Reliability Engineering and System Safety 137 (2015) 129–140

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

EM algorithm for one-shot device testing with competing risks under exponential distribution

N. Balakrishnan [a], H.Y. So [a], M.H. Ling [b,*]

[a] Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S4K1
[b] Department of Mathematics and Information Technology, The Hong Kong Institute of Education, Hong Kong SAR, China

ARTICLE INFO

Article history:
Received 22 May 2014
Received in revised form
13 November 2014
Accepted 30 December 2014
Available online 13 January 2015

Keywords:
EM algorithm
Inequality constrained least squares
One-shot device
Competing risks
Masked data
Exponential distribution
ED01 Data

ABSTRACT

This paper provides an extension of the work of Balakrishnan and Ling [1] by introducing a competing risks model into a one-shot device testing analysis under an accelerated life test setting. An Expectation Maximization (EM) algorithm is then developed for the estimation of the model parameters. An extensive Monte Carlo simulation study is carried out to assess the performance of the EM algorithm and then compare the obtained results with the initial estimates obtained by the Inequality Constrained Least Squares (ICLS) method of estimation. Finally, we apply the EM algorithm to a clinical data, ED01, to illustrate the method of inference developed here.

1. Introduction

One-shot devices often have multiple components that can cause failure. For example, a fire extinguisher contains a cylinder, a valve and chemicals inside; an automobile air bag contains a crash sensor, an inflator and an air bag; and for any packed food (which is also a kind of one-shot device), there are different causes for food expiry such as the growth of microorganism in the package, the moisture level and the food deterioration due to oxidation. A failure of any of the components will result in the failure of the product. However, by the very nature of one-shot devices, each product can be used only once at a specified time (Inspection Time, *IT*) and it gets destroyed afterwards. The outcome from each of the devices is therefore binary, either left-censored (failure) or right-censored (success). The data, which is a collection of these outcomes, will be both left- and right-censored, an extreme form of interval censoring. For those failed units, we will normally check for the cause responsible for the failure. Thus, the information collected from a life-test on one-shot devices in this case will include the status of the unit at inspection time as well as the cause of failure in case the unit has failed.

The one-shot device testing considered in this work will be conducted in an accelerated life test (ALT) setting since we will be often interested in the reliability assessment of highly reliable products. If the products are tested under normal conditions, the failure times of the products will be very large, thus resulting in a very large testing time. ALT will shorten the lifetimes of products by increasing the stress levels, and we can use several stress factors such as temperature and humidity for this purpose. After estimating the parameters under high stress conditions, we can extrapolate the life characteristics such as mean life time and failure rates from high stress conditions to normal operating conditions; see Meeter and Meeker [2] and Meeker et al. [3].

Expectation Maximization (EM) algorithm will be adopted in this paper. It is a powerful technique for obtaining the Maximum Likelihood Estimates (MLEs) in the case of complicated likelihood with missing data. The data obtained from one-shot device testing are both left- and right-censored as mentioned above and so the missing information is usually large. The maximum likelihood estimators will not be in a closed-form. For this reason, the EM algorithm will be a natural way to handle the estimation problem in this case. It enables an efficient determination of the maximum likelihood estimates. Considerable work has been done on estimating the model parameters by the use of EM algorithm: Ng et al. [4] developed EM algorithm for log-normal and Weibull distribution based on progressively censored data; Nandi and Dewan [5] developed EM algorithm for estimating the parameters of the bivariate Weibull distribution under random censoring scheme; and Pal [6] has discussed an EM algorithm for cure rate models. The EM algorithm has also been employed in estimating the masked causes under a competing risk model; see Park [7] and Craiu and Duchesne [8].

* Corresponding author.
E-mail addresses: bala@mcmaster.ca (N. Balakrishnan), sohy@math.mcmaster.ca (H.Y. So), amhling@ied.edu.hk (M.H. Ling).

http://dx.doi.org/10.1016/j.ress.2014.12.014
0951-8320/

For and cost effectiveness as well as convenience, the devices in ALT are often inspected at certain discrete time points. ALT in this form is referred to as intermittent inspection. The one-shot model is a special case of intermittent inspection when each device under test will be inspected only once and gets destroyed afterwards. Sohn [9] provided a model and optimal test plans for analyzing the life characteristics from a stockpile of one-shot products under the logistic lifetime distribution. Yates and Mosleh [10] proposed a Bayesian method for the reliability estimation in an aerospace system. Newby [11] assumed the one-shot device having repairable components and developed a monitoring and maintenance procedure. Fan et al. [12] used Bayesian approaches for analyzing highly reliable one-shot devices. They suggested three priors for the Bayesian estimation: Exponential, Normal and Beta, and their simulation results show that all three priors perform similarly when the data possess enough information. However, if the data possess zero-failure cases, normal prior has been recommended. Fan and Chang [13] also focused on the zero-failure reliability test of high quality one-shot devices with Bayesian analysis. Chuang [14] developed statistical inference for the mean-time-to-failure and reliability of one-shot devices under the Weibull life time distribution. Yang [15] studied the one-shot device under a Brownian degradation process using the Bayesian approach. Balakrishnan and Ling [1] developed an EM algorithm for the estimation of parameters of one-shot device model under exponential distribution with a single stress factor, and Balakrishnan and Ling [16] further extended it to multiple stress factors. Balakrishnan and Ling [17,18] developed an EM algorithm for the same problem under Weibull and gamma lifetime distributions, respectively.

Some work has been done on the analysis of system with binomial subsystems and components, but not under the ALT form. Martz et al. [19] analyzed series systems of binomial sub-systems and components in a Bayesian setup without the linkage to the stress levels. Martz and Wailer [20] considered a more complex system that includes both series and parallel components.

However, none of these works consider the competing risks setting. With competing risks, the model becomes more complicated than all those considered earlier and so the corresponding estimation problem becomes quite complex. But, this competing risks model is more realistic since many one-shot devices do contain multiple components that could cause the failure of the device. This is the motivation for us to consider here the one-shot device model with competing risks.

In this paper, we will first specify the one-shot device testing model with competing risks in Section 2. We will assume that the lifetime distribution is exponential and that there are no masked causes of failure in the data. For convenience, we will confine our attention to the case of two competing risks corresponding to the failure of each device and then develop the EM algorithm in Section 3. The extension to the case of multiple competing risks can be done in a natural way. For evaluating the performance of the developed EM algorithm, a simulation study is conducted in Section 4. In Section 5, the proposed method is compared with the Fisher-scoring method. The modification of the algorithm for handling data with masked cause of failure is then discussed in Section 6. An illustrative example with a set of clinical data, ED01, is analyzed in Section 8. Finally, some concluding remarks are made in Section 9.

2. Model specification

Let us consider the testing of electro-explosive devices. The structure of an ordinary electro-explosive device is displayed in Fig. 1. Let us now assume that there are only two causes responsible for the failure of detonation, say, burnout of resistance wires (part 20 in Fig. 1) as Cause 1 and leakage of organic fuel (part 6 in Fig. 1) as Cause 2. An accelerated life test for such one-shot devices is set up as follows:

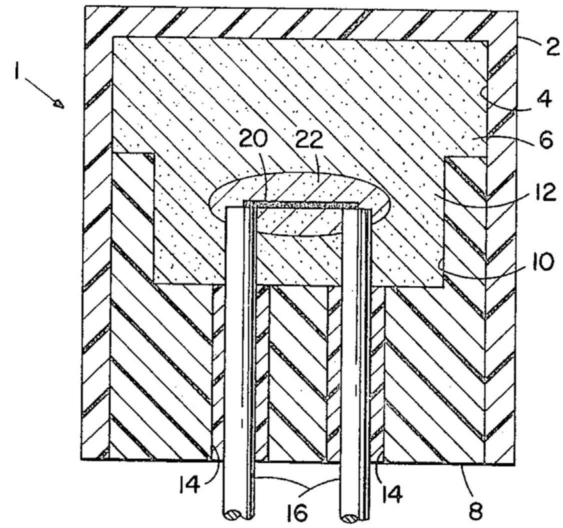

Fig. 1. An electro-explosive device designed by Thomas and Betts [21].

1. the tests are only checked at inspection times IT_i, for $i=1,\ldots,I$,
2. the devices are tested under different temperatures (as stress levels) w_j, for $j=1,\ldots,J$,
3. there are K_{ij} devices tested at IT_i and w_j,
4. the number of devices failed due to the rth cause at IT_i and w_j is denoted by D_{rij}, for $r=1,\ldots,R$,
5. the number of devices that survive (successfully detonated) at IT_i and w_j is denoted by $S_{ij}=K_{ij}-\sum_{r=1}^{R} D_{rij}$.

Let us denote the random variable for the failure time due to Causes 1 and 2 by T_{rijk}, for $r=1,2$, $i=1,\ldots,I$, $j=1,\ldots,J$ and $k=1,\ldots,K_{ij}$, respectively. In this work, we assume that T_{rijk} are independent of each other and follows exponential distribution with rate parameter λ_{rj} with p.d.f.

$$f_{rj}(t)=\lambda_{rj}e^{-\lambda_{rj}t}, r=1,2, j=1,\ldots,J,$$

and with c.d.f.

$$F_r(IT_i\,|\,w_j)=\int_0^{IT_i} f_{rj}(t)\,dt=1-e^{-\lambda_{rj}IT_i}, r=1,2, i=1,\ldots,I, j=1,\ldots,J,$$

where λ_{rj} is the failure rate of the rth component in the device under temperature w_j. Of course, t_{rijk} will be used to denote the realization of the r.v. T_{rijk}. The relationship between λ_{rj} and w_j is assumed to be a log-link function of the form

$$\lambda_{rj}(\boldsymbol{\alpha})=\alpha_{r0}\exp(\alpha_{r1}w_j), \quad \alpha_{r0},\alpha_{r1},w_j\geq 0. \tag{1}$$

We define Δ_{ijk} to be the indicator for the kth device under temperature w_j and inspection time IT_i. When the device is successfully detonated, we will set $\Delta_{ijk}=0$. However, if the device fails to detonate, we will identify (by careful inspection) the specific cause responsible for the failure. If Risk r is the cause for the failure, we will denote this event by $\Delta_{ijk}=r$, for r=1,2. Mathematically, the indicator Δ_{ijk} is then defined as

$$\Delta_{ijk}=\begin{cases}0 & \text{if } \min(T_{1ijk},T_{2ijk})>IT_i,\\ 1 & \text{if } T_{1ijk}<\min(T_{2ijk},IT_i),\\ 2 & \text{if } T_{2ijk}<\min(T_{1ijk},IT_i),\end{cases} \tag{2}$$

and then δ_{ijk} will be used for the realization of Δ_{ijk}.

For example, if we conduct the ALT under temperatures $w_j=35,45,55,65\ ^\circ\text{C}$ and with inspection times $IT_i=10,20,30$

Table 1
An example of one-shot device testing data under temperatures 35, 45, 55, 65 (in °C) and inspection times 10, 20, 30 (in hours) with 2 competing causes of failure.

	$\delta_{ijk}=0$	$\delta_{ijk}=1$	$\delta_{ijk}=2$
$IT_1=10$			
$w_1=35$	$s_{11}=9$	$d_{111}=1$	$d_{211}=0$
$w_2=45$	$s_{12}=9$	$d_{112}=1$	$d_{212}=0$
$w_3=55$	$s_{13}=9$	$d_{113}=0$	$d_{213}=1$
$w_4=65$	$s_{14}=9$	$d_{114}=0$	$d_{214}=1$
$IT_2=20$			
$w_1=35$	$s_{21}=10$	$d_{121}=0$	$d_{221}=0$
$w_2=45$	$s_{22}=9$	$d_{122}=1$	$d_{222}=0$
$w_3=55$	$s_{23}=8$	$d_{123}=2$	$d_{223}=0$
$w_4=65$	$s_{24}=8$	$d_{124}=1$	$d_{224}=1$
$IT_3=30$			
$w_1=35$	$s_{31}=8$	$d_{131}=2$	$d_{231}=0$
$w_2=45$	$s_{32}=8$	$d_{132}=2$	$d_{232}=0$
$w_3=55$	$s_{33}=8$	$d_{133}=2$	$d_{233}=0$
$w_4=65$	$s_{34}=2$	$d_{134}=3$	$d_{234}=5$

with $K_{ij}=10$ devices placed under every testing condition, then the data observed from such an experiment will be as in Table 1.

