Ann Inst Stat Math (2018) 70:583–604
https://doi.org/10.1007/s10463-017-0598-9

Statistical inferences based on INID progressively type II censored order statistics

M. Razmkhah[1] · **S. Simriz**[1]

Received: 24 April 2015 / Revised: 4 January 2017 / Published online: 28 February 2017

Abstract Suppose that the failure times of the units placed on a life-testing experiment are independent but nonidentically distributed random variables. Under progressively type II censoring scheme, distributional properties of the proposed random variables are presented and some inferences are made. Assuming that the random variables come from a proportional hazard rate model, the formulas are simplified and also the amount of Fisher information about the common parameters of this family is calculated. The results are also extended to a fixed covariates model. The performance of the proposed procedure is investigated via a real data set. Some numerical computations are also presented to study the effect of the proportionality rates in view of the Fisher information criterion. Finally, some concluding remarks are stated.

Keywords Fisher information · Maximum likelihood estimator · Cramer–Rao lower bound · Proportional hazard rate family · Exponential family · Weibull distribution · Fixed covariates model

1 Introduction

Censored sampling arises in a life test whenever the experimenter does not observe the lifetimes of all experimental units. The model of progressive type II censoring is of importance in the field of reliability and life testing. In this censoring scheme, n units are simultaneously placed on a lifetime test, and when the ith failure time occurs, R_i surviving units are randomly censored from the experiment, $1 \leq i \leq m$.

✉ M. Razmkhah
razmkhah_m@um.ac.ir

[1] Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, P. O. Box 1159, Mashhad 91775, Iran

Thus, if m failure times are observed, then $R_1 + \cdots + R_m$ units are censored; here, $\mathbf{R} = (R_1, \ldots, R_m)$ denotes the progressive censoring plan. In the special case of $R_1 = \cdots = R_{m-1} = 0$ and $R_m = n - m$, the progressive censoring scheme coincides with the type II censoring scheme. Statistical inferences based on progressively type II censored order statistics in the case of independent and identically distributed (IID) random variables have been extensively investigated by several authors. Balakrishnan et al. (2001) studied the point and interval estimation for both location and scale parameters of the two-parameter exponential distribution based on progressively type II censored samples. Burkschat et al. (2006) investigated the optimal plans in the model of progressive type II censoring for a location-scale family of distributions. In statistical inferences, the Fisher information (FI) plays an important role in the estimation problem of unknown parameters through the Cramer–Rao inequality and its association with the asymptotic properties of the maximum likelihood estimator (MLE). Under certain regularity conditions, the FI about the real parameter θ contained in a random variable X with probability density function (pdf) $f(x;\theta)$ is defined by $I_X(\theta) = -\mathrm{E}\big(\frac{\partial^2}{\partial\theta^2} \log f(X;\theta)\big)$; see, for example, Lehmann and Casella (1998, p. 116). Zheng and Park (2004) expressed the FI contained in the progressively type II censored order statistics as a summation of a single integral involving the hazard rate function. They also obtained a closed form for the FI about the scale parameter of the exponential and Weibull distributions with equal removal at each stage. Then, Abo-Eleneen (2008) proposed an indirect approach for computing the FI in these statistics that simplified the calculations. Balakrishnan et al. (2008) determined optimal plans for a variety of lifetime distributions by employing maximum FI as an optimality criterion. For a detailed description of the IID case of progressively type II censored samples, one may refer to the books by Balakrishnan and Aggarwala (2000) and Balakrishnan and Cramer (2014). Furthermore, an overview of various developments that have been considered about the properties of progressively type II censored order statistics and inferential procedures based on them is provided by Balakrishnan (2007). The author also suggested some potential problems of interest for further research.

The model of progressive type II censoring was generalized by Balakrishnan and Cramer (2008) to the case of independent and nonidentical distributed (INID) random variables. They developed the basic distribution theory for order statistics in this case. Also, Fischer et al. (2008) studied a mixture representation for the joint distribution function of progressively type II censored order statistics from heterogeneous distributions and illustrated the applications of this representation to stochastic orderings and inequalities. Cramer and Lenz (2010) and Mao and Hu (2010) investigated the positive association and the stochastic properties of these statistics, respectively. Recently, Rezapour et al. (2013) investigated some more properties of progressively type II censored order statistics in the INID case.

Although the joint density function of the INID progressively type II censored order statistics was expressed by Balakrishnan and Cramer (2008), we are not aware about any investigations about the problem of parameter estimation or computing the information measures in this case. It seems that the main reason is the functional form of the joint density function of these statistics which has been expressed as a summation over all permutations of $\{1, \ldots, n\}$, which cannot be used to make any inferences about the common parameters of various distributions. In this paper, we first consider a vector of random variables including the INID progressively type II censored order statistics and then derive an explicit expression for the corresponding likelihood function which may be applied to make inference about the parameters of interest.

The rest of this paper is as follows. In Sect. 2, the model of interest is described. In order to identify the failed units and those that are removed from the experiment, two indicator variables are defined and based on them the appropriate likelihood function for the parameter of interest is derived. The probability functions of the INID progressively type II censored order statistics are presented in Sect. 3. Section 4 focuses on INID random variables coming from a proportional hazard rate family; the probability functions are simplified and the amount of FI about the common parameters of interest is derived. More details are presented when the exponential family and Weibull distribution are the baseline distributions of this model. Some results are extended to a fixed covariates model in Sect. 5. The performance of the proposed procedure in the paper is investigated via a real data set in Sect. 6. In Sect. 7, the FI contained in the INID progressively type II censored order statistics about the common parameter of a proportional hazard rate family with the baseline Weibull distribution is numerically computed. Some concluding remarks are presented in Sect. 8.

2 Model description

Let $X_1, \ldots, X_n$ be the lifetimes of n units which are independently and simultaneously placed on a test for which X_r comes from cumulative distribution function (cdf) $F_r(x; \boldsymbol{\theta})$ with corresponding pdf $f_r(x; \boldsymbol{\theta})$, $1 \leq r \leq n$, where $\boldsymbol{\theta}' = (\theta_1, \ldots, \theta_t)$ is the common vector of parameters of the various distributions. Moreover, let $\mathbf{R} = (R_1, \ldots, R_m)$ be the progressive censoring plan with $n = m + \sum_{i=1}^{m} R_i$. For brevity, we denote by $\gamma_j = n - \sum_{i=1}^{j-1} R_i - j + 1$ the number of units remaining in the experiment before the jth failure time.

To obtain the likelihood function of the parameters of interest $\boldsymbol{\theta}$ on the basis of the INID progressively type II censored order statistics, denoted by $X^{\mathbf{R}}_{1:m:n}, \ldots, X^{\mathbf{R}}_{m:m:n}$, let us first define the following random variable to identify the failed units on a lifetime test

$$\Delta_i^{(j)} = \begin{cases} 1, & \text{if the lifetime of the } i\text{th unit coincides with the } j\text{th failure time,} \\ 0, & \text{otherwise,} \end{cases}$$

where for each $j = 1, \ldots, m$, $\sum_{i=1}^{n} \Delta_i^{(j)} = 1$. Also, after the jth failure time, R_j of surviving units are removed from the experiment. Therefore, we use the following random variable to specify the units removed from the test

$$H_i^{(j)} = \begin{cases} 1, & \text{if the } i\text{th unit is removed from the test after the } j\text{th failure time,} \\ 0, & \text{otherwise,} \end{cases}$$

such that for each $j = 1, \ldots, m$, $\sum_{i=1}^{n} H_i^{(j)} = R_j$. Notice that for a fixed j, say j_0, if $\Delta_i^{(j_0)} = 1$ or $H_i^{(j_0)} = 1$, then for other values of j ($j \neq j_0$), $\Delta_i^{(j)} = 0$ and $H_i^{(j)} = 0$.

According to the above random variables, it is reasonable to use the following random vector to make inferences about the parameters of interest

$$\mathbf{B} = \left\{ X_{j:m:n}^{\mathbf{R}}, \Delta_i^{(j)}, H_i^{(j)}, 1 \leq j \leq m, 1 \leq i \leq n \right\}.$$

To determine the likelihood function, note that in progressive type II censoring scheme, n units are simultaneously placed on test, and at the first failure time, R_1 of $n-1$ remaining units are randomly removed from the experiment, so there is an integer s_1 ($1 \leq s_1 \leq n$) such that $\Delta_{s_1}^{(1)} = 1$, and also there are R_1 integers $i_1, \ldots, i_{R_1}$, not equal to s_1, for which $H_{i_j}^{(1)} = 1$ ($1 \leq j \leq R_1$). At the first failure time of the remaining $n - R_1 - 1$ units, R_2 units are randomly censored; hence, there is another integer s_2 ($s_2 \in \{1, \ldots, n\} \backslash \{s_1, i_1, \ldots, i_{R_1}\}$) such that $\Delta_{s_2}^{(2)} = 1$, and also there are R_2 integers, say $i_{R_1+1}, \ldots, i_{R_1+R_2}$, which belong to $\{1, \ldots, n\} \backslash \{s_1, i_1, \ldots, i_{R_1}, s_2\}$ for which $H_{i_j}^{(2)} = 1$ ($R_1 + 1 \leq j \leq R_1 + R_2$). This procedure continues to arrive at the mth failure time in a sample of size n. It is clear that the lifetimes of the removed units from the test after the jth failure time are greater than the jth failure time. Therefore, by taking into account the above scenario, the likelihood function of $\boldsymbol{\theta}$ on the basis of the data set $\mathbf{B}$ can be presented as

$$L(\boldsymbol{\theta}) = \left(\prod_{j=1}^{m} \gamma_j \right) \prod_{j=1}^{m} \left\{ \prod_{i=1}^{n} [f_i(x_j; \boldsymbol{\theta})]^{\delta_i^{(j)}} [\bar{F}_i(x_j; \boldsymbol{\theta})]^{\eta_i^{(j)}} \right\}, \tag{1}$$

where x_j, $\delta_i^{(j)}$ and $\eta_i^{(j)}$ are the observed values of $X_{j:m:n}^{\mathbf{R}}$, $\Delta_i^{(j)}$ and $H_i^{(j)}$, respectively. Notice that the constant γ_j represents the number of ways in which the jth progressively type II censored order statistic may occur.

