_n(\theta_0)\right\}$$

and if $\theta\in\Theta\setminus B_0$, we can write

$$\begin{aligned}\hat{\phi}_n(\theta)&=\hat{\phi}_n(\theta)-\phi(\theta)+\phi(\theta)\\&\geq\hat{\phi}_n(\theta)-\phi(\theta)+\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)\\&\geq-\sup_{\theta\in\Theta}\left|\hat{\phi}_n(\theta)-\phi(\theta)\right|+\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)\end{aligned}$$

so that

$$\inf_{\theta\in\Theta\setminus B_0}\hat{\phi}_n(\theta)>\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)-\sup_{\theta\in\Theta}\left|\hat{\phi}_n(\theta)-\phi(\theta)\right|.$$

Therefore,

$$\begin{aligned}\{\hat{\theta}\notin B_0\}&\subset\left\{\inf_{\theta\in\Theta\setminus B_0}\hat{\phi}_n(\theta)<\hat{\phi}_n(\theta_0)\right\}\\&\subset\left\{\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)-\sup_{\theta\in\Theta}\left|\hat{\phi}_n(\theta)-\phi(\theta)\right|<\hat{\phi}_n(\theta_0)\right\}\\&\subset\left\{\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)-\sup_{\theta\in\Theta}\left|\hat{\phi}_n(\theta)-\phi(\theta)\right|\right.\\&\qquad\left.<\left|\hat{\phi}_n(\theta_0)-\phi(\theta_0)\right|\right\}\\&\subset\left\{\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)<2\sup_{\theta\in\Theta}\left|\hat{\phi}_n(\theta)-\phi(\theta)\right|\right\}.\end{aligned}$$

As a consequence, we have

$$P(\hat{\theta}\notin B_0)\leq P\left(\sup_{\theta\in\Theta}\left|\hat{\phi}_n(\theta)-\phi(\theta)\right|>\frac{1}{2}\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)\right).$$

From the fact that $\sup_{\theta\in\Theta}|\hat{\phi}_n(\theta)-\phi(\theta)|=O_P(n^{-1/2})$, and that $\inf_{\theta\in\Theta\setminus B_0}\phi(\theta)=\phi(\bar{\theta})>0$, we have $\lim_{n\to+\infty}P\{\hat{\theta}\notin B_0\}=0$, and so $\hat{\theta}\xrightarrow{P}\theta$. □

Proof of Theorem 4.4: First, we have

$$\frac{\partial\hat{\phi}_n(\hat{\theta})}{\partial\theta}-\frac{\partial\hat{\phi}_n(\theta_0)}{\partial\theta}=-\frac{\partial\hat{\phi}_n(\theta_0)}{\partial\theta}=-2Q(\theta)\left(\bar{F}(\theta_0)-\hat{\bar{F}}\right).$$

On the other hand, we can also write

$$\frac{\partial\hat{\phi}_n(\hat{\theta})}{\partial\theta}-\frac{\partial\hat{\phi}_n(\theta_0)}{\partial\theta}=-\frac{\partial\hat{\phi}_n(\theta_0)}{\partial\theta}=\frac{\partial^2\hat{\phi}_n(\theta^*)}{\partial\theta\partial\theta^T}(\hat{\theta}-\theta_0),$$

and so it follows that

$$\frac{\partial^2\hat{\phi}_n(\theta^*)}{\partial\theta\partial\theta^T}\sqrt{n}(\hat{\theta}-\theta_0)=-2Q(\theta_0)\sqrt{n}\left(\bar{F}(\theta_0)-\hat{\bar{F}}\right),\quad(20)$$

where $\bar{F}(\theta)=(\bar{F}(T_1,\theta),\ldots,\bar{F}(T_m,\theta))^T$.