We also denote p_{0ij}, p_{1ij} and p_{2ij} for the survival probability, failure probability due to Cause 1 and failure probability due to Cause 2, respectively, which are as follows:

$$p_{0ij}=(1-F_1(IT_i\,|\,w_j))(1-F_2(IT_i\,|\,w_j))=\exp(-(\lambda_{1j}(\boldsymbol{\alpha})+\lambda_{2j}(\boldsymbol{\alpha}))IT_i), \quad (3)$$

$$p_{1ij}=\left(\frac{\lambda_{1j}(\boldsymbol{\alpha})}{\lambda_{1j}(\boldsymbol{\alpha})+\lambda_{2j}(\boldsymbol{\alpha})}\right)(1-\exp(-(\lambda_{1j}(\boldsymbol{\alpha})+\lambda_{2j}(\boldsymbol{\alpha}))IT_i)), \quad (4)$$

$$p_{2ij}=\left(\frac{\lambda_{2j}(\boldsymbol{\alpha})}{\lambda_{1j}(\boldsymbol{\alpha})+\lambda_{2j}(\boldsymbol{\alpha})}\right)(1-\exp(-(\lambda_{1j}(\boldsymbol{\alpha})+\lambda_{2j}(\boldsymbol{\alpha}))IT_i)). \quad (5)$$

Now the data collected at temperatures $\boldsymbol{w}=\{w_j, j=1,2\ldots,J\}$ and inspection times $\boldsymbol{IT}=\{IT_i, i=1,\ldots,I\}$ are the numbers of devices with the indicator values $\delta_{ijk}=0, \delta_{ijk}=1$ and $\delta_{ijk}=2$, which are denoted by S_{ij}, D_{1ij} and D_{2ij}, respectively. Then, the likelihood function of $\boldsymbol{\alpha}=\{\alpha_{10},\alpha_{11},\alpha_{20},\alpha_{21}\}$ is given by

$$L(\boldsymbol{\alpha}\,|\,\boldsymbol{\delta},\boldsymbol{IT},\boldsymbol{w})=\prod_{i=1}^{I}\prod_{j=1}^{J}p_{0ij}^{S_{ij}}p_{1ij}^{D_{1ij}}p_{2ij}^{D_{2ij}}, \quad (6)$$

where $K_{ij}=S_{ij}+D_{1ij}+D_{2ij}$.

3. EM algorithm

It is known that the EM algorithm is an efficient way to determine the Maximum Likelihood Estimates (MLEs) of the model parameters in the presence of missing data arising from censoring and masking effects; see Casella and Berger [22] and McLachlan and Krishnan [23] for details. It involves approximating the missing data by their expectation, given the observed data and the current estimates of the parameters (E-step), and then maximizing the corresponding likelihood function to obtain the updated parameter estimates (M-step). If we repeat the E-step and M-step iteratively, by the monotonicity of EM algorithm, the numerical values of the MLEs will be achieved to the desired level of accuracy.

In the example of electro-explosive devices, the parameters of interest are α_{r0}, α_{r1}, $r=1,2$, and the data that are not observable are the true lifetimes of the devices. Let us denote that lifetime by $T_{rijk}^{(\Delta_{ijk})}$ defined as

$$T_{rijk}^{(\Delta_{ijk})}=\begin{cases} T_{rijk}\,|\,\min(T_{1ijk},T_{2ijk})>IT_i & \text{when } \Delta_{ijk}=0 \\ T_{rijk}\,|\,T_{1ijk}<\min(T_{2ijk},IT_i) & \text{when } \Delta_{ijk}=1 \\ T_{rijk}\,|\,T_{2ijk}<\min(T_{1ijk},IT_i) & \text{when } \Delta_{ijk}=2 \end{cases}. \quad (7)$$

Let us denote $\boldsymbol{\alpha}=(\alpha_{10},\alpha_{11},\alpha_{20},\alpha_{21})$ and $\boldsymbol{\alpha}'$ for the current estimate of $\boldsymbol{\alpha}$. Then, the complete data likelihood is given by

$$l^{\text{complete}}(\boldsymbol{\alpha})=\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K_{ij}}\log\left(f_{1j}(T_{1ijk}^{(\delta_{ijk})})\right)+\log\left(f_{2j}(T_{2ijk}^{(\delta_{ijk})})\right)$$
$$=\sum_{i=1}^{I}\sum_{j=1}^{J}l_{ij}^{\text{complete}}(\boldsymbol{\alpha}),$$

where

$$l_{ij}^{\text{complete}}(\boldsymbol{\alpha})=\sum_{k=1}^{K_{ij}}\log(\lambda_{1j}(\boldsymbol{\alpha}))-\lambda_{1j}(\boldsymbol{\alpha})T_{1ijk}^{(\delta_{ijk})}+\log(\lambda_{2j}(\boldsymbol{\alpha}))-\lambda_{2j}(\boldsymbol{\alpha})T_{2ijk}^{(\delta_{ijk})}$$
$$=K_{ij}\left(\log(\lambda_{1j}(\boldsymbol{\alpha}))+\log(\lambda_{2j}(\boldsymbol{\alpha}))\right)$$
$$-\lambda_{1j}(\boldsymbol{\alpha})\sum_{k=1}^{K_{ij}}T_{1ijk}^{(\delta_{ijk})}-\lambda_{2j}(\boldsymbol{\alpha})\sum_{k=1}^{K_{ij}}T_{2ijk}^{(\delta_{ijk})}.$$

3.1. E-step

In the E-step of the EM algorithm, we shall take the expected value of the missing data, given the observed data and the current parameter estimates, to approximate the missing data. It is given by

$$E(l_{ij}^{\text{complete}}(\boldsymbol{\alpha})\,|\,\boldsymbol{\alpha}')=K_{ij}\left(\log(\lambda_{1j}(\boldsymbol{\alpha}))+\log(\lambda_{2j}(\boldsymbol{\alpha}))\right)$$
$$-\lambda_{1j}\left(S_{ij}E(T_{1ijk}^{(0)}\,|\,\boldsymbol{\alpha}')+D_{1ij}E(T_{1ijk}^{(1)}\,|\,\boldsymbol{\alpha}')+D_{2ij}E(T_{1ijk}^{(2)}\,|\,\boldsymbol{\alpha}')\right)$$
$$-\lambda_{2j}\left(S_{ij}E(T_{2ijk}^{(0)}\,|\,\boldsymbol{\alpha}')+D_{1ij}E(T_{2ijk}^{(1)}\,|\,\boldsymbol{\alpha}')+D_{2ij}E(T_{2ijk}^{(2)}\,|\,\boldsymbol{\alpha}')\right), \quad (8)$$

where the conditional expectations, $E(T_{rijk}^{(\delta_{ijk})}|\boldsymbol{\alpha}')$, are as calculated in Table 2. It is important to note here that the expectations are in simple closed-form since the lifetimes here are assumed to follow an exponential distribution.

Then, the partial objective function $Q_{ij}(\boldsymbol{\alpha}\,|\,\boldsymbol{\alpha}')$ for $l_{ij}^{\text{complete}}(\boldsymbol{\alpha})$ is

$$Q_{ij}(\boldsymbol{\alpha}\,|\,\boldsymbol{\alpha}',IT_i)=E(l_{ij}^{\text{complete}}(\boldsymbol{\alpha})\,|\,\boldsymbol{\alpha}')$$
$$=K_{ij}\left(\log(\lambda_{1j}(\boldsymbol{\alpha}))+\log(\lambda_{2j}(\boldsymbol{\alpha}))\right)-\lambda_{1j}g_{1ij}(\boldsymbol{\alpha}')-\lambda_{2j}g_{i2}(\boldsymbol{\alpha}'), \quad (9)$$

Table 2
The conditional expected values of missing data, with $\lambda'_{rj}=\lambda_{rj}(\boldsymbol{\alpha}')$, $r=1,2$, for different cases.

Case	No. of cases	$E\left(T_{1ijk}^{(\delta_{ijk})}\middle|\boldsymbol{\alpha}'\right)$
$\delta_{ijk}=0$	S_{ij}	$IT_i+\frac{1}{\lambda'_{1j}}$
$\delta_{ijk}=1$	D_{1ij}	$\frac{1}{\lambda'_{1j}+\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
$\delta_{ijk}=2$	D_{2ij}	$\frac{1}{\lambda'_{1j}}+\frac{1}{\lambda'_{1j}+\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
Case	No. of cases	$E\left(T_{2ijk}^{(\delta_{ijk})}\middle|\boldsymbol{\alpha}'\right)$
$\delta_{ijk}=0$	S_{ij}	$IT_i+\frac{1}{\lambda'_{2j}}$
$\delta_{ijk}=1$	D_{1ij}	$\frac{1}{\lambda'_{2j}}+\frac{1}{\lambda'_{1j}+\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
$\delta_{ijk}=2$	D_{2ij}	$\frac{1}{\lambda'_{1j}+\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$

where

$$g_{1ij}(\boldsymbol{\alpha}')=\left(S_{ij}E(T_{1ijk}^{(0)}|\boldsymbol{\alpha}')+D_{1ij}E(T_{1ijk}^{(1)}|\boldsymbol{\alpha}')+D_{2ij}E(T_{1ijk}^{(2)}|\boldsymbol{\alpha}')\right), \tag{10}$$

$$g_{2ij}(\boldsymbol{\alpha}')=\left(S_{ij}E(T_{2ijk}^{(0)}|\boldsymbol{\alpha}')+D_{1ij}E(T_{2ijk}^{(1)}|\boldsymbol{\alpha}')+D_{2ij}E(T_{2ijk}^{(2)}|\boldsymbol{\alpha}')\right). \tag{11}$$

3.2. M-Step

From (9), the objective function for maximizing the overall likelihood function $l^{\text{complete}}(\boldsymbol{\alpha})$ will be the summation of the partial objective functions, $Q_{ij}(\boldsymbol{\alpha}|\boldsymbol{\alpha}', IT_i), i=1,\dots,I, j=1,\dots,J$, which is given by

$$Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')=\sum_{i=1}^{I}\sum_{j=1}^{J}Q_{ij}(\boldsymbol{\alpha}|\boldsymbol{\alpha}',\boldsymbol{IT}).$$

Upon substituting the link function for the failure rates specified in (1), and differentiating the objective function with respect to $\alpha_{r0},\alpha_{r1}, r=1,2$, we will obtain the gradient vector $\overrightarrow{G}$, with the first derivatives of $Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')$, as

$$\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{r0}}=\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}}{\alpha_{r0}}-\sum_{i=1}^{I}\sum_{j=1}^{J}\exp(\alpha_{r1}w_j)g_{rij}(\boldsymbol{\alpha}'), \tag{12}$$

$$\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{r1}}=\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}w_j-\sum_{i=1}^{I}\sum_{j=1}^{J}\alpha_{r0}w_j\exp(\alpha_{r1}w_j)g_{rij}(\boldsymbol{\alpha}'), \tag{13}$$

for $j=0,1,\dots,J$. The gradient vector, $\overrightarrow{G}$, is then given by

$$\overrightarrow{G}(\boldsymbol{\alpha}|\boldsymbol{\alpha}')=\left(\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{10}},\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{11}},\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{20}},\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{21}}\right)'. \tag{14}$$

It is evident that the first derivatives involve nonlinear terms, and so we have to find the estimates by numerical methods. To solve likelihood equations, we consider the updating equations

$$\alpha_{r1}^{(h+1)}=\alpha_{r1}^{(h)}-\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}c_j\exp(\alpha_{r1}^{(h)}w_j)g_{rij}(\boldsymbol{\alpha}')}{\sum_{i=1}^{I}\sum_{j=1}^{J}c_jw_j\exp(\alpha_{r1}^{(h)}w_j)g_{rij}(\boldsymbol{\alpha}')}, \tag{15}$$

$$\hat{\alpha}_{r0}=\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}}{\sum_{i=1}^{I}\sum_{j=1}^{J}\exp(\hat{\alpha}_{r1}w_j)g_{rij}(\boldsymbol{\alpha}')}, \tag{16}$$

where r=1,2, and $c_j=\left(w_j-\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}w_j}{\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}}\right)$.