Remark 1 The likelihood function of $\boldsymbol{\theta}$ in (1) can be used for the following cases:

- Let $X_1, \ldots, X_n$ be INID random variables with k ($k < n$) different distributions such that n_j of them come from the cdf $F_j(x; \boldsymbol{\theta})$, $1 \leq j \leq k$, for which $n = \sum_{j=1}^{k} n_j$; then, the likelihood function of $\boldsymbol{\theta}$ in (1) can also be applied to make inference about $\boldsymbol{\theta}$.
- In the case of usual order statistics, we have $m = n$ as well as $R_j = 0$. In this case, $\gamma_j = n - j + 1$ and $H_i^{(j)} = 0$, for each $1 \leq i \leq n$ and $1 \leq j \leq n$. Therefore, using (1), the likelihood function of $\boldsymbol{\theta}$ can be presented as

$$L(\boldsymbol{\theta}) = n! \prod_{j=1}^{n} \prod_{i=1}^{n} [f_i(x_j; \boldsymbol{\theta})]^{\delta_i^{(j)}}.$$

- Since for each $j = 1, \ldots, m$, $\sum_{i=1}^{n} \Delta_i^{(j)} = 1$ and $\sum_{i=1}^{n} H_i^{(j)} = R_j$, the likelihood function of $\boldsymbol{\theta}$ in (1) converts to the joint density function of progressively type II censored order statistics in the case of IID random variables; see, for example, Balakrishnan and Aggarwala (2000).

3 Joint probability functions

In this section, the probability of the events $\{X^{\mathbf{R}}_{k:m:n} > y, \Delta_i^{(k)} = 1\}$ and $\{X^{\mathbf{R}}_{k:m:n} > y, H_i^{(k)} = 1\}$ is presented in Theorems 1 and 2, respectively, which will be used in the next sections and also are useful in the other statistical inferences. Accordingly, the joint probability functions of $(X^{\mathbf{R}}_{k:m:n}, \Delta_i^{(k)})$ and $(X^{\mathbf{R}}_{k:m:n}, H_i^{(k)})$, the marginal pdf of $X^{\mathbf{R}}_{k:m:n}$ and also the probability of the events $\{\Delta_i^{(k)} = 1\}$ and $\{H_i^{(k)} = 1\}$ may be derived.

Theorem 1 *For each $i = 1, \ldots, n$, $k = 1, \ldots, m$ and every positive real value y, we get*

$$
\begin{aligned}
Pr\left(X^{\mathbf{R}}_{k:m:n} > y, \Delta_i^{(k)} = 1\right) &= \left\{\prod_{j=1}^{k-1} \binom{\gamma_j - 1}{R_j}\right\}^{-1} \\
&\times \sum_{E_i^{(k-1)}} \int_y^{\infty} \int_0^{x} \int_0^{x_{k-1}} \cdots \int_0^{x_2} \prod_{j=1}^{k-1} \left(f_{s_j}(x_j; \boldsymbol{\theta}) \prod_{t \in C^{(j)}} \bar{F}_t(x_j; \boldsymbol{\theta}) \right) \\
&\times f_i(x; \boldsymbol{\theta}) \left(\prod_{t \in \Lambda_i^{(k-1)}} \bar{F}_t(x; \boldsymbol{\theta}) \right) dx_1 \ldots dx_{k-2} dx_{k-1} dx,
\end{aligned} \tag{2}
$$

where the summation index $E_i^{(k)}$ extends over all permutations $(s_1, \ldots, s_k, i_1, \ldots, i_{\sum_{r=1}^{k} R_r})$ of the integers $\{1, \ldots, n\} \backslash \{i\}$ for which for $1 \le j \le k$, $i_{\sum_{r=1}^{j-1} R_r + 1} < \cdots < i_{\sum_{r=1}^{j} R_r}$; further,

$$
C^{(j)} = \left\{ i_{\sum_{r=1}^{j-1} R_r + 1}, \ldots, i_{\sum_{r=1}^{j} R_r} \right\}, \tag{3}
$$

$$
\Lambda_i^{(k)} = \{1, \ldots, n\} \backslash \{s_1, \ldots, s_k, i_1, \ldots, i_{\sum_{r=1}^{k} R_r}, i\}. \tag{4}
$$

Proof Notice that before the kth failure time, the units $X_{s_1}, \ldots, X_{s_{k-1}}$ are failed and the units $X_{i_1}, \ldots, X_{i_{\sum_{r=1}^{k-1} R_r}}$ are removed randomly from the test. Indeed, at the jth ($1 \le j \le k-1$) failure time, R_j units are randomly removed from $(\gamma_j - 1)$ remaining units on test. Therefore, the probability of the event that units indexed by $i_{\sum_{r=1}^{j-1} R_r}, \ldots, i_{\sum_{r=1}^{j} R_r}$ are removed from the test at the jth failure time is equal to $1/\binom{\gamma_j - 1}{R_j}$. Hence, by summing up over all permutations of failures and random removals, we get

$$\Pr\left(X_{k:m:n}^{\mathbf{R}} > y, \Delta_i^{(k)} = 1\right) = \left\{\prod_{j=1}^{k-1} \binom{\gamma_j - 1}{R_j}\right\}^{-1} \sum_{E_i^{(k-1)}} \Pr\left(X_i > y,\right.$$

$$\left. X_i < Z_i^{(k-1)}, \quad X_i > X_{s_k} > X_{s_{k-1}} > \cdots > X_{s_1}, \cap_{j=1}^{k-1}\{Y^{(j)} > X_{s_j}\}\right), \quad (5)$$

where $\cap$ stands for the intersection of the events; for $k \geq 1$, we have

$$Z_i^{(k)} = \min\{X_1, \ldots, X_n\} \setminus \left\{X_{s_1}, \ldots, X_{s_k}, X_{i_1}, \ldots, X_{i_{\sum_{r=1}^{k} R_r}}, X_i\right\},$$

with $Z_i^{(0)} = \min\{X_1, \ldots, X_n\} \setminus \{X_i\}$ and

$$Y^{(j)} = \min\left\{X_{i_{\sum_{r=1}^{j-1} R_r + 1}}, \ldots, X_{i_{\sum_{r=1}^{j} R_r}}\right\}.$$

Using the independence property of the random variables in (5) and performing some algebraic calculations, we find

$$\Pr\left(X_{k:m:n}^{\mathbf{R}} > y, \Delta_i^{(k)} = 1\right) = \left\{\prod_{j=1}^{k-1} \binom{\gamma_j - 1}{R_j}\right\}^{-1}$$

$$\times \sum_{E_i^{(k-1)}} \int_y^{\infty} \int_0^{x} \int_0^{x_{k-1}} \int_0^{x_{k-2}} \cdots \int_0^{x_2} \left(\prod_{j=1}^{k-1} f_{s_j}(x_j; \boldsymbol{\theta}) \bar{F}_{Y^{(j)}}(x_j; \boldsymbol{\theta})\right)$$

$$\times f_i(x; \boldsymbol{\theta}) \bar{F}_{Z_i^{(k-1)}}(x; \boldsymbol{\theta}) \mathrm{d}x_1 \ldots \mathrm{d}x_{k-3} \mathrm{d}x_{k-2} \mathrm{d}x_{k-1} \mathrm{d}x.$$

Therefore, the result follows. □

Corollary 1 *Using Theorem 1, the following results may be deduced:*

1. *The probability for the event that the lifetime of the* i*th unit coincides with the* k*th failure time may be obtained for each* $i = 1, \ldots, n$ *and* $k = 1, \ldots, m$*. That is,*

$$Pr\left(\Delta_i^{(k)} = 1\right) = Pr\left(X_{k:m:n}^{\mathbf{R}} > 0, \Delta_i^{(k)} = 1\right). \quad (6)$$

2. *The joint probability function of* $(X_{k:m:n}, \Delta_i^{(k)})$ *at point* $(y, 1)$ *may be obtained as*

$$f_{X_{k:m:n}^{\mathbf{R}}, \Delta_i^{(k)}}(x, 1) = -\frac{d}{dy} Pr\left(X_{k:m:n}^{\mathbf{R}} > y, \Delta_i^{(k)} = 1\right).$$