Let us now introduce

$$\hat{I}_n=\frac{\partial^2\hat{\phi}_n(\theta^*)}{\partial\theta\partial\theta^T},\quad\text{and}\quad I^*=\frac{\partial^2\phi(\theta^*)}{\partial\theta\partial\theta^T}.$$

Formula (20) can then be rewritten as

$$\left((\hat{I}_n-I^*)+I^*\right)\sqrt{n}(\hat{\theta}-\theta_0)=Q(\theta_0)\sqrt{n}(\bar{F}-\hat{\bar{F}}).\quad(21)$$

Because $\hat{\theta}$ is consistent, using the third part of Lemma 4.2, we obtain

$$(\hat{I}_n-I^*)\sqrt{n}(\hat{\theta}-\theta_0)=o_P(1)$$

so (21) becomes

$$I^*\sqrt{n}(\hat{\theta}-\theta_0)=Q(\theta_0)\sqrt{n}(\bar{F}-\hat{\bar{F}})+o_P(1).\quad(22)$$

Let λ_n^* be the minimum eigen value of I^*, and λ_0 be the minimum eigen value of I_0. Because $I^*\xrightarrow{P}I_0$, we have $\lambda_n^*\xrightarrow{P}\lambda_0>0$. Now, we have

$$\left\|I^*\sqrt{n}(\hat{\theta}_n-\theta_0)\right\|^2\geq(\lambda_n^*)^2)\left\|\sqrt{n}(\hat{\theta}_n-\theta_0)\right\|^2.$$

For a positive real number A, we have

$$\begin{aligned}&P\left(\left\|\sqrt{n}(\hat{\theta}_n-\theta_0)\right\|>A\right)\\&\quad\leq P\left(\left\|\sqrt{n}(\hat{\theta}_n-\theta_0)\right\|>A;\lambda_n^*>\lambda_0/2\right)+P\left(\lambda_n^*\leq\lambda_0/2\right)\\&\quad\leq P\left(\left\|I^*\sqrt{n}(\hat{\theta}_n-\theta_0)\right\|>A\lambda_0/2\right)+P\left(\lambda_n^*\leq\lambda_0/2\right).\end{aligned}$$

From the fact that $P(\lambda_n^*\leq\lambda_0/2)\to0$, and that $P(\|I^*\sqrt{n}(\hat{\theta}_n-\theta_0)\|>A\lambda_0/2)$ may be arbitrarily small by choosing A large enough, the above inequality leads to

$$\sqrt{n}(\hat{\theta}_n-\theta_0)=O_P(1).\quad(23)$$

Formula (20) becomes

$$\left((\hat{I}_n - I^*) + (I^* - I_0) + I_0\right)\sqrt{n}(\hat{\theta} - \theta_0) = Q(\theta_0)\sqrt{n}(\bar{F} - \hat{\bar{F}}). \tag{24}$$

Upon using the fact that $I^* \to I_0$, the third part of Lemma 4.2, and formula (23), we see that (24) implies that

$$I_0\sqrt{n}(\hat{\theta} - \theta_0) = Q(\theta_0)\sqrt{n}(\bar{F} - \hat{\bar{F}}) + o_P(1).$$

Therefore,

$$\sqrt{n}(\hat{\theta} - \theta_0) = M\sqrt{n}(\bar{F} - \hat{\bar{F}}) + o_P(1)$$

which, with Corollary 3.9, and Assumption (A3), proves the central limit result. The consistency of $\hat{M}$ then follows from Assumptions (A2), and (A3). □

ACKNOWLEDGMENT

The authors express their sincere thanks to the Managing Editor Dr. J. Rupe, the Associate Editor, and two anonymous Referees for their constructive criticisms and useful suggestions which led to a considerable improvement over an earlier version of this manuscript.