3.3. Inequality constrained least squares estimate as initial value

Determining a good initial value is always a challenging problem. If the initial values are far from the true parameters, the algorithm may take a large number of steps to converge or even may result in divergence. Fan et al. [12] assumed that the reliability under w_j and IT_i should be around the true probabilities p_{0ij},p_{1ij},p_{2ij}, for $i=1,\dots,I$ and $j=1,\dots,J$. Intuitively, the empirical estimates $\hat{p}_{0ij},\hat{p}_{1ij},\hat{p}_{2ij}$ can be defined as $S_{ij}/K_{ij},D_{1ij}/K_{ij},D_{2ij}/K_{ij}$. If one of them is zero, then it will be hard to determine the initial value. Zero frequency problem has been discussed considerably in the literature, and Lee and Cohen [24] suggested using

$$\left(\tilde{p}_{0ij},\tilde{p}_{1ij},\tilde{p}_{2ij}\right)=\left(\frac{S_{ij}+1}{K_{ij}+3},\frac{D_{1ij}+1}{K_{ij}+3},\frac{D_{2ij}+1}{K_{ij}+3}\right) \tag{17}$$

as modified empirical estimates. Recalling (3), we can write $\lambda_{1j}(\boldsymbol{\alpha})+\lambda_{2j}(\boldsymbol{\alpha})$ as $-\ln(p_{0ij})/IT_i$. Then, by (4) and (5), we can rewrite

$$\ln(\alpha_{r0})+\alpha_{r1}w_j=\ln(p_{rij})-\ln(1-p_{0ij})+\ln(-\ln(p_{0ij}))-\ln(IT_i) \tag{18}$$

for $r=1,2$. By replacing p_{rij} by the modified empirical estimates in (17), we can obtain the least squares estimate of $\boldsymbol{\alpha}$ by minimizing the quadratic form

$$S(\boldsymbol{\alpha})=\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}\sum_{r=1}^{2}\left(\hat{y}_{rij}-\ln(\alpha_{r0})-\alpha_{r1}w_j-\ln(IT_i)\right)^2, \tag{19}$$

where

$$\hat{y}_{rij}=\ln(\tilde{p}_{rij})-\ln(1-\tilde{p}_{0ij})+\ln(-\ln(\tilde{p}_{0ij})). \tag{20}$$

By simple algebra, we derive the (weighted) least squares estimates as follows:

$$\hat{\alpha}_{r1}^{\text{LSE}}=\left\{\begin{array}{c}\left(\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}\right)\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}w_j(\hat{y}_{rij}-\ln(IT_i))-\\ \left(\sum_{i=1}^{I}\sum_{j=1}^{J}w_jK_{ij}\right)\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}(\hat{y}_{rij}-\ln(IT_i))\end{array}\right\}$$
$$\times\left\{\left(\sum_{i=1}^{I}\sum_{j=1}^{J}w_j^2K_{ij}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}\right)-\left(\sum_{i=1}^{I}\sum_{j=1}^{J}w_jK_{ij}\right)^2\right\}^{-1},$$

$$\hat{\alpha}_{r0}^{\text{LSE}}=\exp\left\{\begin{array}{c}\left\{\begin{array}{c}\left(\sum_{i=1}^{I}\sum_{j=1}^{J}w_j^2K_{ij}\right)\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}(\hat{y}_{rij}-\ln(IT_i))-\\ \left(\sum_{i=1}^{I}\sum_{j=1}^{J}w_jK_{ij}\right)\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}w_j(\hat{y}_{rij}-\ln(IT_i))\end{array}\right\}\\ \times\left\{\left(\sum_{i=1}^{I}\sum_{j=1}^{J}w_j^2K_{ij}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}\right)-\left(\sum_{i=1}^{I}\sum_{j=1}^{J}w_jK_{ij}\right)^2\right\}^{-1}\end{array}\right\}.$$

If the test is conducted under balanced setting with K_{ij} all being equal for all $i=1,\dots,I$ and $j=1,\dots,J$, we will then obtain the least squares estimates as

$$\hat{\alpha}_{r1}^{\text{LSE},*}=\frac{J\sum_{i=1}^{I}\sum_{j=1}^{J}w_j\hat{y}_{rij}-\sum_{i=1}^{J}w_j\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{y}_{rij}}{IJ\sum_{j=1}^{J}w_j^2-I\left(\sum_{j=1}^{J}w_j\right)^2},$$

$$\hat{\alpha}_{r0}^{\text{LSE},*}=\exp\left\{\frac{1}{IJ}\left(\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{y}_{rij}-\hat{\alpha}_{r1}^{\text{LSE},*}I\sum_{j=1}^{J}w_j-J\sum_{i=1}^{I}\ln(IT_i)\right)\right\},$$

which have been used by Fan et al. [12] as the hyper-parameters in one of their prior distributions for the Bayesian estimation. However, in most accelerated life tests, increasing stress levels will also increase the hazard rate of every component which means the parameters are all positive. The least squares estimates are not guaranteed to be positive, and so may violate this assumption. To have a least squares estimate with non-negative constraints, Liew [25] suggested using Inequality Constrainted Least Squares (ICLS) method: if $\hat{\alpha}_{r1}^{LSE}$ is negative, then the modified estimates are

$$\hat{\alpha}_{r1}^{\text{ICLS}}=0,$$
$$\ln\left(\hat{\alpha}_{r0}^{\text{ICLS}}\right)=\ln\left(\hat{\alpha}_{r0}^{\text{LSE}}\right)+\hat{\alpha}_{r1}^{\text{LSE}}\times\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}^2w_j}{\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}^2},$$

and if K_{ij} are equal, we get

$$\hat{\alpha}_{r1}^{\text{ICLS}}=0,$$
$$\ln\left(\hat{\alpha}_{r0}^{\text{ICLS}}\right)=\ln\left(\hat{\alpha}_{r0}^{\text{LSE}}\right)+\hat{\alpha}_{r1}^{\text{LSE}}\sum_{j=1}^{J}w_j/J.$$

If $\hat{\alpha}_{r1}^{\text{LSE}}$ is positive, we will set ICLS estimates as

$$\hat{\alpha}_{r1}^{\text{ICLS}}=\hat{\alpha}_{r1}^{\text{LSE}},\hat{\alpha}_{r0}^{\text{ICLS}}=\hat{\alpha}_{r0}^{\text{LSE}}, \tag{21}$$

for r=1,2. If $\hat{\alpha}_{r1}^{\text{ICLS}}=0$, we will add a small number to it and make it positive, say, $\hat{\alpha}_{r1}^{\text{ICLS}}=10^{-14}$. We assume that the ICLS estimate is

Table 3
Parameter values used in the simulation of devices with high reliability.

Parameters	Symbols	Values
Risk 1	$\alpha_{10},\ \alpha_{11}$	(0.0005, 0.05)
Risk 2	$\alpha_{20},\ \alpha_{21}$	(5×10^{-5}, 0.08)
Temperature (°C)	$w_1,\ w_2,\ w_3,\ w_4$	(35, 45, 55, 65)
Inspection time (days)	$IT_1,\ IT_2,\ IT_3$	(10, 20, 30)
Sample size	$K_1,\ K_2,\ K_3$	(10, 50, 100)

close to the true parameter so that we can use the initial estimate for α_{rm} in the EM algorithm as $\hat{\alpha}_{rm}^{\text{ICLS}}$, for $r=1,2$ and $m=0,1$.

4. Simulation study

In this section, we will evaluate the performance of the EM algorithm developed in the preceding section. We consider an accelerated life test with $I=3$ inspection times and $J=4$ levels of temperatures, and then repeat the experiment with different sample sizes with K_{ij} devices being allocated to each level of temperature and inspection time. For simplicity, we allocate the same number of devices to each testing condition and denote the number as $K=K_{ij}$ for $i=1,\ldots,I$ and $j=1,\ldots,J$. The setting used in this simulation study is as specified in Table 3. The maximum iteration number is set as 10,000 and the algorithm is stopped when the sum of squares of the difference of the successive parameter estimates is less than 10^{-10}. Since we regard the EM estimate to be an improvement of the initial estimate obtained by ICLS, we can provide a comparison between these two with the ICLS estimate serving as a reference.

Notice that Risk 2 has a larger intercept (0.08) than Risk 1 (0.05). This means that Risk 2, the leakage of organic fuel, is more sensitive to the temperature as compared to Risk 1, the burnout of resistance wire. This simulation setting imitates the higher chances of having cracks with higher temperature while the resistance wire is not so sensitive to change in temperature. Also, Risk 1 has a larger intercept (0.0005) than Risk 2 ($5\times\ 10^{-5}$). This setting mimics the fact that most of the common failures are due to the disconnection of resistance wires than due to ignition. The disconnection may be due to shocks (which is not the type of stress we are interested in) based on regular use, for example. The intercepts are small so that we can say that the products are of high reliability.

Table 4 shows the Bias and MSEs of the MLEs determined from the EM algorithm. From Table 4, we see that the EM algorithm performs very well and that the obtained estimates are much better than the initial estimate given by ICLS. Both Bias and MSE decrease with increasing K. For $K=10$, the Bias is of the size 10^{-3}, while the size of MSE is around 10^{-5}. But, for $K=100$, the Bias is of the size 10^{-4}, while the size of MSE is around 10^{-6}.

We may also be interested in some useful quantities at room temperature (the used condition, 25 °C, such as the Reliability (survival probability at the normal operating condition) p_{0i} at IT_i, the expected lifetime $E(T)$ and the probability of failure due to Cause 1 given failure, P.d1. The formulas for estimating these quantities are all presented in Table 5.

In Table 6, we observe that the EM algorithm improves the initial estimates given by ICLS quite considerably. Both Bias and MSE are much smaller under the EM algorithm than under ICLS. It is also of interest to observe that all the EM estimates are negatively biased. The EM algorithm does not make use of the prior belief that the devices are of high reliability and so the EM algorithm tends to underestimate the survival probabilities and the life expectancies.

Table 4
The Bias and MSE of the MLEs of model parameters of products with high reliability under EM algorithm and ICLS.

	Bias		MSE	
$\alpha_{10}=0.0005$	EM	ICLS	EM	ICLS
$K=10$	6.186e−04	2.57e−03	4.964e−06	1.07e−05
$K=50$	9.567e−05	3.886e−04	1.532e−07	4.122e−07
$K=100$	6.18e−05	1.958e−04	7.045e−08	1.419e−07
$\alpha_{11}=0.05$	EM	ICLS	EM	ICLS
$K=10$	2.148e−03	−2.352e−02	7.85e−04	7.013e−04
$K=50$	−3.452e−05	−6.421e−03	1.293e−04	1.509e−04
$K=100$	−3.603e−04	−3.457e−03	6.937e−05	8.526e−05
$\alpha_{20}=5\times10^{-5}$	EM	ICLS	EM	ICLS
$K=10$	1.59e−03	2.084e−03	1.859e−03	5.727e−06
$K=50$	2.637e−05	2.11e−04	8.126e−09	7.486e−08
$K=100$	1.001e−05	7.759e−05	2.156e−09	1.19e−08
$\alpha_{21}=0.08$	EM	ICLS	EM	ICLS
$K=10$	1.922e−02	−5.248e−02	6.467e−03	2.869e−03
$K=50$	1.565e−03	−2.292e−02	3.573e−04	6.454e−04
$K=100$	1.031e−03	−1.217e−02	1.519e−04	2.516e−04

Table 5
Quantities of interest under room temperature with $\hat{\lambda}_r=\hat{\alpha}_{r0}e^{25\hat{\alpha}_{r1}}$, for $r=1,2$.