3. *By summing up both sides of* (2) *over* i *from 1 up to* n*, the survival function of* $X_{k:m:n}$ *may be derived. That is,*

$$Pr\left(X^{\mathbf{R}}_{k:m:n} > y\right) = \sum_{i=1}^{n} Pr\left(X^{\mathbf{R}}_{k:m:n} > y,\, \Delta_i^{(k)} = 1\right).$$

Theorem 2 *For each* $i = 1, \ldots, n$*,* $k = 1, \ldots, m$ *and every positive real value* y*, we have*

$$\begin{aligned} Pr\left(X^{\mathbf{R}}_{k:m:n} > y,\, H_i^{(k)} = 1\right) &= \frac{R_k}{\gamma_k - 1}\left\{\prod_{j=1}^{k-1}\binom{\gamma_j - 1}{R_j}\right\}^{-1} \\ &\times \sum_{E_i^{(k-1)}} \int_y^{\infty}\int_0^{z}\int_0^{x_{k-1}} \cdots \int_0^{x_2} \left\{\prod_{j=1}^{k-1} f_{s_j}(x_j;\boldsymbol{\theta}) \prod_{t\in C^{(j)}} \bar{F}_t(x_j;\boldsymbol{\theta})\right\} \bar{F}_i(z;\boldsymbol{\theta}) \\ &\times \left\{\sum_{v\in\Lambda_i^{(k-1)}} f_v(z;\boldsymbol{\theta}) \prod_{\substack{t\in\Lambda_i^{(k-1)}\\ t\neq v}} \bar{F}_t(z;\boldsymbol{\theta})\right\} dx_1\cdots dx_{k-2}dx_{k-1}dz, \end{aligned} \tag{7}$$

where $E_i^{(k)}$*,* $C^{(j)}$ *and* $\Lambda_i^{(k)}$ *are as defined in Theorem* 1.

Proof As mentioned in the proof of Theorem 1, the units indexed by $i_{\sum_{r=1}^{j-1} R_r}$, $\ldots, i_{\sum_{r=1}^{j} R_r}$ are removed from the test at the jth failure time with probability $1/\binom{\gamma_j-1}{R_j}$, $1 \le j \le k-1$. Then, after the kth failure time, $(\gamma_k - 1)$ units remain in the experiment. Therefore, occurring the event $\{H_i^{(k)} = 1\}$ is equivalent to removing the ith unit from the test as one of the R_k units which must be randomly censored. So, we get

$$\begin{aligned} \Pr\left(X^{\mathbf{R}}_{k:m:n} > y,\, H_i^{(k)} = 1\right) &= \frac{R_k}{\gamma_k - 1}\left\{\prod_{j=1}^{k-1}\binom{\gamma_j - 1}{R_j}\right\}^{-1} \sum_{E_i^{(k-1)}} \Pr\Big(Z_i^{(k-1)} > y, \\ &X_i > Z_i^{(k-1)} > X_{s_{k-1}} > X_{s_{k-2}} > \cdots > X_{s_1}, \cap_{j=1}^{k-1}\{Y^{(j)} > X_{s_j}\}\Big) \\ &= \frac{R_k}{\gamma_k - 1}\left\{\prod_{j=1}^{k-1}\binom{\gamma_j - 1}{R_j}\right\}^{-1} \sum_{E_i^{(k-1)}} \int_y^{\infty}\int_0^{z}\int_0^{x_{k-1}}\int_0^{x_{k-2}} \cdots \int_0^{x_2} \\ &\left(\prod_{j=1}^{k-1} f_{s_j}(x_j;\boldsymbol{\theta})\bar{F}_{Y^{(j)}}(x_j;\boldsymbol{\theta})\right) \bar{F}_i(z;\boldsymbol{\theta}) f_{Z_i^{(k-1)}}(z;\boldsymbol{\theta}) \\ &\times \mathrm{d}x_1 \ldots \mathrm{d}x_{k-3}\mathrm{d}x_{k-2}\mathrm{d}x_{k-1}\mathrm{d}z, \end{aligned}$$

where

$$f_{Z_i^{(k-1)}}(z;\boldsymbol{\theta}) = \sum_{v\in\Lambda_i^{(k-1)}} f_v(z;\boldsymbol{\theta}) \prod_{\substack{t\in\Lambda_i^{(k-1)}\\ t\neq v}} \bar{F}_t(z;\boldsymbol{\theta}).$$

Hence, the relation (7) is derived. □

Corollary 2 *By use of Theorem 2, the following results are deduced:*

1. *The probability of the event that the ith unit is censored from a lifetime test after the kth failure time, for each $i = 1,\ldots,n$ and $k = 1,\ldots,m$, is given by*

$$Pr\left(H_i^{(k)} = 1\right) = Pr\left(X_{k:m:n}^{\mathbf{R}} > 0,\, H_i^{(k)} = 1\right). \tag{8}$$

2. *The joint probability of $(X_{k:m:n}, H_i^{(k)})$ at point $(y, 1)$ may be derived as follows*

$$f_{X_{k:m:n}^{\mathbf{R}},H_i^{(k)}}(x,1) = -\frac{d}{dy}Pr\left(X_{k:m:n}^{\mathbf{R}} > y,\, H_i^{(k)} = 1\right).$$

Note. In the special case of $k = 1$, we get $E_i^{(1)} = \emptyset$. Hence, the probabilities in (2) and (7) are simplified as

$$\Pr\left(X_{1:m:n}^{\mathbf{R}} > y,\, \Delta_i^{(1)} = 1\right) = \int_y^\infty f_i(x;\boldsymbol{\theta}) \prod_{\substack{j=1\\ j\neq i}}^n \bar{F}_j(x;\boldsymbol{\theta})\, \mathrm{d}x, \tag{9}$$

and

$$\Pr\left(X_{1:m:n}^{\mathbf{R}} > y,\, H_i^{(1)} = 1\right) = \frac{R_1}{n-1}\int_y^\infty \sum_{\substack{v=1\\ v\neq i}}^n f_v(z;\boldsymbol{\theta}) \prod_{\substack{j=1\\ j\neq v}}^n \bar{F}_j(z;\boldsymbol{\theta})\, \mathrm{d}z, \tag{10}$$

respectively.

Remark 2 Based on the proposed procedure in this paper, the following expressions deduce:

– Surely, one of the surviving units in the lifetime test is the kth failure, i.e.,

$$\sum_{i=1}^n \Pr\left(\Delta_i^{(k)} = 1\right) = 1, \quad k = 1,\ldots,m. \tag{11}$$

Therefore, in the case of IID random variables, it is trivial that

$$\Pr\left(\Delta_1^{(k)} = 1\right) = \cdots = \Pr\left(\Delta_n^{(k)} = 1\right) = \frac{1}{n}.$$

– The ith unit on test is either one of the m failures or one of those that are removed from the test. Hence,

$$\sum_{k=1}^{m} \left\{ \Pr\big(\Delta_i^{(k)} = 1\big) + \Pr\big(H_i^{(k)} = 1\big) \right\} = 1, \quad i = 1, \ldots, n.$$

– Certainly, there are R_k units censored from the test at the kth $(1 \le k \le m)$ failure time, such that each of them with probability one is one of the working units on the test. Therefore,

$$\sum_{i=1}^{n} \Pr\left(H_i^{(k)} = 1\right) = R_k, \quad k = 1, \ldots, m. \tag{12}$$

Hence, in the case of IID random variables, it is obvious that

$$\Pr\left(H_1^{(k)} = 1\right) = \cdots = \Pr\left(H_n^{(k)} = 1\right) = \frac{R_k}{n}.$$

4 Inference based on proportional hazard rate family

Let $X_1, \ldots, X_n$ be a sample of independent random variables for which $X_i,\ 1 \le i \le n$, has the survival function

$$\bar{F}_i(x; \boldsymbol{\theta}) = [\bar{G}(x; \boldsymbol{\theta})]^{\lambda_i}, \tag{13}$$

where $\boldsymbol{\theta}' = (\theta_1, \ldots, \theta_t)$ is a common vector of parameters; $\lambda_1, \ldots, \lambda_n$ are known positive constants and $\bar{G} = 1 - G$ is the survival function of the baseline distribution. This family is well known as proportional hazard rate family with proportionality rates $\lambda_1, \ldots, \lambda_n$; see, for example, Lawless (2003). This family includes several well-known lifetime distributions such as exponential, Weibull, Pareto and Burr type-XII. For example, suppose that n units which have been made by different companies and the corresponding lifetimes come from different distributions are independently placed on a life test. Also, suppose that there exists a hierarchical relation among these distributions through relation (13). If one considers the first distribution as the baseline distribution, that is, $\lambda_1 = 1$, then $\lambda_2, \ldots, \lambda_n$ represent the proportionality rates of the other distributions with respect to the first (baseline) distribution. In this section, we would like to make inference about the common parameters of the lifetime distributions in the proportional hazard rate family.