REFERENCES

[1] N. Balakrishnan and A. Cohen, *Order Statistics and Inference: Estimation Methods*. Boston: Academic Press, 1991.
[2] T. Fleming and D. Harrington, *Counting Processes and Survival Analysis*. New York: Wiley, 1991.
[3] P. Andersen, Ø. Borgan, R. Gill, and N. Keiding, *Statistical Models Based on Counting Processes*. New York: Springer-Verlag, 1993.
[4] J. Klein and M. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer-Verlag, 1997.
[5] N. Balakrishnan and R. Aggarwala, *Progressive Censoring: Theory, Methods, and Applications*. Boston: Birkhäuser, 2000.
[6] L. Bordes, "Non-parametric estimation under progressive censoring," *J. Statist. Plann. Inf.*, vol. 119, pp. 171–189, 2004.
[7] S. Alvarez-Andrade, N. Balakrishnan, and L. Bordes, "Proportional hazards regression under progressive Type–II censoring," *Ann. Inst. Statist. Math.*, vol. 61, no. 4, pp. 887–903, 2009.
[8] N. Balakrishnan, S. Alvarez-Andrade, and L. Bordes, "Homogeneity tests based on several progressively Type–II censored samples," *J. Multivar. Anal.*, vol. 98, no. 6, pp. 1195–1213, 2007.
[9] N. Balakrishnan, "Progressive censoring methodology: An appraisal (with discussions)," *Test*, vol. 16, pp. 211–296, 2007.
[10] N. Balakrishnan and D. Han, "Optimal progressive Type–II censoring schemes for non-parametric confidence intervals of quantiles," *Comm. Statist. Simul. Comput.*, vol. 36, no. 6, pp. 1247–1262, 2007.
[11] M. Burke, Nonparametric estimation of a survival function under progressive Type–I multistage censoring The University of Calgary, 2004, Research Paper.
[12] E. Gouno, A. Sen, and N. Balakrishnan, "Optimal step-stress test under progressive Type–I censoring," *IEEE Trans. Reliab.*, vol. 53, no. 3, pp. 388–393, 2004.
[13] D. Han, N. Balakrishnan, A. Sen, and E. Gouno, "Corrections on optimal step-stress test under progressive Type–I censoring," *IEEE Trans. Reliab.*, vol. 55, pp. 613–614, 2006.
[14] N. Balakrishnan and D. Han, "Optimal step-stress testing for progressively Type-I censored data from exponential distribution," *J. Statist. Plann. Inf.*, vol. 139, pp. 1782–1798, 2009.
[15] N. Balakrishnan and Q. Xie, "Exact inference for a simple step-stress model with Type–II hybrid censored data from the exponential distribution," *J. Statist. Plann. Inf.*, vol. 8, no. 137, pp. 2543–2563, 2007.
[16] N. Balakrishnan and Q. Xie, "Exact inference for a simple step-stress model with Type–I hybrid censored data from the exponential distribution," *J. Statist. Plann. Inf.*, vol. 137, no. 11, pp. 3268–3290, 2007.
[17] M. N. Patel and A. V. Gajjar, "Some results on maximum likelihood estimators of parameters of exponential distribution under Type–I progressive censoring with changing failure rates," *Comm. Statist. Theo. Meth.*, vol. 24, no. 9, pp. 2421–2435, 1995.
[18] D. J. Bartholomew, "The sampling distribution of an estimate arising in life testing," *Technometrics*, vol. 5, no. 3, pp. 361–374, 1963.
[19] P. Kendell and R. Anderson, "An estimation problem in life testing," *Technometrics*, vol. 13, pp. 289–309, 1971.

Narayanaswamy Balakrishnan (M'08) is a Professor of Statistics at McMaster University, Hamilton, Ontario, Canada. He received his B.Sc., and M.Sc. degrees in Statistics from University of Madras, India, in 1976, and 1978, respectively. He finished his Ph.D. in Statistics from Indian Institute of Technology, Kanpur, India, in 1981. He is a Fellow of the American Statistical Association, and a Fellow of the Institute of Mathematical Statistics. He is currently the Editor-in-Chief of Communications in Statistics, and Executive Editor of Journal of Statistical Planning and Inference. His research interests include distribution theory, ordered data analysis, censoring methodology, reliability, survival analysis, and statistical quality control.

Laurent Bordes is a Professor of Statistics at the University of Pau, France. He finished his Ph.D. in Statistics from Bordeaux University, Bordeaux, France, in 1996. He was an Associate Professor at Compiegne University of Technology, France, from 1997 to 2006. He is currently the head of the department of mathematics at Pau University. His research interests include semi-/non-parametric asymptotic theory, missing data, mixture of distributions, reliability, survival analysis, and degradation models.

Xuejing Zhao works at the School of Mathematics and Statistics at Lanzhou University, P. R. China. He finished his Ph.D. in Systems Optimization and Security from University of Technology of Troyes (France) in 2009. His research interests include censoring methodology, reliability and system safety, maintenance optimization, survival analysis, and actuarial mathematics.