Quantity	Estimate
(Reliability) p_{0i}	$\exp(-(\hat{\lambda}_1+\hat{\lambda}_2)IT_i)$
(Expected lifetime) $E(T)$	$(\hat{\lambda}_1+\hat{\lambda}_2)^{-1}$
$P(T_1<T_2 \mid \min(T_1,T_2)<IT)$ (P.d1)	$\hat{\lambda}_1/(\hat{\lambda}_1+\hat{\lambda}_2)$

Table 6
The Bias and MSE of various useful quantities of devices with high reliability under EM algorithm and ICLS.

	Bias		MSE	
$p_{01}=0.9791$	EM	ICLS	EM	ICLS
$K=10$	−5.648e−03	−6.548e−02	3.919e−04	4.64e−03
$K=50$	−1.273e−03	−1.186e−02	4.597e−05	2.012e−04
$K=100$	−8.041e−04	−5.657e−03	2.407e−05	6.274e−05
$p_{02}=0.9586$	EM	ICLS	EM	ICLS
$K=10$	−1.067e−02	−1.236e−01	1.408e−03	1.643e−02
$K=50$	−2.447e−03	−2.303e−02	1.75e−04	7.554e−04
$K=100$	−1.55e−03	−1.101e−02	9.189e−05	2.376e−04
$p_{03}=0.9385$	EM	ICLS	EM	ICLS
$K=10$	−1.512e−02	−1.75e−01	2.865e−03	3.278e−02
$K=50$	−3.527e−03	−3.353e−02	3.746e−04	1.596e−03
$K=100$	−2.242e−03	−1.609e−02	1.973e−04	5.062e−04
$E(T)=472.9$	EM	ICLS	EM	ICLS
$K=10$	1.54e+02	−3.57e+02	4.895e+05	1.281e+05
$K=50$	1.827e+01	−1.552e+02	2.885e+04	3.036e+04
$K=100$	6.498e+00	−8.437e+01	1.326e+04	1.468e+04
P.d1 = 0.8253	EM	ICLS	EM	ICLS
$K=10$	−5.921e−02	−2.581e−01	5.513e−02	7.478e−02
$K=50$	−1.54e−02	−1.202e−01	1.113e−02	2.291e−02
$K=100$	−5.101e−03	−6.232e−02	5.138e−03	9.42e−03

In addition, we may also want to compare the estimation methods at different levels of reliability. For this reason, we simulate data sets for different choices of the parameters and set $(\alpha_{10},\alpha_{11},\alpha_{20},\alpha_{21})=(0.001,0.05,0.0001,0.08)$ which represent devices with moderate reliability, and $(\alpha_{10},\alpha_{11},\alpha_{20},\alpha_{21})=$ (0.005, 0.05, 0.0005, 0.08) which represent devices with low reliability.

Table 7
Parameter values used in simulation of devices with moderate reliability.

Parameters	Symbols	Values
Risk 1	$\alpha_{10},\ \alpha_{11}$	(0.001, 0.05)
Risk 2	$\alpha_{20},\ \alpha_{21}$	(0.0001, 0.08)
Temperature (°C)	$w_1,\ w_2,\ w_3,\ w_4$	(35, 45, 55, 65)
Inspection time (days)	$IT_1,\ IT_2,\ IT_3$	(10, 20, 30)
Sample size	$K_1,\ K_2,\ K_3$	(10, 50, 100)

Table 8
Parameter values used in simulation of devices with low reliability.

Parameters	Symbols	Values
Risk 1	$\alpha_{10},\ \alpha_{11}$	(0.005, 0.05)
Risk 2	$\alpha_{20},\ \alpha_{21}$	(0.0005, 0.08)
Temperature (°C)	$w_1,\ w_2,\ w_3,\ w_4$	(35, 45, 55, 65)
Inspection time (days)	$IT_1,\ IT_2,\ IT_3$	(10, 20, 30)
Sample size	$K_1,\ K_2,\ K_3$	(10, 50, 100)

Keeping other settings the same as in the case of devices with high reliability, the setting for the simulation of products with moderate and low reliabilities are summarized in Tables 7 and 8, respectively. The results obtained from the corresponding simulation studies are presented in Tables 9 and 10, respectively.

From Tables 9 and 10, we observe that the EM algorithm still performs better than the ICLS estimate. Also, the estimates from the EM algorithm have generally lower Bias and lower MSE, compared to those in Table 4. This incidentally agrees with the observation in Balakrishnan and Ling [1] that the EM algorithm performs better for devices with low or moderate reliabilities. Here again, we may be interested in quantities specified in Table 5 and the corresponding results are presented in Tables 11 and 12.

From Tables 11 and 12, we observe that the EM algorithm is still a better estimation method as compared to ICLS. Again, the bias and MSEs are lower in these tables than those in Table 6, which agrees with the result in Balakrishnan and Ling [16]. One reason for the inferior performance of the EM algorithm in case of high reliability is that the MLEs depend solely on the observed data. Samples with devices of high reliability will have small number of observed failures and, therefore, the estimates will not be as accurate as those from samples with devices of moderate or low reliabilities.

Table 9
The Bias and MSE of the parameters of devices with moderate reliability under EM algorithm and ICLS.

	Bias		MSE	
$\alpha_{10}=0.001$	EM	ICLS	EM	ICLS
$K=10$	6.827e−04	2.281e−03	5.487e−06	1.118e−05
$K=50$	8.333e−05	3.434e−04	3.11e−07	5.5e−07
$K=100$	5.051e−05	1.801e−04	1.282e−07	2.176e−07
$\alpha_{11}=0.05$	EM	ICLS	EM	ICLS
$K=10$	8.065e−04	−1.55e−02	4.408e−04	4.119e−04
$K=50$	5.427e−04	−2.993e−03	8.273e−05	9.344e−05
$K=100$	6.514e−05	−1.672e−03	3.623e−05	4.65e−05
$\alpha_{20}=0.0001$	EM	ICLS	EM	ICLS
$K=10$	1.938e−04	1.633e−03	4.118e−07	3.97e−06
$K=50$	2.424e−05	1.678e−04	1.051e−08	5.539e−08
$K=100$	1.524e−05	7.341e−05	4.353e−09	1.445e−08
$\alpha_{21}=0.08$	EM	ICLS	EM	ICLS
$K=10$	5.553e−03	−4.14e−02	1.613e−03	1.857e−03
$K=50$	8.517e−04	−1.309e−02	1.743e−04	2.824e−04
$K=100$	3.394e−05	−6.486e−03	9.604e−05	1.421e−04

Table 10
The Bias and MSE of the parameters of devices with low reliability under EM algorithm and ICLS.

	Bias		MSE	
$\alpha_{10}=0.005$	EM	ICLS	EM	ICLS
$K=10$	9.951e−04	6.748e−03	2.415e−05	8.894e−05
$K=50$	1.507e−04	2.117e−03	2.768e−06	8.543e−06
$K=100$	3.602e−05	1.258e−03	1.318e−06	3.252e−06
$\alpha_{11}=0.05$	EM	ICLS	EM	ICLS
$K=10$	2.272e−03	−1.935e−02	2.594e−04	4.809e−04
$K=50$	5.515e−04	−8.077e−03	4.377e−05	9.405e−05
$K=100$	4.33e−04	−5.177e−03	2.086e−05	4.21e−05
$\alpha_{20}=0.0005$	EM	ICLS	EM	ICLS
$K=10$	2.018e−04	1.985e−03	1.716e−06	7.218e−06
$K=50$	4.182e−05	3.807e−04	8.113e−08	3.287e−07
$K=100$	7.737e−06	1.808e−04	2.9e−08	8.523e−08
$\alpha_{21}=0.08$	EM	ICLS	EM	ICLS
$K=10$	4.908e−03	−2.994e−02	5.136e−04	1.035e−03
$K=50$	7.506e−04	−1.036e−02	8.427e−05	1.781e−04
$K=100$	6.92e−04	−5.949e−03	4.078e−05	7.465e−05

5. Comparison to Fisher-Scoring Method

Another natural comparison for the EM algorithm is with the Fisher-Scoring Method for the MLEs. It requires solving the score functions, $V(\boldsymbol{\theta}) = \partial l(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$, by the use of the updating equation

$$\boldsymbol{\theta}^{(h+1)} = \boldsymbol{\theta}^{(h)} + I_{obs}^{-1}\left(\boldsymbol{\theta}^{(h)}\right) V\left(\boldsymbol{\theta}^{(h)}\right), \tag{22}$$

where $I_{obs}(\boldsymbol{\theta}) = -\partial^2 l(\boldsymbol{\theta})/(\partial\boldsymbol{\theta})^2$ is the observed Fisher information matrix. We can compare this direct maximization method with the EM algorithm by different criteria.

5.1. Convergence, computational time and tolerance

Here we compare the number of divergent cases, the average computational time in seconds used (with R version 3.1.1 with CPU i5 2500 K, 3.30 GHz) and the tolerance of converged cases. The Fisher-Scoring Method was implemented with maximum update size of 0.01. The results obtained are presented in Table 13. Both methods seem to run quite fast with the average time for each iteration being about one-tenth of a second. The computational time for the Fisher-Scoring Method seems to be less than that of the EM algorithm. However, the trade-off is based on the divergence. For small sample sizes, the Fisher-Scoring Method resulted in 103 and 3 divergent cases in high reliability and moderate reliability situations, respectively, while the EM algorithm converged in all the cases. This shows the advantage of the EM algorithm in cases of small samples and high reliability.

Besides, the Fisher-Scoring Method in this case have to be controlled by limiting the update size to be 0.01 in order to stabilize the iterations. However, the EM algorithm does not have any such constraints. This is because the Fisher-Scoring Method requires solving a system of $2R=4$ equations which becomes difficult, while the EM algorithm only requires solving $R=2$ equations in each step. This advantage of the EM algorithm in terms of stability becomes more evident for larger number of competing risks, R.

5.2. Robustness of initial values

If the initial values of the EM algorithm as well as the Fisher-Scoring Method are far removed from the MLEs, the methods may

Table 11
The Bias and MSE of various useful quantities of devices with moderate reliability under EM algorithm and ICLS.

	Bias		MSE	
$p_{01}=0.9586$	EM	ICLS	EM	ICLS
$K=10$	−7.014e−03	−5.862e−02	6.857e−04	4.107e−03
$K=50$	−7.213e−04	−1.001e−02	9.714e−05	2.186e−04
$K=100$	−7.358e−04	−5.088e−03	4.563e−05	8.694e−05
$p_{02}=0.9189$	EM	ICLS	EM	ICLS
$K=10$	−1.276e−02	−1.083e−01	2.402e−03	1.386e−02
$K=50$	−1.286e−03	−1.898e−02	3.542e−04	7.839e−04
$K=100$	−1.365e−03	−9.667e−03	1.67e−04	3.15e−04
$p_{03}=0.8808$	EM	ICLS	EM	ICLS
$K=10$	−1.74e−02	−1.501e−01	4.747e−03	2.635e−02
$K=50$	−1.71e−03	−2.698e−02	7.267e−04	1.582e−03
$K=100$	−1.897e−03	−1.378e−02	3.437e−04	6.421e−04
$E\ (T)=236.4$	EM	ICLS	EM	ICLS
$K=10$	3.825e+01	−1.349e+02	4.346e+04	1.894e+04
$K=50$	9.03e+00	−3.791e+01	3.597e+03	3.49e+03
$K=100$	1.958e+00	−2.007e+01	1.585e+03	1.936e+03
P.d1 = 0.8253	EM	ICLS	EM	ICLS
$K=10$	−4.02e−02	−2.12e−01	3.196e−02	5.525e−02
$K=50$	−1.164e−02	−7.176e−02	6.136e−03	1.147e−02
$K=100$	−5.642e−03	−3.396e−02	2.912e−03	4.714e−03

Table 12
The Bias and MSE of various useful quantities of devices with low reliability under EM algorithm and ICLS.