First of all, to determine the joint probability functions of $(X_{k:m:n}^{\mathbf{R}}, \Delta_i^{(k)})$ and $(X_{k:m:n}^{\mathbf{R}}, H_i^{(k)})$, let us define $D_j = C^{(j)} \cup \{s_j\}$ and $A_j = \Lambda_i^{(j)} \cup \{i\}$ with $C^{(j)}$ and $\Lambda_i^{(j)}$ $(1 \le j \le k)$ being as defined in (3) and (4), respectively. In fact, $A_1 = \{1, \ldots, n\}$. Moreover, for each set A, we define

$$\varphi(y; A) = \frac{1}{\sum_{j \in A} \lambda_j} [\bar{G}(y; \boldsymbol{\theta})]^{\sum_{j \in A} \lambda_j}, \tag{14}$$

with $\varphi(y;\emptyset)=0$; further, we consider the set $V^{(k)}=\{(v_1,\ldots,v_k); v_1=+1, v_j=\pm1, 2\le j\le k\}$ with $V^{(0)}=\emptyset$. Then, by using (2) and performing some algebraic calculations, it can be shown that for the proportional hazard rate family (13), for $1\le k\le m$, we get

$$\begin{aligned}
\Pr\left(X^{\mathbf{R}}_{k:m:n}>y,\Delta_i^{(k)}=1\right)&=\frac{\lambda_i}{\prod_{j=1}^{k-1}\binom{\gamma_j-1}{R_j}}\sum_{E_i^{(k-1)}}\left\{\left(\prod_{j=1}^{k-1}\lambda_{s_j}\right)\right.\\
&\left.\times\sum_{V^{(k-1)}}\frac{\prod_{j=1}^{k-1}v_j}{\prod_{j=1}^{k-1}\sum_{r\in W_j}\lambda_r}\left(\varphi(y;A_k)-\varphi(y;A_k-W_{k-1})\right)\right\}\\
&:=\Psi_1\left(k,i;\varphi(y;\cdot)\right),
\end{aligned}\tag{15}$$

where $A-B$ stands for the difference of set B from A. Moreover, $W_0=A_1$, $W_1=D_1$ and for $2\le j\le k$,

$$W_j=\begin{cases}D_j, & \text{if } v_j=+1\\ W_{j-1}\cup D_j, & \text{if } v_j=-1.\end{cases}$$

Similarly, by use of the (7), for $1\le k\le m$, we obtain

$$\begin{aligned}
\Pr\left(X^{\mathbf{R}}_{k:m:n}>y,H_i^{(k)}=1\right)&=\frac{R_k}{(\gamma_k-1)\prod_{j=1}^{k-1}\binom{\gamma_j-1}{R_j}}\sum_{E_i^{(k-1)}}\left\{\left(\prod_{j=1}^{k-1}\lambda_{s_j}\right)\right.\\
&\left.\times\left(\sum_{j\in\Lambda_i^{(k-1)}}\lambda_j\right)\sum_{V^{(k-1)}}\frac{\prod_{j=1}^{k-1}v_j}{\prod_{j=1}^{k-1}\sum_{r\in W_j}\lambda_r}\left(\varphi(y;A_k)-\varphi(y;A_k-W_{k-1})\right)\right\}\\
&:=\Psi_2\left(k,i;\varphi(y;\cdot)\right),
\end{aligned}\tag{16}$$

where $\varphi(y;\cdot)$ is as defined in (14).

Note. Using (9) and (10), the expressions in (15) and (16) may be simplified for the special case of $k=1$, as follows:

$$\Psi_1\left(1,i;\varphi(y;\cdot)\right)=\lambda_i\varphi(y;A_1)$$

and

$$\Psi_2\left(1,i;\varphi(y;\cdot)\right)=\frac{R_1}{n-1}\left(\sum_{j=1}^{n}\lambda_j-\lambda_i\right)\varphi(y;A_1),$$

respectively.

Remark 3 Using (15), the probability functions $f_{X^{\mathbf{R}}_{k:m:n},\Delta_i^{(k)}}(y,1)$ and $f_{X^{\mathbf{R}}_{k:m:n}}(y)$ and from (16), the joint probability function $f_{X^{\mathbf{R}}_{k:m:n},H_i^{(k)}}(y,1)$ may be obtained. Moreover, the failed and censored probabilities in the proportional hazard rate family are derived via (6) and (8), respectively, which are free of the baseline distribution and are only depend on the proportionality rates, $\lambda_1,\ldots,\lambda_n$, and the progressive censoring plan, $\mathbf{R}=(R_1,\ldots,R_m)$.

Using (1) and (13), the log-likelihood function of $\boldsymbol{\theta}$ based on the data set $\mathbf{B}$ is

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^{m}\log\gamma_k + \sum_{k=1}^{m}\sum_{i=1}^{n}\delta_i^{(k)}\log\lambda_i + \sum_{k=1}^{m}\log h(x_k;\boldsymbol{\theta}) + \sum_{k=1}^{m}\sum_{i=1}^{n}\left\{\lambda_i\left(\delta_i^{(k)}+\eta_i^{(k)}\right)\log\bar{G}(x_k;\boldsymbol{\theta})\right\}, \tag{17}$$

where $h=g/\bar{G}$ is the hazard rate function of the baseline distribution with cdf G and pdf g. Using (17), the FI contained in the data set $\mathbf{B}$ about $\boldsymbol{\theta}'=(\theta_1,\ldots,\theta_t)$ is given by the matrix

$$I_{\mathbf{B}}(\boldsymbol{\theta}) = [I_{r,s}(\boldsymbol{\theta})]; \quad r,s=1,\ldots,t, \tag{18}$$

where

$$\begin{aligned} I_{r,s}(\boldsymbol{\theta}) &= -E\left(\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_r\partial\theta_s}\right)\\ &= -\sum_{k=1}^{m}\mathrm{E}\left(\frac{\partial^2}{\partial\theta_r\partial\theta_s}\log h(X^{\mathbf{R}}_{k:m:n};\boldsymbol{\theta})\right)\\ &\quad -\sum_{k=1}^{m}\sum_{i=1}^{n}\lambda_i\mathrm{E}\left[\left(\Delta_i^{(k)}+H_i^{(k)}\right)\left(\frac{\partial^2}{\partial\theta_r\partial\theta_s}\log\bar{G}(X^{\mathbf{R}}_{k:m:n};\boldsymbol{\theta})\right)\right]\\ &= -\sum_{k=1}^{m}\int_0^\infty\left(\frac{\partial^2\log h(y;\boldsymbol{\theta})}{\partial\theta_r\partial\theta_s}\right)f_{X^{\mathbf{R}}_{k:m:n}}(y;\boldsymbol{\theta})\mathrm{d}y\\ &\quad -\sum_{k=1}^{m}\sum_{i=1}^{n}\lambda_i\int_0^\infty\left(f_{X^{\mathbf{R}}_{k:m:n},\Delta_i^{(k)}}(y,1)+f_{X^{\mathbf{R}}_{k:m:n},H_i^{(k)}}(y,1)\right)\\ &\quad\times\left(\frac{\partial^2\log\bar{G}(y;\boldsymbol{\theta})}{\partial\theta_r\partial\theta_s}\right)\mathrm{d}y. \end{aligned} \tag{19}$$

Remark 4 For the special case of $\lambda_1=\cdots=\lambda_n=1$ in the family (13), i.e., when $X_1,\ldots,X_n$ are IID random variables, for fixed k, the random variables $\Delta_i^{(k)}$ and $H_i^{(k)}$ $(1\le i\le n)$ are independent of $X^{\mathbf{R}}_{k:m:n}$, while in the INID case this is not true. Therefore, using (11), (12) and (19), the FI contained in the data set $\mathbf{B}$ about $\boldsymbol{\theta}$ may be obtained from (18), such that $I_{r,s}(\boldsymbol{\theta})$ is simplified as follows

$$
\begin{aligned}
I_{r,s}(\boldsymbol{\theta}) &= -\sum_{k=1}^{m} \mathrm{E}\left(\frac{\partial^2}{\partial\theta_r\partial\theta_s}\log h(X^{\mathbf{R}}_{k:m:n};\boldsymbol{\theta})\right) \\
&\quad -\sum_{k=1}^{m}\sum_{i=1}^{n}\lambda_i\left(\Pr\left(\Delta_i^{(k)}=1\right)+\Pr\left(H_i^{(k)}=1\right)\right)\mathrm{E}\left(\frac{\partial^2\log\bar{G}(X^{\mathbf{R}}_{k:m:n};\boldsymbol{\theta})}{\partial\theta_r\partial\theta_s}\right) \\
&= -\sum_{k=1}^{m}\left\{\mathrm{E}\left(\frac{\partial^2\log h(X^{\mathbf{R}}_{k:m:n};\boldsymbol{\theta})}{\partial\theta_r\partial\theta_s}\right)+(R_k+1)\mathrm{E}\left(\frac{\partial^2\log\bar{G}(X^{\mathbf{R}}_{k:m:n};\boldsymbol{\theta})}{\partial\theta_r\partial\theta_s}\right)\right\}.
\end{aligned}
$$

In the sequel, the exponential family and Weibull distributions are considered as the baseline cdf for the proportional hazard rate family and some details are investigated.