	Bias		MSE	
$p_{01}=0.8094$	EM	ICLS	EM	ICLS
$K=10$	−3.912e−03	−6.929e−02	3.746e−03	7.845e−03
$K=50$	−8.305e−04	−2.55e−02	6.949e−04	1.34e−03
$K=100$	4.782e−04	−1.526e−02	3.282e−04	5.653e−04
$p_{02}=0.6551$	EM	ICLS	EM	ICLS
$K=10$	−2.588e−03	−1.043e−01	9.192e−03	1.738e−02
$K=50$	−6.495e−04	−3.994e−02	1.804e−03	3.281e−03
$K=100$	1.102e−03	−2.414e−02	8.586e−04	1.419e−03
$p_{03}=0.5303$	EM	ICLS	EM	ICLS
$K=10$	1.194e−03	−1.182e−01	1.301e−02	2.191e−02
$K=50$	4.944e−05	−4.693e−02	2.645e−03	4.529e−03
$K=100$	1.736e−03	−2.865e−02	1.266e−03	2.005e−03
$E(T)=47.29$	EM	ICLS	EM	ICLS
$K=10$	4.6e+00	−1.222e+01	4.512e+02	2.344e+02
$K=50$	7.88e−01	−5.519e+00	5.771e+01	6.486e+01
$K=100$	6.159e−01	−3.521e+00	2.689e+01	3.187e+01
P.d1 = 0.8253	EM	ICLS	EM	ICLS
$K=10$	−5.509e−03	−8.024e−02	8.857e−03	1.355e−02
$K=50$	−2.244e−03	−1.736e−02	1.804e−03	2.451e−03
$K=100$	2.696e−04	−6.343e−03	8.891e−04	1.067e−03

not converge to produce MLEs. Here, we compare the convergence of the algorithm if the initial value is removed from the MLEs. We adopted an approach similar to the one in Park [26]. First, we run the EM algorithm with maximum number of iterations to be 10,000 and assume the resulting estimate, $\hat{\boldsymbol{\alpha}}$, to be the true MLE. Second, we keep $\hat{\alpha}_{r0}, r=1,2$, as the initial estimates of the intercept parameters, $\alpha^*_{r0}, r=1,2$, and then the initial estimates of the slope parameters are set to be

$$\alpha^*_{11}=\hat{\alpha_{11}}+d\cos(2k\pi/N),$$
$$\alpha^*_{21}=\hat{\alpha_{21}}+d\sin(2k\pi/N), \quad (23)$$

where $k=1,\ldots,N$ and d is the distance from the true MLE. We then ran the EM algorithm and the Fisher-Scoring Method on the same data with these initial estimates focusing on small sample size and high reliability. For this purpose, we set $N=60$ and the maximum number of iterations to be 1000. We considered three values of distance, d, namely, 0.01, 0.1 and 1. The total number of cases evaluated in each distance is $N\times 1000=60{,}000$. The number of divergent cases out of these 60,000 that were observed for these two methods is presented in Table 14.

In Table 14, we observe that the number of divergent cases for the Fisher-Scoring Method is much larger than that of the EM algorithm, especially for $d=0.01$ and $d=1$. The number of divergent cases for the EM algorithm increases approximately linearly with d since it requires more number of iterations to reach convergence. For the Fisher-Scoring Method, the large number of divergent cases suggests that the method is quite sensitive to even small changes in the initial values. The EM algorithm seems to be relatively insensitive to the initial values which is in agreement with the findings of Park [26]. However, some divergent cases are still observed for the EM algorithm. It is therefore important to give a good and reasonable initial estimate for the EM algorithm, and we have observed in our study that the ICLS estimate presented in Section 3.3 seems to work well for this purpose.

5.3. Missing information principle

One-shot device testing results in highly censored data, and consequently the amount of missing information is quite large. When the EM algorithm is adopted for the determination of MLEs, we can extract the missing information matrix by using the Missing Information Principle, proposed by Louis [27]. According to this principle, the observed information matrix can be calculated as

$$I_{obs}=I_{complete}-I_{missing}, \quad (24)$$

where $I_{complete}$, I_{obs} and $I_{missing}$ are the complete, observed and missing information matrices, respectively. The complete information matrix can be computed as $I_{complete}=-E\left[\partial^2(l_c(\boldsymbol{\alpha}))/\partial\boldsymbol{\alpha}^2\right]$. The complete likelihood l_c is immediately available from the implementation of the EM algorithm. For the observation information matrix, I_{obs} is the observed Fisher information in the Fisher-Scoring Method. With the use of missing information principle in (24), the missing information matrix can then be calculated easily, and it can be shown that $I_{missing}=-E[\partial^2(f_{rj}(t_{rij}\,|\,\boldsymbol{\Delta},\boldsymbol{\alpha}))/\partial\boldsymbol{\alpha}^2]$; see Balakrishnan and Ling [17], for example. With the knowledge of missing information matrix, experimenters can plan future one-shot device testing experiments with specific stress levels and sample size allocations in order to reduce the missing information to a desired level, for example.

6. Masked causes of failure

Sometimes, the cause of the failure of a device cannot be identified. In this case, the cause is said to be "masked". To adapt our method to this case, we denote the indicator $\Delta_{ijk}=-1$ and modify (2) as follows:

$$\Delta_{ijk}=\begin{cases}-1 & \text{for}\{\min(T_{1ijk},T_{2ijk})\le IT_i\}\cap\{\text{masked}\},\\ 0 & \text{for}\{\min(T_{1ijk},T_{2ijk})\ge IT_i\},\\ 1 & \text{for}\{T_{1ijk}<\min(IT_i,T_{2ijk})\}\cap\{\text{not masked}\},\\ 2 & \text{for}\{T_{2ijk}<\min(IT_i,T_{1ijk})\}\cap\{\text{not masked}\}.\end{cases} \quad (25)$$

For simplicity, we may assume that the occurrence of masked causes is independent of the underlying unobserved causes

Table 13
The comparison of the number of divergent cases, the average computational time and the tolerance of the Fisher-scoring method and the EM algorithm.

Reliability		High		Moderate		Low	
K		FS	EM	FS	EM	FS	EM
10	# of div.	103	0	3	0	0	0
	time(s)	2.924e–02	3.655e–01	2.304e–02	1.037e–01	1.157e–02	3.305e–02
	tolerance	1.319e–11	9.766e–11	1.07e–11	9.552e–11	1.18e–11	8.812e–11
50	# of div.	0	0	0	0	0	0
	time(s)	1.656e–02	1.148e–01	1.297e–02	5.996e–02	7.012e–03	2.289e–02
	tolerance	1.311e–11	9.758e–11	1.304e–11	9.555e–11	1.137e–11	8.604e–11
100	# of div.	0	0	0	0	0	0
	time(s)	1.278e–02	9.414e–02	1.009e–02	4.815e–02	5.728e–03	2.033e–02
	tolerance	1.23e–11	9.725e–11	1.109e–11	9.507e–11	9.417e–12	8.53e–11

Table 14
The number of divergent cases for the EM algorithm and the Fisher-scoring method with different initial values.

Distance (d)	EM	FS
0.01	221	2417
0.1	2133	2379
1	19,031	59,971

responsible for the failure; then,

$$E(T_{rijk}|\{\text{conditions of } T_{1ijk} \text{ and } T_{2ijk}\}\cap\{\text{masking event}\})$$
$$=E(T_{rijk}|\{\text{conditions of } T_{1ijk} \text{ and } T_{2ijk}\}). \tag{26}$$

Hence, the expectations of $T_{rijk}^{(\delta_{ijk})}$, when $\delta_{ijk}=0,1,2$, with masked causes are the same as those without masked causes. Let us denote the number of devices failed with a masked cause in the condition IT_i and w_j by M_{ij}. Moreover, the sum of the number of failures with masked, first and second causes of failure, M_{ij}, D_{1ij}, D_{2ij}, and the number of devices survived S_{ij} will be equal to K_{ij}, the number of devices tested at the condition IT_i and w_j; that is, $K_{ij}=M_{ij}+S_{ij}+D_{1ij}+D_{2ij}$. We derive expectations of T_{rijk} under every value of Δ_{ijk} and these are presented in Table 15.

Then, (8) will be modified as

$$\begin{aligned}E(l_c(\lambda_{1j},\lambda_{2j})|\boldsymbol{\alpha}')=&K_{ij}\left(\log(\lambda_{1j}(\boldsymbol{\alpha}))+\log(\lambda_{2j}(\boldsymbol{\alpha}))\right)\\&-\lambda_{1j}\Big(M_{ij}E(T_{1ijk}^{(-1)}|\boldsymbol{\alpha}')+S_{ij}E(T_{1ijk}^{(0)}|\boldsymbol{\alpha}')\\&+D_{1ij}E(T_{1ijk}^{(1)}|\boldsymbol{\alpha}')+D_{2ij}E(T_{1ijk}^{(2)}|\boldsymbol{\alpha}')\Big)\\&-\lambda_{2j}\Big(M_{ij}E(T_{2ijk}^{(-1)}|\boldsymbol{\alpha}')+S_{ij}E(T_{2ijk}^{(0)}|\boldsymbol{\alpha}')\\&+D_{1ij}E(T_{2ijk}^{(1)}|\boldsymbol{\alpha}')+D_{2ij}E(T_{2ijk}^{(2)}|\boldsymbol{\alpha}')\Big),\end{aligned} \tag{27}$$

and (10) and (11) will correspondingly be modified as

$$\begin{aligned}g_{1ij}(\boldsymbol{\alpha}')=&\Big(M_{ij}E(T_{1ijk}^{(-1)}|\boldsymbol{\alpha}')+S_{ij}E(T_{1ijk}^{(0)}|\boldsymbol{\alpha}')\\&+D_{1ij}E(T_{1ijk}^{(1)}|\boldsymbol{\alpha}')+D_{2ij}E(T_{1ijk}^{(2)}|\boldsymbol{\alpha}')\Big),\end{aligned} \tag{28}$$

$$\begin{aligned}g_{2ij}(\boldsymbol{\alpha}')=&\Big(M_{ij}E(T_{2ijk}^{(-1)}|\boldsymbol{\alpha}')+S_{ij}E(T_{2ijk}^{(0)}|\boldsymbol{\alpha}')\\&+D_{1ij}E(T_{2ijk}^{(1)}|\boldsymbol{\alpha}')+D_{2ij}E(T_{2ijk}^{(2)}|\boldsymbol{\alpha}')\Big),\end{aligned} \tag{29}$$

respectively. We then substitute (27), (28) and (29) back into (9) and follow the same maximization procedure. The EM algorithm for the case of masked causes of failures is thus obtained.

Table 15
The conditional expected values of missing data with masked causes corresponding to different cases.

Case	No. of cases	$E\left(T_{1ijk}^{(\delta_{ijk})}\right)$
$\delta_{ijk}=-1$	M_{ij}	$\frac{1}{\lambda'_{1j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
$\delta_{ijk}=0$	S_{ij}	$IT_i+\frac{1}{\lambda'_{1j}}$
$\delta_{ijk}=1$	D_{1ij}	$\frac{1}{\lambda'_{1j}+\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
$\delta_{ijk}=2$	D_{2ij}	$\frac{1}{\lambda'_{1j}}+\frac{1}{\lambda'_{1j}+\lambda'_{ij}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
Case	No. of cases	$E\left(T_{2ijk}^{(\delta_{ijk})}\right)$
$\delta_{ijk}=-1$	M_{ij}	$\frac{1}{\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
$\delta_{ijk}=0$	S_{ij}	$IT_i+\frac{1}{\lambda'_{2j}}$
$\delta_{ijk}=1$	D_{1ij}	$\frac{1}{\lambda'_{2j}}+\frac{1}{\lambda'_{1j}+\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$
$\delta_{ijk}=2$	D_{2ij}	$\frac{1}{\lambda'_{1j}+\lambda'_{2j}}-\frac{IT_i\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}{1-\exp(-(\lambda'_{1j}+\lambda'_{2j})IT_i)}$

7. A goodness of fit test

We can measure the goodness of fit of the proposed model with a distance-based test statistic of the form

$$M=\max_{ij}(|S_{ij}-\hat{S}_{ij}|,|D_{1ij}-\hat{D}_{1ij}|,|D_{2ij}-\hat{D}_{2ij}|), \tag{30}$$

where $i=1,\dots,I, j=1,\dots,J$, and $(\hat{S}_{ij},\hat{D}_{rij})$ are the expected number of successes and failures based on the proposed model. Such a test has been proposed by Balakrishnan and Ling [16–18] to validate their statistical models in the case of ordinary one-shot device testing. The statistic M quantifies the distance between the fitted model and the observed data. If the model does not fit the data, the distance M will be large. If the model is true, $(S_{ij}, D_{1ij}, D_{2ij})$ should follow the multinomial distribution with probabilities close to $(\hat{p}_{0ij},\hat{p}_{1ij},\hat{p}_{2ij})$ which are obtained from (3), (4) and (5) with parameters replaced by the corresponding estimates. Then, the exact p-value can be found as follows:

$$\begin{aligned}p&-\text{value}\\&=Pr(\max_{ij}(|S_{ij}-\hat{S}_{ij}|,|D_{1ij}-\hat{D}_{1ij}|,|D_{2ij}-\hat{D}_{2ij}|)>M)\\&=1-Pr(|\max_{ij}(|S_{ij}-\hat{S}_{ij}|,|D_{1ij}-\hat{D}_{1ij}|,|D_{2ij}-\hat{D}_{2ij}|)\le M)\end{aligned}$$

$$
\begin{aligned}
&= 1 - Pr(\max(|S_{ij} - \hat{S}_{ij}|, |D_{1ij} - \hat{D}_{1ij}|, |D_{2ij} - \hat{D}_{2ij}|) \le M \text{ for all } i,j) \\
&= 1 - \prod_{i,j} Pr(\max(|S_{ij} - \hat{S}_{ij}|, |D_{1ij} - \hat{D}_{1ij}|, |D_{2ij} - \hat{D}_{2ij}|) \le M) \\
&= 1 - \prod_{i,j} Pr\left(\left\{|S_{ij} - \hat{S}_{ij}| \le M\right\} \cap \left\{|D_{1ij} - \hat{D}_{1ij}| \le M\right\}\right. \\
&\quad \left.\cap \left\{|D_{2ij} - \hat{D}_{2ij}| \le M\right\}\right) \\
&= 1 - \prod_{i,j}\left(\sum_{D_{2ij} = b_{2ij}^{l}}^{b_{2ij}^{u}} \sum_{D_{1ij} = b_{1ij}^{l}(D_{2ij})}^{b_{1ij}^{u}(D_{2ij})} \frac{K_{ij}!}{S_{ij}! D_{1ij}! D_{2ij}!} \hat{p}_{0ij}^{S_{ij}} \hat{p}_{1ij}^{D_{1ij}} \hat{p}_{2ij}^{D_{2ij}}\right), \qquad (31)
\end{aligned}
$$

where $S_{ij} = K_{ij} - D_{1ij} - D_{2ij}$, $b_{1ij}^{l} = \max(0, \lceil \hat{D}_{1ij} - M \rceil, \lceil \hat{D}_{1ij} + \hat{D}_{2ij} - M \rceil - D_{2ij})$, $b_{1ij}^{u}(D_{2ij}) = \min(K_{ij} - D_{2ij}, \lfloor \hat{D}_{1ij} + M \rfloor, \lfloor \hat{D}_{1ij} + \hat{D}_{2ij} + M \rfloor - D_{2ij})$, $b_{2ij}^{l}(D_{2ij}) = \max(0, \lceil \hat{D}_{2ij} - M \rceil)$ and $b_{2ij}^{u} = \min(K_{ij}, \lfloor \hat{D}_{2ij} - M \rfloor)$ for $i = 1, \ldots, I, j = 1, \ldots, J$. If the exact p-value is smaller than the desired level, say, 0.05, then we can conclude that the proposed model does not fit the data well.

8. An illustrative example

For an illustrative example, we apply our model to a dataset described in Lindsey and Ryan [28] (Table 1). The table shows the experimental results conducted by National Center for Toxicological Research in 1974. 3355 out of 24,000 female mice are randomized to a control group ($w=0$) or groups that will be injected with a high dose (150 parts per million) of a known carcinogen, called 2-AAF ($w=1$), to different parts of the body. The inspection times used on the mice are 12, 18 and 33 months and the outcome of mice will be death without tumor (DNT), with tumor (DWT), and sacrificed without tumor (SNT) and with tumor (SWT). In this analysis, we ignore the information about parts of mouse bodies where the drugs were injected. We combine SNT and SWT as the sacrificed group ($r=0$), and denote the cause of DNT as natural death ($r=1$) and the cause of DWT as death due to cancer ($r=2$). These modified data are presented in Table 16 and the estimates of model parameters obtained by the proposed estimation method are presented in Table 17.

From Table 17, we note that the EM estimate of α_{11} is negative which means that the drug will decrease the hazard rate of natural death. The reason for this is that the carcinogenic drug will increase the chance of getting tumor which will of course decrease the chances of death without tumor, meaning that $\text{P.d1}_{w=0} > \text{P.d1}_{w=1}$. The EM estimate of α_{21} is positive suggesting that the drug is indeed carcinogenic. Apparently, there are no significant differences between the life expectancies on mice under the drug and under no drug.

Table 16
The number of mice sacrificed ($r=0$) and died (without tumor $r=1$, with tumor $r=2$) from the ED01 experiment data.

	$\delta_{ijk}=0$	$\delta_{ijk}=1$	$\delta_{ijk}=2$
$IT_1 = 12$			
$w=0$	115	22	8
$w=1$	110	49	16
$IT_2 = 18$			
$w=0$	780	42	8
$w=1$	540	54	26
$IT_3 = 33$			
$w=0$	675	200	85
$w=1$	510	64	51

Table 17
The estimates of various parameters of interest under the EM algorithm and ICLS method.

Parameter	EM	ICLS
α_{10}	6.169e−03	5.295e−03
α_{11}	−1.28e−01	2.219e−02
α_{20}	2.36e−03	1.656e−03
α_{21}	2.477e−01	6.427e−01
$p_{01\mid w=0}$	9.027e−01	9.2e−01
$p_{02\mid w=0}$	8.577e−01	8.824e−01
$p_{03\mid w=0}$	7.547e−01	7.95e−01
$E(T\mid w=0)$	1.172e+02	1.439e+02
$\text{P.d1}_{w=0}$	7.233e−01	7.618e−01
$p_{01\mid w=1}$	9.036e−01	9.023e−01
$p_{02\mid w=1}$	8.589e−01	8.572e−01
$p_{03\mid w=1}$	7.566e−01	7.538e−01
$E(T\mid w=1)$	1.183e+02	1.168e+02
$\text{P.d1}_{w=1}$	6.423e−01	6.322e−01

9. Concluding remarks

In this work, a model for one-shot devices with two competing risks is introduced. EM algorithm has been developed for the estimation of model parameters under exponential distribution for lifetimes. Although the EM algorithm takes more computational time and effort than the Fisher-Scoring Method, the algorithm reduces the problem of solving a system of $2R$ equations into that of a system of R-equations. That makes the estimation robust to the choice of initial estimates and improves the convergence rate. We can also obtain the missing information matrix naturally from the EM algorithm. The performance of the EM algorithm is evaluated through simulations and the obtained results show that the EM algorithm works well for devices with medium and low reliability. The EM algorithm has also been modified to adapt to data with masked causes of failures. Finally, we have applied the algorithm to the ED01 data to illustrate all the inferential methods developed here.

For further study, it will be of interest to compare the performance of the EM algorithm with Bayesian estimation under different prior distributions. Since K_{ij} affects the estimation accuracy, the optimal allocation of K_{ij} can also be determined subjected to some cost and time constraints. It will also be of interest to derive optimal test plans with optimal stress levels and inspection times similar to those done in Sohn [9]. We can also extend the model to R competing risks, where $R > 2$. Furthermore, we may also consider extending the model to dependent competing risks which will be of practical interest. There are many bivariate exponential distributions (see Chapter 47 in Kotz et al. [29]) that could be used for this purpose.

Finally, it will be worthwhile to extend the model to risks with Weibull or gamma lifetimes along the lines of Balakrishnan and Ling [17,18]. However, such an extension is not straightforward and it would involve a lot of work in terms of derivations and numerical algorithms resulting from the change to these lifetime distributions. Even under a simple one-shot device model, Balakrishnan and Ling [17,18] did lengthy and meticulous derivations and calculations. Hence, this will be a challenging task to extend the obtained results to more general lifetime distributions.

One important point in the present work is that the probability of failure due to Cause 1, P.d1, does not depend on Inspection Time IT. This is because of the constant hazard rate of the exponential distribution. However, when we consider Weibull and gamma lifetime distributions, P.d1 will not be a constant over IT anymore. Work on these problems is currently under progress and we hope to report these findings in a future paper.

Acknowledgments

The authors thank the referees for their valuable comments on an earlier version of this manuscript led to this improved version.

Appendix A

A.1. Derivations of conditional expectations

The derivations of formulas in Table 2 are provided in this subsection.

Case	Cause 1	Cause 2
$\delta_{ijk}=0$	$f_{1j}(t_1 \mid T_{1ijk}>IT_i)$	$f_{2j}(t_2 \mid T_{2ijk}>IT_i)$
$\delta_{ijk}=1$	$f_{1j}(t_1 \mid T_{1ijk}<\min(IT_i,T_{2ijk}))$	$f_{2j}(t_2 \mid T_{1ijk}<\min(IT_i,T_{2ijk}))$
$\delta_{ijk}=2$	$f_{1j}(t_1 \mid T_{2ijk}<\min(IT_i,T_{1ijk}))$	$f_{2j}(t_2 \mid T_{2ijk}<\min(IT_i,T_{1ijk}))$

First, we have

$$\begin{aligned}E(T_{1ijk}\mid T_{1ijk}>IT_i)&=\frac{\int_{IT_i}^{\infty} t\lambda_{1j}\exp(-\lambda_{1j}t)\,dt}{\exp(-\lambda_{1j}IT_i)}\\&=\exp(\lambda_{1j}IT_i)\left([-t\exp(-\lambda_{1j}t)]_{IT_i}^{\infty}+\int_{IT_i}^{\infty}\exp(-\lambda_{1j}t)\,dt\right)\\&=\exp(\lambda_{1j}IT_i)\left(IT_i\exp(-\lambda_{1j}IT_i)+\frac{\exp(-\lambda_{1j}IT_i)}{\lambda_{1j}}\right)\\&=IT_i+\frac{1}{\lambda_{1j}}.\end{aligned}$$

Similarly, we get

$$E(T_{2ijk}\mid T_{2ijk}>IT_i)=IT_i+\frac{1}{\lambda_{2j}}.$$

The following formula is useful in our derivations:

$$\begin{aligned}\int_0^A t\lambda\exp(-\lambda t)\,dt&=[-t\exp(-\lambda t)]_0^A+\int_0^A\exp(-\lambda t)\,dt\\&=-A\exp(-\lambda A)+\frac{1-\exp(-\lambda A)}{\lambda}.\end{aligned}$$