4.1 Baseline exponential family

Let $X_1, \ldots, X_n$ be independent random variables for which X_i $(1 \le i \le n)$ comes from the model (13) with the baseline one-parameter exponential family with the cdf

$$
G(x;\theta) = 1 - \mathrm{e}^{-\beta(\theta)D(x)}, \tag{20}
$$

where $\beta(\theta)$ and $D(x)$ are positive and differentiable functions and θ is a real-valued parameter. Notice that by the Cramer–Rao lower bound, the variance of any estimator of $\vartheta(\theta)$, any differentiable function of θ, is related to the inverse of the FI about θ. Also, the FI plays a valuable role in the asymptotic properties of the MLE. Hence, in spite of computing the FI about θ we would like to study the estimation problem of $\beta_1(\theta) = 1/\beta(\theta)$, the hazard rate function of the baseline distribution at ν: $\beta_2(\theta) = h(\nu;\theta) = g(\nu;\theta)/\bar{G}(\nu;\theta) = \beta(\theta)d(\nu)$, where $d(\nu)$ is the first derivation of $D(\nu)$, and the survival function at ν: $\beta_3(\theta) = \bar{G}(\nu;\theta) = \mathrm{e}^{-\beta(\theta)D(\nu)}$ on the basis of the data set $\mathbf{B}$.

Theorem 3 *Let $X_1, \ldots, X_n$ be independent random variables for which X_i $(1 \le i \le n)$ comes from a proportional hazard rate family with survival function in* (13). *Then, the random variable*

$$
T(\mathbf{B}) = \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{n}\lambda_i\left(\Delta_i^{(j)}+H_i^{(j)}\right)D(X^{\boldsymbol{R}}_{j:m:n}) \tag{21}
$$

is a sufficient statistic for θ on the basis of the data set $\mathbf{B}$ *if and only if the baseline distribution in* (13) *belongs to the exponential family with cdf* (20).

Proof (i) Sufficiency: If the baseline distribution in (13) belongs to the exponential family with the cdf (20), then the joint density (1) is also in the one-parameter exponential family; thus, the result deduces.

(ii) Necessary: By sufficiency of $T(\mathbf{B})$, it is deduced that the observed value of $T(\mathbf{B})$ must be appeared in the last term of the log-likelihood function of θ in (17); moreover, from the third term in the right hand side of (17), we get

$$\prod_{j=1}^{m} h(x_j;\theta) = \eta(x_1,\ldots,x_m)K(\theta),$$

where $(x_1,\ldots,x_m)$ is the observed value of $(X^{\mathbf{R}}_{1:m:n},\ldots,X^{\mathbf{R}}_{m:m:n})$ and $K(\cdot)$ and $\eta(\cdot)$ are some positive functions of θ. Therefore, there exist some positive functions, say $d(\cdot)$ and $\beta(\cdot)$ for which $h(x_j;\theta) = d(x_j)\beta(\theta)$ or equivalently $\bar{G}(x_j;\theta) = \exp\{-\beta(\theta)D(x_j)\}$. □

Remark 5 Let the baseline distribution in a proportional hazard rate model belong to the exponential family with cdf (20); then, using the properties of the one-parameter exponential family (see, for example, Lehmann and Casella 1998), we get

- The FI contained in the data set $\mathbf{B}$ about θ is

$$I_{\mathbf{B}}(\theta) = m\left(\frac{\beta'(\theta)}{\beta(\theta)}\right)^2, \tag{22}$$

  where depends on neither the progressive censoring plan, $\mathbf{R} = (R_1,\ldots,R_n)$, nor proportionality rates, $\lambda_1,\ldots,\lambda_n$.
- The statistic $T(\mathbf{B})$ in (21) is complete, and also using its sufficiency in Theorem 3, it is the best (most efficient) unbiased estimator (BUE) for $\beta_1(\theta)$ on the basis of the data set $\mathbf{B}$ in the sense that its variance attains the Cramer–Rao lower bound.
- The statistic $T(\mathbf{B})$ in (21) is also the consistent MLE of $\beta_1(\theta)$, and using (22) it is asymptotically distributed as $N\left(\beta_1(\theta), \frac{[\beta_1(\theta)]^2}{m}\right)$, where $N(\mu,\sigma^2)$ stands for the normal distribution with mean μ and variance σ^2. Hence, an asymptotic $100(1-\alpha)\%$ confidence interval for $\beta_1(\theta)$ is as follows:

$$\left(\frac{T(\mathbf{B})}{1+\frac{z_{\alpha/2}}{\sqrt{m}}}, \frac{T(\mathbf{B})}{1-\frac{z_{\alpha/2}}{\sqrt{m}}}\right), \tag{23}$$

  where z_α is the αth upper quantile of the standard normal distribution. Notice that the confidence interval in (23) can be used for large values of m such that $\sqrt{m} > z_{\alpha/2}$. It is obvious that larger values of m lead to more reliable confidence interval.

Corollary 3 *Using the invariance property of the MLE, the following results deduce:*

- $\frac{d(\nu)}{T(\mathbf{B})}$ *is the MLE of* $\beta_2(\theta)$*,*
- $e^{-\frac{D(\nu)}{T(\mathbf{B})}}$ *is the MLE of* $\beta_3(\theta)$*,*

where $T(\mathbf{B})$ *is as defined in* (21). *Since* $\beta_2(\theta)$ *and* $\beta_3(\theta)$ *are monotone functions of* $\beta_1(\theta)$*, the asymptotic confidence intervals for these parameters can be obtained using* (23).

4.2 Baseline Weibull distribution

In this section, we assume that the baseline cdf in model (13) is the two-parameter Weibull distribution. The Weibull distribution appears very frequently in practical problems as the most widely used lifetime distribution model. A random variable X is said to have the two-parameter Weibull distribution, denoted by $\mathrm{We}(\theta_1, \theta_2)$, if its cdf is

$$G(x; \theta_1, \theta_2) = 1 - \mathrm{e}^{-\left(\frac{x}{\theta_2}\right)^{\theta_1}}, \quad x > 0,$$

where θ_1 and θ_2 are the shape and scale parameters, respectively.

4.2.1 θ_2 unknown, θ_1 known

Let the shape parameter be known and assume without loss of generality that $\theta_1 = 1$; then, the $\mathrm{We}(1, \theta_2)$ distribution belongs to the exponential family in (20). Therefore, using (22) we get

$$\theta_2^2 I_{\mathbf{B}}(\theta_2) = m,$$

which depends on neither the progressive censoring plan nor proportionality rates. On the other hand, using Remark 5,

$$T^*(\mathbf{B}) = \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_i \left(\Delta_i^{(j)} + H_i^{(j)}\right) X_{j:m:n}^{\mathbf{R}} \tag{24}$$

is a complete sufficient statistic, and also it is the BUE and the MLE of $\beta_1(\theta_2) = \theta_2$. Note that since $T^*(\mathbf{B})$ in (24) is a linear function of $X_{j:m:n}^{\mathbf{R}}$s, it is also the best linear unbiased estimator (BLUE) of θ_2. Moreover, using (23), an asymptotic $100(1-\alpha)\%$ confidence interval for θ_2 is

$$\left(\frac{T^*(\mathbf{B})}{1 + \frac{z_{\alpha/2}}{\sqrt{m}}}, \frac{T^*(\mathbf{B})}{1 - \frac{z_{\alpha/2}}{\sqrt{m}}}\right). \tag{25}$$

Furthermore, by Corollary 3, we have

- $\frac{1}{T^*(\mathbf{B})}$ is the MLE of $\beta_2(\theta_2) = 1/\theta_2$,
- $\mathrm{e}^{-\frac{\nu}{T^*(\mathbf{B})}}$ is the MLE of $\beta_3(\theta_2) = \mathrm{e}^{-\frac{\nu}{\theta_2}}$,

where $T^*(\mathbf{B})$ is as defined in (24). Also, using (25), the asymptotic confidence intervals for $\beta_2(\theta_2)$ and $\beta_3(\theta_2)$ can be derived.

4.2.2 θ_1 unknown, θ_2 known

When the scale parameter is known, without loss of generality we consider the $\mathrm{We}(\theta_1, 1)$ distribution. Therefore, the ith population in a proportional hazard rate family has cdf

$$F_i(x;\theta_1) = 1 - \mathrm{e}^{-\lambda_i x^{\theta_1}}. \tag{26}$$

Using (19) and (26), it can be shown that the FI contained in the data set $\mathbf{B}$ about θ_1 is given by

$$\begin{aligned}\theta_1^2 I_{\mathbf{B}}(\theta_1) &= m + \sum_{k=1}^{m}\sum_{i=1}^{n}\lambda_i\left(\Psi_1(k,i;\mu(\cdot)) + \Psi_2(k,i;\mu(\cdot))\right)\\ &:= \psi(m,n,\boldsymbol{\lambda}),\end{aligned} \tag{27}$$

where $\boldsymbol{\lambda} = (\lambda_1,\ldots,\lambda_n)$ represents the proportionality rates; the functions $\Psi_1\big(k,i;\mu(\cdot)\big)$ and $\Psi_2\big(k,i;\mu(\cdot)\big)$ are as defined in (15) and (16), respectively. Moreover, by doing some algebraic calculations, for each set A, we get

$$\begin{aligned}\mu(A) &= \int_0^\infty y\mathrm{e}^{-y\left(\sum_{j\in A}\lambda_j\right)}(\log y)^2\mathrm{d}y\\ &= \frac{1}{(\sum_{j\in A}\lambda_j)^2}\left\{\frac{\pi^2}{6} - 2\left(\gamma + \log\left(\sum_{j\in A}\lambda_j\right)\right) + \left(\gamma + \log\left(\sum_{j\in A}\lambda_j\right)\right)^2\right\},\end{aligned}$$

where γ is the Euler's constant and $\mu(\emptyset) = 0$ (see, for example, Balakrishnan et al. (2008)).