The probability for the case $\Delta=1$:

$$\begin{aligned}P(T_{1ijk}<\min(IT_i,T_{2ijk}))&=\int_0^{\infty}\int_0^{\min(IT_i,t_2)} f_{1j}(t_1)f_{2j}(t_2)\,dt_1dt_2\\&=\int_0^{IT_i}\int_0^{t_2}\lambda_{1j}\exp(-\lambda_{1j}t_1)\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_1\,dt_2\\&\quad+\int_0^{IT_i}\lambda_{1j}\exp(-\lambda_{1j}t_1)\,dt_1\int_{IT_i}^{\infty}\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_2\\&=\int_0^{IT_i}(1-\exp(-\lambda_{ij}t_2))\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_2\\&\quad+\int_0^{IT_i}\lambda_{1j}\exp(-\lambda_{1j}t_1)\,dt_1\int_{IT_i}^{\infty}\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_2\\&=1-\exp(-\lambda_{2j}IT_i)-\left(\frac{\lambda_{2j}}{\lambda_{1j}+\lambda_{2j}}\right)(1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i))\\&\quad+\exp(-\lambda_{2j}IT_i)-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)\\&=\left(\frac{\lambda_{1j}}{\lambda_{1j}+\lambda_{2j}}\right)(1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i))\\&=P(T_{1ijk}<T_{2ijk})P(\min(T_{1ijk},T_{2ijk})<IT_i).\end{aligned}$$

Similarly, we find

$$\begin{aligned}P(T_{2ijk}<\min(IT_i,T_{1ijk}))&=\left(\frac{\lambda_{2j}}{\lambda_{1j}+\lambda_{2j}}\right)(1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i))\\&=P(T_{2ijk}<T_{1ijk})P(\min(T_{1ijk},T_{2ijk})<IT_i).\end{aligned}$$

We then find

$$\begin{aligned}&E(T_{1ijk}\mid T_{1ijk}<\min(IT_i,T_{2ijk}))\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{\infty}\int_0^{\min(IT_i,t_2)} t_1f_{1j}(t_1)f_{2j}(t_2)\,dt_1dt_2\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{IT_i}\int_0^{t_2}\lambda_{1j}t_1\exp(-\lambda_{1j}t_1)\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_1dt_2\\&+\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{IT_i}\lambda_{1j}t_1\exp(-\lambda_{1j}t_1)\,dt_1\int_{IT_i}^{\infty}\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_2\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{IT_i}-\lambda_{2j}t_2\exp(-(\lambda_{1j}+\lambda_{2j})t_2)+\frac{\lambda_{2j}}{\lambda_{1j}}\exp(-\lambda_{2j}t_2)dt_2\\&-\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{IT_i}\frac{\lambda_{2j}}{\lambda_{1j}}\exp(-(\lambda_{1j}+\lambda_{2j})t_2)dt_2\\&-\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)\\&+\frac{\exp(-\lambda_{2j}IT_i)-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}{\lambda_{1j}P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\left[\frac{\lambda_{2j}}{\lambda_{1j}+\lambda_{2j}}\left(IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)\right.\right.\\&\left.-\frac{1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}{\lambda_{1j}+\lambda_{2j}}\right).\\&+\frac{1}{\lambda_{1j}}(1-\exp(-\lambda_{2j}IT_i))\\&-\frac{\lambda_{1j}}{\lambda_{1j}(\lambda_{1j}+\lambda_{2j})}(1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i))\\&\left.-IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)+\frac{1}{\lambda_{1j}}\exp(-\lambda_{2j}IT_i)-\frac{1}{\lambda_{1j}}\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)\right]\\&=\frac{1}{\lambda_{1j}+\lambda_{2j}}-\frac{IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}{1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}.\end{aligned}$$

Similarly, we find

$$E(T_{2ijk}\mid T_{2ijk}<\min(IT_i,T_{1ijk}))=\frac{1}{\lambda_{1j}+\lambda_{2j}}-\frac{IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}{1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}$$

and

$$\begin{aligned}&E(T_{2ijk}\mid T_{1ijk}<\min(IT_i,T_{2ijk}))\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{\infty}\int_0^{\min(IT_i,t_2)} t_2f_{1j}(t_1)f_{2j}(t_2)\,dt_1dt_2\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{IT_i}\int_0^{t_2}\lambda_{1j}\exp(-\lambda_{1j}t_1)\,dt_1\lambda_{2j}t_2\exp(-\lambda_{2j}t_2)\,dt_2\\&+\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{IT_i}\lambda_{1j}\exp(-\lambda_{1j}t_1)\,dt_1\int_{IT_i}^{\infty}t_2\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_2\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\int_0^{IT_i}(1-\exp(-\lambda_{1j}t_2))t_2\lambda_{2j}\exp(-\lambda_{2j}t_2)\,dt_2\\&+\left(IT_i+\frac{1}{\lambda_{2j}}\right)\frac{\exp(-\lambda_{2j}IT_i)(1-\exp(-\lambda_{1j}IT_i))}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\\&=\frac{1}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\left[\frac{1-\exp(-\lambda_{2j}IT_i)}{\lambda_{2j}}-IT_i\exp(-\lambda_{2j}IT_i)\right.\\&\left.-\frac{\lambda_{2j}}{\lambda_{1j}+\lambda_{2j}}\left(\frac{1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}{\lambda_{1j}+\lambda_{2j}}-IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)\right)\right]\\&+\left(IT_i+\frac{1}{\lambda_{2j}}\right)\frac{\exp(-\lambda_{2j}IT_i)(1-\exp(-\lambda_{1j}IT_i))}{P(T_{1ijk}<\min(IT_i,T_{2ijk}))}\\&=\frac{1}{\lambda_{2j}}+\frac{1}{\lambda_{1j}+\lambda_{2j}}-\frac{IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}{1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}.\end{aligned}$$

We also analogously find

$$E(T_{1ijk}\mid T_{2ijk}<\min(IT_i,T_{1ijk}))=\frac{1}{\lambda_{1j}}+\frac{1}{\lambda_{1j}+\lambda_{2j}}-\frac{IT_i\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}{1-\exp(-(\lambda_{1j}+\lambda_{2j})IT_i)}.$$

Thus, we obtain all the expressions presented in Table 2. Next, for the case of masked causes, we find

$$\begin{aligned}E(T_1|\Delta=-1)&=E(T_1|\min(T_1,T_2)\le IT)\\&=\frac{1}{P(\min(T_1,T_2)\le IT)}\\&\quad\times\int\int_{\min(T_1,T_2)<IT}t_1\exp(-\lambda_1t_1)t_2\exp(-\lambda_2t_2)\,dt_1dt_2\\&=\frac{1}{P(\min(T_1,T_2)\le IT)}\\&\quad\times\left(\int_0^\infty\int_0^\infty t_1\exp(-\lambda_1t_1)t_2\exp(-\lambda_2t_2)\,dt_1dt_2-\right.\\&\quad\left.\int\int_{\min(T_1,T_2)\ge IT}t_1\exp(-\lambda_1t_1)t_2\exp(-\lambda_2t_2)\,dt_1dt_2\right)\\&=\frac{1}{P(\min(T_1,T_2)\le IT)}\times\\&\quad\left(\frac{1}{\lambda_1}-\exp(-\lambda_2IT)\int_{IT}^\infty t_1\exp(-\lambda_1t_1)\,dt_1\right)\\&=\frac{1}{1-\exp(-(\lambda_1+\lambda_2)IT)}\\&\quad\times\left[\frac{1}{\lambda_1}-\left(IT+\frac{1}{\lambda_1}\right)\exp(-(\lambda_1+\lambda_2)IT)\right]\\&=\frac{1}{\lambda_1}-\frac{IT\exp(-(\lambda_1+\lambda_2)IT)}{1-\exp(-(\lambda_1+\lambda_2)IT)}.\end{aligned}$$

Similarly, we find

$$E(T_2|\Delta=-1)=E(T_2|\min(T_1,T_2)\le IT)=\frac{1}{\lambda_2}-\frac{IT\exp(-(\lambda_1+\lambda_2)IT)}{1-\exp(-(\lambda_1+\lambda_2)IT)}.$$

These are precisely the formulas presented in Table 15.

A.2. Derivation of updating equations for maximization

In Section 3.2, the updating equations are based on Newton's method and their derivation is presented here. Consider solving the gradient equations

$$\left.\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{r0}}\right|_{\hat\alpha_{r0},\hat\alpha_{r1}}=\left.\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}')}{\partial\alpha_{r1}}\right|_{\hat\alpha_{r0},\hat\alpha_{r1}}=0,$$

for $r=1,2$, and they are as given in (12) and (13). We can rewrite them as

$$\hat\alpha_{r0}=\frac{\sum_{i=1}^I\sum_{j=1}^J K_{ij}}{\sum_{i=1}^I\sum_{j=1}^J\exp(\hat\alpha_{r1}w_j)g_{rij}(\boldsymbol{\alpha}')},$$

$$\sum_{i=1}^I\sum_{j=1}^J K_{ij}w_j=\sum_{i=1}^I\sum_{j=1}^J\hat\alpha_{r0}w_j\exp(\hat\alpha_{r1}w_j)g_{rij}(\boldsymbol{\alpha}'),$$

and reduce them to

$$\sum_{i=1}^I\sum_{j=1}^J c_j\exp(\hat\alpha_{r1}w_j)g_{rij}(\boldsymbol{\alpha}')=0,$$

where

$$c_j=\left(w_j-\frac{\sum_{i=1}^I\sum_{j=1}^J K_{ij}w_j}{\sum_{i=1}^I\sum_{j=1}^J K_{ij}}\right)$$

and $r=1,2$. We solve the equation by using Newton's method with the updating formula for the h th iteration as

$$\alpha_{r1}^{(h+1)}=\alpha_{r1}^{(h)}-\frac{\sum_{i=1}^I\sum_{j=1}^J c_j\exp(\alpha_{r1}^{(h)}w_j)g_{rij}(\boldsymbol{\alpha}')}{\sum_{i=1}^I\sum_{j=1}^J c_jw_j\exp(\alpha_{r1}^{(h)}w_j)g_{rij}(\boldsymbol{\alpha}')}.$$

After obtaining the converged $\hat\alpha_{r1}$, we obtain

$$\hat\alpha_{r0}=\frac{\sum_{i=1}^I\sum_{j=1}^J K_{ij}}{\sum_{i=1}^I\sum_{j=1}^J\exp(\hat\alpha_{r1}w_j)g_{rij}(\boldsymbol{\alpha}')}.$$

A.3. Derivations of ICLS

In Section 3.3, we adopted the inequality constrained least squares (ICLS) method proposed by Liew [25], which provides solution to the following constrained optimization problem:

$$\min_b Z=\frac{1}{2}(y-Xb)'(y-Xb), \tag{32}$$

subjected to

$$Ab\ge c. \tag{33}$$

Let $\hat\beta^{\mathrm{LSE}}=(X'X)^{-1}X'y$ be the least squares estimate to (32) without any constraint. If $\hat\beta^{LSE}$ satisfies the constraint stated in the inequality (33), then the solution to the constrained problem will be $\hat\beta^{ICLS}=\hat\beta^{\mathrm{LSE}}$.