Using (1), the MLE of θ_1 based on the data set $\mathbf{B}$, denoted by $\hat{\theta}_{1\mathbf{B}}$, is the solution of the following equation

$$\frac{m}{\theta_1} = \sum_{r=1}^{m}\left(\log X^{\mathbf{R}}_{r:m:n}\right)\left\{\left(X^{\mathbf{R}}_{r:m:n}\right)^{\theta_1}\sum_{i=1}^{n}\lambda_i\left(\Delta_i^{(r)} + H_i^{(r)}\right) - 1\right\}. \tag{28}$$

It is trivial that the left-hand side of (28) is a positive decreasing function with respect to θ_1. Also, it can be shown that the right-hand side of (28) is an increasing function of θ_1 for which it converges to a real positive constant as θ_1 tends to infinity. Therefore, the existence and uniqueness of the MLE of θ_1 are confirmed.

Since $\hat{\theta}_{1\mathbf{B}}$ is asymptotically distributed as $N(\theta_1, \frac{\theta_1^2}{\psi(m,n,\boldsymbol{\lambda})})$, an asymptotic $100(1-\alpha)\%$ confidence interval for θ_1 is

$$\left(\frac{\hat{\theta}_{1\mathbf{B}}}{1 + \frac{z_{\alpha/2}}{\sqrt{\psi(m,n,\boldsymbol{\lambda})}}}, \frac{\hat{\theta}_{1\mathbf{B}}}{1 - \frac{z_{\alpha/2}}{\sqrt{\psi(m,n,\boldsymbol{\lambda})}}}\right), \tag{29}$$

where $\psi(m,n,\boldsymbol{\lambda})$ is as defined in (27). Notice that $\psi(m,n,\boldsymbol{\lambda})$ is an increasing function in m; therefore, the confidence interval in (29) can be used for large values of m such that $\sqrt{\psi(m,n,\boldsymbol{\lambda})} > z_{\alpha/2}$.

4.2.3 Both of θ_1 and θ_2 unknown

When both of shape and scale parameters are unknown, the ith population in model (13) has the cdf

$$F_i(x;\theta_1,\theta_2)=1-\mathrm{e}^{-\lambda_i\left(\frac{x}{\theta_2}\right)^{\theta_1}},\quad x>0. \tag{30}$$

It can be shown that the MLEs of θ_1 and θ_2 on the basis of the data set $\mathbf{B}$ are the solutions of the following equations:

$$\begin{cases}\sum_{k=1}^m\sum_{i=1}^n\lambda_i\left(\Delta_i^{(k)}+H_i^{(k)}\right)\left(\frac{X_{k:m:n}^{\mathbf{R}}}{\theta_2}\right)^{\theta_1}\log\left(\frac{X_{k:m:n}^{\mathbf{R}}}{\theta_2}\right)=\frac{m}{\theta_1}+\sum_{k=1}^m\log\left(\frac{X_{k:m:n}^{\mathbf{R}}}{\theta_2}\right),\\ \sum_{k=1}^m\sum_{i=1}^n\lambda_i\left(\Delta_i^{(k)}+H_i^{(k)}\right)\left(\frac{X_{k:m:n}^{\mathbf{R}}}{\theta_2}\right)^{\theta_1}=m.\end{cases}$$

Moreover, using (19) and (30), the FI contained in the data set $\mathbf{B}$ about $\boldsymbol{\theta}'=(\theta_1,\theta_2)$ is given by

$$I(\boldsymbol{\theta})=\begin{pmatrix}I_{11} & I_{12}\\ I_{12} & I_{22}\end{pmatrix},$$

where, by doing some algebraic calculations, it can be shown that

$$\theta_1^2 I_{11}=-\theta_1^2E\left(\frac{\partial^2}{\partial\theta_1^2}\ell(\boldsymbol{\theta})\right)=\psi(m,n,\boldsymbol{\lambda}),$$

where $\ell(\boldsymbol{\theta})$ stands for the log-likelihood function of $\boldsymbol{\theta}$ on the basis of the data set $\mathbf{B}$ and $\psi(m,n,\boldsymbol{\lambda})$ is as defined in (27). Moreover,

$$\begin{aligned}\theta_2^2 I_{22}&=-\theta_2^2E\left(\frac{\partial^2}{\partial\theta_2^2}\ell(\boldsymbol{\theta})\right)\\&=-m\theta_1+\theta_1(\theta_1+1)\sum_{k=1}^m\sum_{i=1}^n\lambda_i\left(\Psi_1\left(k,i;\nu(\cdot)\right)+\Psi_2\left(k,i;\nu(\cdot)\right)\right),\end{aligned}$$

and

$$\begin{aligned}\theta_2 I_{12}&=-\theta_2E\left(\frac{\partial^2}{\partial\theta_1\partial\theta_2}\ell(\boldsymbol{\theta})\right)\\&=m-\sum_{k=1}^m\sum_{i=1}^n\lambda_i\left(\Psi_1\left(k,i;\omega(\cdot)\right)+\Psi_2\left(k,i;\omega(\cdot)\right)\right),\end{aligned}$$

where $\Psi_1\big(k, i; \omega(\cdot)\big)$ and $\Psi_2\big(k, i; \omega(\cdot)$ are as defined in (15) and (16), respectively. Further, for each set A, we have

$$\nu(A) = \int_0^\infty y\mathrm{e}^{-y(\sum_{j\in A}\lambda_j)}\mathrm{d}y = \left(\sum_{j\in A}\lambda_j\right)^{-2}$$

and

$$\omega(A) = \int_0^\infty y(\log y + 1)\mathrm{e}^{-y(\sum_{j\in A}\lambda_j)}\mathrm{d}y = \frac{2 - \gamma - \log\left(\sum_{j\in A}\lambda_j\right)}{\left(\sum_{j\in A}\lambda_j\right)^2},$$

where γ is the Euler's constant.

Note that asymptotic confidence intervals for any function of $\boldsymbol{\theta}' = (\theta_1, \theta_2)$, say $\xi : R^2 \to R$, may be derived through

$$\xi(\hat{\boldsymbol{\theta}}) - \xi(\boldsymbol{\theta}) \to N\left(0, \left(\frac{\partial\xi(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)' I^{-1}(\boldsymbol{\theta})\left(\frac{\partial\xi(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\right), \tag{31}$$

where $I^{-1}(\boldsymbol{\theta})$ is the inverse of the FI matrix $I(\boldsymbol{\theta})$. For example, when the reliability of the baseline population at point x is of interest, we have $\xi(\boldsymbol{\theta}) = \exp\{-(x/\theta_2)^{\theta_1}\}$. Hence,

$$\left(\frac{\partial\xi(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)' = \left(-\left(\frac{x}{\theta_2}\right)^{\theta_1}\log\left(\frac{x}{\theta_2}\right)\mathrm{e}^{-\left(\frac{x}{\theta_2}\right)^{\theta_1}}, \frac{\theta_1}{\theta_2}\left(\frac{x}{\theta_2}\right)^{\theta_1}\mathrm{e}^{-\left(\frac{x}{\theta_2}\right)^{\theta_1}}\right).$$

Therefore, by determining the proportionality rates and doing some numerical computations, the asymptotic confidence intervals may be obtained.

5 Fixed covariates model

Suppose that $X_1, \ldots, X_n$ are independent random variables representing the lifetimes of n units such that for $i = 1, \ldots, n$, X_i has an exponential distribution with parameter $\lambda_i = \exp\{-\mathbf{y}_i'\boldsymbol{\beta}\}$, where $\mathbf{y}_i' = (y_{i1}, \ldots, y_{ip})$ is the observation of the covariates associated with X_i and $\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_p)$ is the regression coefficient. Hence, the pdf of X_i is given by

$$f_i(x) = \mathrm{e}^{\mathbf{y}_i'\boldsymbol{\beta}}\exp\{-x\mathrm{e}^{\mathbf{y}_i'\boldsymbol{\beta}}\}, \quad x > 0.$$

If the units are placed on a life-testing experiment under type II censoring scheme, then using (1), the log-likelihood function of the vector $\boldsymbol{\beta}$ is

$$l(\boldsymbol{\beta}) = \sum_{k=1}^m \log(\gamma_k) + \sum_{k=1}^m\sum_{i=1}^n\left\{\delta_i^{(k)}\mathbf{y}_i'\boldsymbol{\beta} - x_k\mathrm{e}^{\mathbf{y}_i'\boldsymbol{\beta}}\left(\delta_i^{(k)} + \eta_i^{(k)}\right)\right\}.$$

Therefore, the MLE of $\boldsymbol{\beta}$ is the solution of the following equation

$$\sum_{k=1}^{m}\sum_{i=1}^{n}\delta_i^{(k)}\mathbf{y}_i=\sum_{k=1}^{m}\sum_{i=1}^{n}x_k\left(\delta_i^{(k)}+\eta_i^{(k)}\right)\mathbf{y}_i\mathrm{e}^{\mathbf{y}_i'\boldsymbol{\beta}}.$$

Furthermore, the FI matrix of the $\boldsymbol{\beta}$ is $I(\boldsymbol{\beta})=[I_{r,s}(\boldsymbol{\beta})]$, where for $r, s=1,\ldots,p$,

$$\begin{aligned}I_{r,s}(\boldsymbol{\beta})&=-E\left(\frac{\partial^2\ell(\boldsymbol{\beta})}{\partial\beta_r\partial\beta_s}\right)\\&=\sum_{k=1}^{m}\sum_{i=1}^{n}y_{ir}y_{is}\mathrm{e}^{\mathbf{y}_i'\boldsymbol{\beta}}E\left(X_{k:m:n}\left(\Delta_i^{(k)}+H_i^{(k)}\right)\right)\\&=\sum_{k=1}^{m}\sum_{i=1}^{n}y_{ir}y_{is}\mathrm{e}^{\mathbf{y}_i'\boldsymbol{\beta}}\left(\varphi_1^*(k,i;m,n,\boldsymbol{\beta})-\varphi_2^*(k,i;m,n,\boldsymbol{\beta})\right),\end{aligned}\tag{32}$$

where by doing some algebraic calculations, it can be shown that

$$\begin{aligned}\varphi_1^*(k,i;m,n,\boldsymbol{\beta})&=\int_0^\infty xf_{X_{k:m:n},\Delta_i^{(k)}}(x,1)\mathrm{d}x\\&=\frac{\lambda_i}{\prod_{j=1}^{k-1}\binom{\gamma_j-1}{R_j}}\sum_{E_i^{(k-1)}}\left\{\mathrm{e}^{-\left(\sum_{j=1}^{k-1}\mathbf{y}_{s_j}'\right)\boldsymbol{\beta}}\sum_{V^{(k-1)}}\frac{\prod_{j=1}^{k-1}\upsilon_j}{\prod_{j=1}^{k-1}\sum_{r\in W_j}\mathrm{e}^{-\mathbf{y}_r'\boldsymbol{\beta}}}\right.\\&\quad\times\left[\left(\sum_{j\in A_k}\mathrm{e}^{-\mathbf{y}_j'\boldsymbol{\beta}}\right)^{-2}-\left(\sum_{j\in A_k-W_{k-1}}\mathrm{e}^{-\mathrm{e}^{-\mathbf{y}_j'\boldsymbol{\beta}}}\right)^{-2}\right]\end{aligned}$$

and

$$\begin{aligned}\varphi_2^*(k,i;m,n,\boldsymbol{\beta})&=\int_0^\infty xf_{X_{k:m:n},H_i^{(k)}}(x,1)\mathrm{d}x\\&=\frac{R_k}{(\gamma_k-1)\prod_{j=1}^{k-1}\binom{\gamma_j-1}{R_j}}\sum_{E_i^{(k-1)}}\left\{\mathrm{e}^{-\left(\sum_{j=1}^{k-1}\mathbf{y}_{s_j}'\right)\boldsymbol{\beta}}\left(\sum_{j\in\Lambda_i^{(k)}}\mathrm{e}^{-\mathbf{y}_j'\boldsymbol{\beta}}\right)\right.\\&\quad\times\sum_{V^{(k-1)}}\frac{\prod_{j=1}^{k-1}\upsilon_j}{\prod_{j=1}^{k-1}\sum_{r\in W_j}\mathrm{e}^{-\mathbf{y}_r'\boldsymbol{\beta}}}\left[\left(\sum_{j\in A_k}\mathrm{e}^{-\mathbf{y}_j'\boldsymbol{\beta}}\right)^{-2}\right.\\&\quad\left.\left.-\left(\sum_{j\in A_k-W_{k-1}}\mathrm{e}^{-\mathbf{y}_j'\boldsymbol{\beta}}\right)^{-2}\right]\right\},\end{aligned}$$

where A_k and W_k are as used in (15). Using (31) and (32), the asymptotic confidence interval for any function of $\boldsymbol{\beta}$ may be derived.

Table 1 Summary descriptions of Boeing 720 jet aircraft data

i	1	2	3	4	5	6	7	8	9	10
X_{3i-2}	194	413	90	74	55	23	97	50	359	50
X_{3i-1}	15	14	10	57	320	261	51	44	9	254
X_{3i}	41	58	60	48	56	87	11	102	12	5
λ_i	1	0.9017	0.9836	0.6803	0.6230	1.3770	1.0656	1.2787	0.4098	0.7705

Table 2 Progressively type II censored order statistics extracted from the data in Table 1

i	$X_{i:m:n}^{\mathbf{R}^*}$	x_i	Censored units after the ith failure time
1	X_{30}	5	X_2 X_{10} X_{12} X_{23} X_{26}
2	X_8	10	X_5 X_6 X_{15} X_{18} X_{22}
3	X_{21}	11	X_3 X_9 X_{16} X_{27} X_{28}
4	X_{20}	51	X_{11} X_{13} X_{19} X_{25} X_{29}
5	X_7	90	X_1 X_4 X_{14} X_{17} X_{24}

6 Application on a real data set

To illustrate the performance of the proposed procedure in this paper, we use a real data set which consists of the time (in H) of successive failures of the air conditioning system in ten Boeing 720 jet aircrafts; see, Proschan (1963) for a detailed description of the data set. He tested and accepted the hypothesis that the successive failure times are IID exponential for each aircraft, but with different failure rates. Therefore, we assume that the corresponding failure times for the ith aircraft come from the cdf $F_i(x;\sigma) = 1 - e^{-\lambda_i \sigma^{-1} x}$, which coincides with a proportional hazard rate family in (13). Since in the assumptions of our model, λ_i's are known parameters, we consider some arbitrary values for $\lambda_1, \ldots, \lambda_{10}$ as presented in Table 1. Moreover, three observations related to the ith aircraft, denoted by X_{3i-2}, X_{3i-1} and X_{3i} $(1 \le i \le 10)$, which have the same distribution as $F_i(x,\sigma)$, are used to estimate the unknown common parameter σ. In fact, we use a sample of size thirty of failure times for which the first three of them come from cdf $F_1(x;\sigma)$, the second three of them come from cdf $F_2(x;\sigma)$ and eventually the last three of them come from cdf $F_{10}(x;\sigma)$. Notice that the common parameter in these distributions is σ which is of interest; moreover, there exist ten known parameters $\lambda_1, \ldots, \lambda_{10}$ which construct the different distributions related together via relation in (13). Summary descriptions are reported in Table 1.

Using the data in Table 1 and by using the progressive censoring plan $\mathbf{R}^* = (5, 5, 5, 5, 5)$, the first five progressively type II censored order statistics have been extracted. The results are tabulated in Table 2. From the entries of this table, the values of $\Delta_i^{(j)}$ and $H_i^{(j)}$ in the data set $\mathbf{B}$ may also be specified.

Using (24) and the data in Table 2, the observed value of the BLUE and also the MLE of σ on the basis of the data set $\mathbf{B}$ is given by

$$t_{\mathbf{B}}^* = \frac{1}{5}\{x_1(\lambda_2 + \lambda_{10} + \lambda_{12} + \lambda_{23} + \lambda_{26} + \lambda_{30})$$

$$
\begin{aligned}
&+x_2\,(\lambda_5+\lambda_6+\lambda_8+\lambda_{15}+\lambda_{18}+\lambda_{22})\\
&+x_3\,(\lambda_3+\lambda_9+\lambda_{16}+\lambda_{21}+\lambda_{27}+\lambda_{28})\\
&+x_4\,(\lambda_{11}+\lambda_{13}+\lambda_{19}+\lambda_{20}+\lambda_{25}+\lambda_{29})\\
&+x_5\,(\lambda_1+\lambda_4+\lambda_7+\lambda_{14}+\lambda_{17}+\lambda_{24})\}\\
=&\ 187.3066. \qquad (33)
\end{aligned}
$$

From (25) and (33), the observed values of asymptotic 90 and 95% confidence intervals for σ based on the data in Table 2 are (108.0554, 702.6551) and (99.8149, 1517.1270), respectively. Notice that in small samples, the asymptotic confidence intervals lead to unreliable results. For this reason, the upper bounds of the confidence intervals for σ are rather large compared to its estimate.