Otherwise,

$$\hat\beta^{ICLS}=\hat\beta^{\mathrm{LSE}}+(X'X)^{-1}A'(A(X'X)^{-1}A')^{-1}(c-A\hat\beta^{\mathrm{LSE}}). \tag{34}$$

For the constraint that $\alpha_{r0}\ge0,\alpha_{r1}\ge0,r=1,2$, the constraint for the least squares problem will be $\alpha_{r1}\ge0$ but with no constraint on $\ln(\alpha_{r0})$. Then,

$$A=(1\quad 0),\quad b=\begin{pmatrix}\alpha_{r1}\\ \ln(\alpha_{r0})\end{pmatrix},\quad c=0,\quad X=\{\boldsymbol{Kw},\boldsymbol{K}\},$$

where

$$\boldsymbol{K}=(K_{11},K_{21},\ldots,K_{I1},K_{12},\ldots,K_{IJ})^t$$

and

$$\boldsymbol{Kw}=(K_{11}w_1,K_{21}w_1,\ldots,K_{I1}w_1,K_{12}w_2,\ldots,K_{IJ}w_J).$$

After some simple algebra, we get

$$\hat\alpha_{r1}^{\mathrm{ICLS}}=0,$$

$$\ln\left(\hat\alpha_{r0}^{\mathrm{ICLS}}\right)=\ln\left(\hat\alpha_{r0}^{\mathrm{LSE}}\right)+\hat\alpha_{r1}^{\mathrm{LSE}}\times\frac{\sum_{i=1}^I\sum_{j=1}^J K_{ij}^2w_j}{\sum_{i=1}^I\sum_{j=1}^J K_{ij}^2},$$

and if K_{ij} are equal, we get

$$\hat\alpha_{r1}^{\mathrm{ICLS}}=0,$$

$$\ln\left(\hat\alpha_{r0}^{\mathrm{ICLS}}\right)=\ln\left(\hat\alpha_{r0}^{\mathrm{LSE}}\right)+\hat\alpha_{r1}^{LSE}\sum_{j=1}^J w_j/J.$$

A.4. Derivation of the exact p-value

In Section 7, boundaries for the summations can be derived as follows. Given i,j, $1\le i\le I$, $1\le j\le J$, the constraint

$$\max(|S_{ij}-\hat S_{ij}|,|D_{1ij}-\hat D_{1ij}|,|D_{2ij}-\hat D_{2ij}|)\le M \tag{35}$$

implies

$$|S_{ij}-\hat S_{ij}|\le M,|D_{1ij}-\hat D_{1ij}|\le M,|D_{2ij}-\hat D_{2ij}|\le M. \tag{36}$$

The boundaries for D_{1ij} and D_{2ij} under the constraint are

$$\begin{aligned}&\lceil\hat S_{ij}-M\rceil\le S_{ij}\le\lfloor\hat S_{ij}+M\rfloor\\&\quad\Rightarrow\lceil\hat D_{1ij}+\hat D_{2ij}-M\rceil\le D_{1ij}+D_{2ij}\le\lfloor\hat D_{1ij}+\hat D_{2ij}+M\rfloor,\end{aligned} \tag{37}$$

$$\lceil\hat D_{1ij}-M\rceil\le D_{1ij}\le\lfloor\hat D_{1ij}+M\rfloor, \tag{38}$$

$$\lceil\hat D_{2ij}-M\rceil\le D_{2ij}\le\lfloor\hat D_{2ij}+M\rfloor. \tag{39}$$

The usual constraints on D_{1ij} and D_{2ij} are

$$0\le D_{1ij}+D_{2ij}\le K_{ij}, \tag{40}$$

$$0\le D_{1ij}\le K_{ij}, \tag{41}$$

$$0 \le D_{2ij} \le K_{ij}. \tag{42}$$

Now, given D_{2ij}, the boundary for D_{1ij} is

$$0 \le D_{1ij} \le K_{ij} - D_{2ij}, \tag{43}$$

$$\lceil \hat{D}_{1ij} + \hat{D}_{2ij} - M \rceil - D_{2ij} \le D_{1ij} \le \lfloor \hat{D}_{1ij} + \hat{D}_{2ij} + M \rfloor - D_{2ij}. \tag{44}$$

Thus, the range for D_{1ij} is

$$\max(0, \lceil \hat{D}_{1ij} - M \rceil, \lceil \hat{D}_{1ij} + \hat{D}_{2ij} - M \rceil - D_{2ij}) \le D_{1ij} \le \min(K_{ij} - D_{2ij}, \lfloor \hat{D}_{1ij} + M \rfloor, \lfloor \hat{D}_{1ij} + \hat{D}_{2ij} + M \rfloor - D_{2ij}), \tag{45}$$

and the range for D_{2ij} is

$$\max(0, \lceil \hat{D}_{2ij} - M \rceil) \le D_{2ij} \le \min(K_{ij}, \lfloor \hat{D}_{2ij} + M \rfloor). \tag{46}$$

Since the range for D_{1ij} depends on D_{2ij}, the probability should be

$$Pr(\max(|S_{ij} - \hat{S}_{ij}|, |D_{1ij} - \hat{D}_{1ij}|, |D_{2ij} - \hat{D}_{2ij}|) \le M) = \sum_{D_{2ij} = b^l_{2ij}}^{b^u_{2ij}} \sum_{d1 = b^l_{1ij}(D_{2ij})}^{b^u_{1ij}(D_{2ij})} \frac{K_{ij}!}{S_{ij} D_{1ij}! D_{2ij}!} p_{0_{ij}}^{S_{ij}} p_{1ij}^{D_{1ij}} p_{2ij}^{D_{2ij}}, \tag{47}$$

where $b^l_{1ij}(D_{2ij}) = \max(0, \lceil \hat{D}_{1ij} - M \rceil, \lceil \hat{D}_{1ij} + \hat{D}_{2ij} - M \rceil - D_{2ij})$, $b^u_{1ij}(D_{2ij}) = \min(K - D_{2ij}, \lfloor \hat{D}_{1ij} + M \rfloor, \lfloor \hat{D}_{1ij} + \hat{D}_{2ij} + M \rfloor - D_{2ij})$, $b^l_{2ij} = \max(0, \lceil \hat{D}_{2ij} - M \rceil)$ and $b^u_{2ij} = \min(K_{ij}, \lfloor \hat{D}_{2ij} + M \rfloor)$.

For example, suppose $K_{ij} = 10, \hat{D}_{1ij} = 2.1, \hat{D}_{2ij} = 3.5, M = 1.7$, then $S_{ij} = K_{ij} - \hat{D}_{1ij} - \hat{D}_{2ij} = 4.4$. The range for D_{2ij} is

$$\max(0, \lceil \hat{D}_{2ij} - M \rceil) = 2 \le D_{2ij} \le \min(K_{ij}, \lfloor \hat{D}_{2ij} + M \rfloor) = 5, \tag{48}$$

and for a given D_{2ij}, the range for D_{1ij} is

$$\max(0, 1, 4 - D_{2ij}) \le D_{1ij} \le \min(10 - D_{2ij}, 3, 7 - D_{2ij}). \tag{49}$$

We summarize these numbers in the following table:

D_{2ij}	$\lvert D_{2ij} - \hat{D}_{2ij} \rvert$	D_{1ij}	$\lvert D_{1ij} - \hat{D}_{1ij} \rvert$	S_{ij}	$\lvert S_{ij} - \hat{S}_{ij} \rvert$
2	1.5	2, 3	0.1, 0.9	6, 5	1.6, 0.6
3	0.5	1, 2, 3	1.1, 0.1, 0.9	6, 5, 4	1.6, 0.6, 0.4
4	0.5	1, 2, 3	1.1, 0.1, 0.9	5, 4, 3	0.6, 0.4, 1.4
5	1.5	1, 2	1.1, 0.1	4, 3	0.4, 1.4 .

Therefore, we can the probability in (47) in the summation form as

$$Pr(\max(|S_{ij} - 4.4|, |D_{1ij} - 2.1|, |D_{2ij} - 3.5|) \le 1.7) = \sum_{D_{2ij} = 2}^{5} \sum_{D_{1ij} = b^l(D_{2ij})}^{b^u(D_{2ij})} \frac{K_{ij}!}{S_{ij}! D_{1ij}! D_{2ij}!} p_{0_{ij}}^{S_{ij}} p_{1ij}^{D_{1ij}} p_{2ij}^{D_{2ij}}, \tag{50}$$

where

$$b^l(D_{2ij}) \begin{cases} 2 & \text{if } D_{2ij} = 2 \\ 1 & \text{if } D_{2ij} = 3, 4, 5 \end{cases},$$

and

$$b^u(D_{2ij}) \begin{cases} 3 & \text{if } D_{2ij} = 2, 3, 4 \\ 2 & \text{if } D_{2ij} = 5 \end{cases}.$$

References

[1] Balakrishnan N, Ling MH. EM algorithm for one-shot device testing under the exponential distribution. Comput Stat Data Anal 2012;56(3):502–9.

[2] Meeter CA, Meeker WQ. Optimum accelerated life tests with a nonconstant scale parameter. Technometrics 1994;36(1):71–83.

[3] Meeker WQ, Escobar LA, Lu CJ. Accelerated degradation tests: modeling and analysis. Technometrics 1998;40(2):89–99.

[4] Ng HKT, Chan PS, Balakrishnan N. Estimation of parameters from progressively censored data using EM algorithm. Comput Stat Data Anal 2002;39(4):371–86.

[5] Nandi S, Dewan I. An EM algorithm for estimating the parameters of bivariate Weibull distribution under random censoring. Comput Stat Data Anal 2010;54(6):1559–69.

[6] Pal S. Likelihood-based inferential methods for some flexible cure rate models [Ph.D. thesis]. McMaster University, Hamilton, Canada; 2013.

[7] Park C. Parameter estimation of incomplete data in competing risks using the EM algorithm. IEEE Trans Reliab 2005;54(2):282–90.

[8] Craiu RV, Duchesne T. Inference based on the EM algorithm for the competing risks model with masked causes of failure. Biometrika 2004;91(3):543–58.

[9] Sohn SY. Accelerated life-tests for intermittent destructive inspection, with logistic failure-distribution. IEEE Trans Reliab 1997;46(1):122–9.

[10] Yates S, Mosleh A. A Bayesian approach to reliability demonstration for aerospace systems. Reliab Maintainab Symp 2006:611–7.

[11] Newby M. Monitoring and maintenance of spares and one shot devices. Reliab Eng Syst Saf 2008;93(4):588–94.

[12] Fan TH, Balakrishnan N, Chang CC. The Bayesian approach for highly reliable electro-explosive devices using one-shot device testing. J Stat Comput Simul 2009;79(9):1143–54.

[13] Fan TH, Chang CC. A Bayesian zero-failure reliability demonstration test of high quality electro-explosive devices. Quality Reliab Eng Int 2009;25(8):913–20.

[14] Chuang HY. Accelerated failure tests of electro-explosive devices under Weibull lifetime distributions [Ph.D. thesis]. Jhongli City, Taiwan: National Central University; 2011.

[15] Yang GC. A Bayesian reliability analysis of accelerated degradation test for destructive devices [Ph.D. thesis]. Jhongli City, Taiwan: National Central University.

[16] Balakrishnan N, Ling MH. Multiple-stress model for one-shot device testing data under exponential distribution. IEEE Trans Reliab 2012;61(3):809–21.

[17] Balakrishnan N, Ling MH. Expectation maximization algorithm for one shot device accelerated life testing with Weibull lifetimes, and variable parameters over stress. IEEE Trans Reliab 2013;62(2):537–51.

[18] Balakrishnan N, Ling MH. Gamma lifetimes and one-shot device testing analysis. Reliab Eng Syst Saf 2014;126:54–64.

[19] Martz H, Wailer R, Fickas E. Bayesian reliability analysis of series systems of binomial subsystems and components. Technometrics 1988;30(2):143–54.

[20] Martz H, Wailer R. Bayesian reliability analysis of complex series/parallel systems of binomial subsystems and components. Technometrics 1990;32(4):407–16.

[21] Thomas WB, Betts RE. Electroexplosive device; 1967.

[22] Casella G, Berger R. Statistical inference. Pacific Grove: Thomson; 2002.

[23] McLachlan GJ, Krishnan T. EM algorithm and its extensions. Hoboken, New Jersey: John Wiley & Sons; 2007.

[24] Lee HL, Cohen MA. A multinomial logit model for the spatial distribution of hospital utilization. J Bus Econ Stat 1985;3(2):159–68.

[25] Liew CK. Inequality constrained least-squares estimation. J Am Stat Assoc 1976;71(355):746–51.

[26] Park C. Parameter estimation from load-sharing system data using the expectation-maximization algorithm. IIE Trans 2013;45(2):147–63.

[27] Louis T. Finding the observed information matrix when using the EM algorithm. J R Stat Soc Ser B 1982;44:226–33.

[28] Lindsey J, Ryan L. A three-state multiplicative model for rodent tumorigenicity experiments. J R Stat Soc Ser C 1993;42(2):283–300.

[29] Kotz S, Balakrishnan N, Johnson NL. Continuous multivariate distributions. Models and applications. Second edition. New York: John Wiley & Sons; 2000 vol. 1.