7 Numerical computations

Let $X_1, \dots, X_n$ be independent random variables from a proportional hazard rate family for which X_i $(1 \le i \le n)$ has the same distribution as presented in (26). In this section, we determine the amount of FI about θ_1 contained in the data set $\mathbf{B}$ for given proportionality rates. Toward this end, we assume that $n = 8$, $m = 4$ and consider five n-tuples of proportionality rates as $\boldsymbol{\lambda}_1 = (1, 1, 1, 1, 1, 1, 1, 1)$ which corresponds to the case of IID random variables from $We(\theta_1, 1)$ distribution, $\boldsymbol{\lambda}_2 = (1, 0.8, 0.6, 0.4, 0.35, 0.3, 0.25, 0.2)$, $\boldsymbol{\lambda}_3 = (1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3)$, $\boldsymbol{\lambda}_4 = (1, 2, 3, 4, 5, 6, 7, 8)$ and $\boldsymbol{\lambda}_5 = (1, 3, 5, 7, 9, 11, 13, 15)$. Using (27), the numerical values of $\theta_1^2 I_{\mathbf{B}}(\theta_1)$ are presented in Table 3, for some choices of progressive censoring plans, $\mathbf{R}$.
From Table 3, it is deduced that:

1. For the cases in which all $\lambda_1, \dots, \lambda_8$ are less than or equal to 1 (such as in $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$ and $\boldsymbol{\lambda}_3$), later removal of working units decreases the amount of FI about θ_1. Therefore, it is better to remove units at the first stage of the test.
2. If all $\lambda_1, \dots, \lambda_8$ are greater than or equal to 1 (such as in $\boldsymbol{\lambda}_4$ and $\boldsymbol{\lambda}_5$), later removal of working units increases the FI of θ_1. Therefore, it is preferred to remove units at the last stage of the test.

8 Concluding remarks

In this paper, we assumed that $X_1, \dots, X_n$ are the lifetimes of n units where are independently and simultaneously placed on a test for which X_r is distributed with cdf $F_r(x; \boldsymbol{\theta})$, $1 \le r \le n$, where $\boldsymbol{\theta}$ is the common vector of parameters of these distributions. The likelihood function of $\boldsymbol{\theta}$ was derived based on INID progressively type II censored order statistics and the indicator random variables that identify the failed units and those that are removed from the experiment. Some major results were obtained regarding the distribution theory of these statistics. To construct the nonidentical distributed random variables, a proportional hazard rate family was considered and it was shown that in this family the probabilities for the events that the ith unit is

Table 3 Values of $\theta_1^2 I_{\mathbf{B}}(\theta_1)$ for proportionality rate $\boldsymbol{\lambda}_i$ $(1 \leq i \leq 6)$ and some choices of **R**

R	$\boldsymbol{\lambda}_1$	$\boldsymbol{\lambda}_2$	$\boldsymbol{\lambda}_3$	$\boldsymbol{\lambda}_4$	$\boldsymbol{\lambda}_5$
(4,0,0,0)	8.7609	13.7163	10.8837	16.5671	23.6522
(0,4,0,0)	8.4512	11.6792	9.6109	18.7637	26.9462
(0,0,4,0)	7.5424	9.0806	7.7709	19.9339	29.0779
(0,0,0,4)	6.2403	5.9955	5.4845	20.4534	30.4925
(3,1,0,0)	8.5275	12.7921	10.2666	17.2597	24.7602
(2,2,0,0)	8.4452	12.2530	9.9300	17.8450	25.6374
(1,3,0,0)	8.4312	11.9100	9.7319	18.3400	26.3511
(3,0,1,0)	8.0046	11.5344	9.3338	17.4715	25.3316
(2,0,2,0)	7.6902	10.3276	8.5278	18.3815	26.7911
(1,0,3,0)	7.5680	9.5784	8.0591	19.2035	28.0246
(3,0,0,1)	7.1959	9.8143	7.9880	17.2368	25.3935
(2,0,0,2)	6.5858	7.8898	6.6682	18.3675	27.3407
(1,0,0,3)	6.3284	6.7428	5.9314	19.4591	29.0324
(0,0,3,1)	6.7236	7.2611	6.3828	19.9365	29.4975
(0,0,2,2)	6.4447	6.5739	5.8864	20.1384	29.9375
(0,0,1,3)	6.3129	6.2147	5.6347	20.3129	30.2555

the kth failure or that it is censored after the kth failure time are free of the baseline distribution. The results were derived in details for the baseline one-parameter exponential and two-parameter Weibull family of distributions, and they also extended to a fixed covariates model with multi-dimensional parameter. In the case of one-parameter exponential family, a real data set was used to illustrate the performance of the proposed procedure. Some numerical computations were also presented to study the effect of the proportionality rates in view of the FI of the shape parameter of a Weibull distribution contained in the data set **B**. The proposed procedure in this paper can be extended to the following cases:

– Statistical inferences for the proportional hazard rate family have been obtained by assuming that $\lambda_1, \ldots, \lambda_n$ are known positive constants for which $\lambda_1 = 1$. When λ_i $(2 \leq i \leq n)$ is unknown, one can estimate it nonparametrically. Suppose that there exists a sample of size N_i from the ith $(1 \leq i \leq n)$ distribution; then, the nonparametric MLE for λ_i is

$$\hat{\lambda}_i = \frac{-N_i}{\sum_{j=1}^{n_i} \log \hat{\bar{G}}(X_{i,j}; \theta)}, \quad 2 \leq i \leq n,$$

where $\hat{\bar{G}}(x; \theta)$ is the empirical estimator of $\bar{G}(x; \theta)$ obtained on the basis of the sample comes from the first (baseline) distribution; see, Razmkhah et al. (2008).

– Let $Y_1, \ldots, Y_n$ be independent random variables for which Y_i $(1 \leq i \leq n)$ is distributed with the cdf $F_i(x; \theta) = [F_0(x; \theta)]^{\beta_i}$, where $F_0(x; \theta)$ is an absolutely continuous cdf and β_i is a known positive constant. The aforementioned identity is well known in the lifetime literature as proportional reversed hazard rate family which includes several well-known distributions such as power function and Burr type III, Fréchet; see, for example, Lawless (2003). In this case, the results of the paper hold with obvious modifications.

- Using the results of Theorem 1 regarding the probability of the event that the lifetime of the ith unit is the kth failure time, we can develop a general approach to robust inference about the parameter of interest in the presence of one or more outliers.

Acknowledgements The authors would like to thank an anonymous referee and the associate editor for their useful comments and constructive criticisms on the original version of this manuscript which led to this considerably improved version. This research was supported by a Grant from Ferdowsi University of Mashhad (MS90212RZM).

References

Abo-Eleneen, Z. A. (2008). Fisher information in type II progressive censored samples. *Communications in Statistics-Theory and Methods*, *37*, 682–691.

Balakrishnan, N. (2007). Progressive censoring methodology: an appraisal. *Test*, *16*, 211–259.

Balakrishnan, N., Aggarwala, R. (2000). *Progressive censoring*: *Theory, methods and applications*. Boston: Birkhauser.

Balakrishnan, N., Burkschat, M., Cramer, E., Hofmann, G. (2008). Fisher information based progressive censoring plans. *Computational Statistics and Data Analysis*, *53*, 366–380.

Balakrishnan, N., Cramer, E. (2008). Progressive censoring from heterogeneous distributions with applications to robustness. *Annals of the Institute of Statistical Mathematics*, *60*, 151–171.

Balakrishnan, N., Cramer, E. (2014). *The art of progressive censoring*: *applications to reliability and quality*. New York: Birkhauser.

Balakrishnan, N., Cramer, E., Kamps, U., Schenk, N. (2001). Progressive type II censored order statistics from exponential distributions. *Statistics*, *35*, 537–556.

Burkschat, M., Cramer, E., Kamps, U. (2006). On optimal schemes in progressive censoring. *Statistics & Probability Letters*, *76*, 1032–1036.

Cramer, E., Lenz, U. (2010). Association of progressively type II censored order statistics. *Journal of Statistical Planning and Inference*, *140*, 576–583.

Fischer, T., Balakrishnan, N., Cramer, E. (2008). Mixture representation for order statistics from INID progressively censoring and its applications. *Journal of Multivariate Analysis*, *99*, 1999–2015.

Lawless, J. L. (2003). *Statistical models and methods for lifetime data* (2nd ed.). New York: Wiley.

Lehmann, E. L., Casella, G. (1998). *Theory of point estimation* (2nd ed.). New York: Springer.

Mao, T., Hu, T. (2010). Stochastic properties of INID progressive type II censored order statistics. *Journal of Multivariate Analysis*, *101*, 1493–1500.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics*, *5*, 375–383.

Razmkhah, M., Ahmadi, J., Khatib, B. (2008). Nonparametric confidence intervals and tolerance limits based on minima and maxima. *Communications in Statistics-Theory and Methods*, *37*, 1525–1542.

Rezapour, M., Alamatsaz, M. H., Balakrishnan, N., Cramer, E. (2013). On properties of progressively Type II censored order statistics arising from dependent and non-identical random variables. *Statistical Methodology*, *10*, 58–71.

Zheng, G., Park, S. (2004). On the Fisher information in multiply censored and progressively censored data. *Communications in Statistics-Theory and Methods*, *23*, 1821–1835.