Optimal step-stress test under Type-I censoring for multivariate exponential distribution

Qiang Guan a,b,*, Yincai Tang a

a School of Finance and Statistics, East China Normal University, Shanghai 200241, China
b Department of Mathematics and Computer Science, Sanming University, Sanming 365004, China

ARTICLE INFO

Received 23 April 2011
Received in revised form 22 November 2011
Accepted 14 February 2012
Available online 22 February 2012

Keywords:
Marshall–Olkin multivariate exponential distribution
D-optimality
V-optimality
Fisher information
Accelerated life-testing
Type-I censoring

ABSTRACT

In this paper, we study a k-step-stress accelerated life test under Type-I censoring. The lifetime of the items follows the multivariate exponential distribution and a cumulative exposure model is considered. We derive the maximum likelihood estimators of the model parameters and establish the asymptotic properties of them. The problem of choosing the optimal time is addressed by using V-optimality as well as D-optimality criteria. Finally, some numerical studies are discussed to illustrate the proposed procedures.

1. Introduction

Life test for highly reliable products is often time-consuming as well as expensive under normal operating conditions since it would take a long period of time to produce a reasonable number of failures for the required analysis. In these situations, standard life-testing methods are not suitable. This difficulty is overcome by accelerated life test (ALT). In order to speed up the testing procedure and to reduce the failure time, ALT allows the experimenter to apply more severe stresses to obtain information from product life more quickly than under normal operating conditions. Then the data from such an ALT will be transformed to estimate the distribution of failure time under usual conditions. Some key references in the ALT include Nelson (1990) and Meeker and Escobar (1998).

One popular ALT is the step-stress ALT (SSALT), where the stress putting on the testing units changes at pre-specified times or upon the occurrence of a fixed number of failures (Nelson, 1990). The former is called SSALT with Type-I censoring and the latter is called SSALT with Type-II censoring. The statistical inferences in this SSALT for exponential distribution have been studied by many authors such as Miller and Nelson (1983), Xiong (1998), Balakrishnan and Xie (2007a,b), Balakrishnan (2009), Fan et al. (2008), and Wang (2010).

Abbreviation: ALT, accelerated life test; AsVar, asymptotic variance; Cdf, cumulative distribution function; MTTF, mean time to failure; MLE, maximum likelihood estimator; Pdf, probability density function; SSALT, step-stress accelerated life test.

The research is supported by Natural Science Foundation of China (10571057).

* Corresponding author at: School of Finance and Statistics, East China Normal University, Shanghai 200241, China.
E-mail address: yctang@stat.ecnu.edu.cn (Y. Tang).

0378-3758/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2012.02.029
During the past two decades, the problem of optimal scheduling of the step-stress test has attracted great attention in the reliability literature. Miller and Nelson (1983) initiated research in this area by assuming complete failure data under two stress levels (simple step-stress) and the assumption of exponential lifetimes. Extension of the above results to the time-censored and the three stress level cases were obtained by Bai et al. (1989) and Khamis and Higgins (1996), respectively. For the general k-level and M-variable case, some numerical investigations was undertaken by Khamis (1997). For a k-step-stress ALT with equal duration steps, Gouno et al. (2004) and Balakrishnan and Han (2009) tackled the problem of determining the optimal stress change points when the available data are exponential progressively Type-I censored. The problem of optimal issues with the cumulative exposure model (Miller and Nelson, 1983) under exponentially distributed life has been studied by several authors such as Balakrishnan et al. (2009), and Wu et al. (2008).

In electrical engineering or survival analysis study, it is quite common that more than one cause of failure for system. In statistical literature, this process is known as the competing risks model. In the analysis of competing risks models, it is assumed that data consist of failure times, and indicators denoting the cause of failure. Competing risks models are studied in the statistical literature, this process is known as the competing risks model. In the analysis of competing risks models, it is assumed that data consist of failure times, and indicators denoting the cause of failure. Competing risks models are studied in the statistical literature by several authors such as Bai and Chun (1991) and Sarhan (2007). Most of these papers assumed that failures occur from any one of p statistically independent causes.

However, for some more complex products, the causes of failure are dependent. To derive the life distribution of such product is very difficult. Fortunately, Marshall and Olkin (1967) proposed a multivariate exponential distribution to depict the life distribution of this product. Several papers have been written on analyzing the multivariate exponential distribution, such as Sarkar (1987), Joe (1990) and Li (2005). These researches focus on the characteristics and properties of the multivariate exponential distribution. Unfortunately, SSALT planning for the multivariate exponential distribution has got little attention in the literature.

In practice, the assumption of exponential distribution for some product life is unsuitable. Moreover, the assumption of equal duration steps may be not the best for k-step-stress life test planning. Therefore, in this paper, we consider unequal duration steps accelerated life tests for products with Marshall–Olkin multivariate exponential distribution. The main focus of this paper is to investigate the choice of optimal change points of the stress levels. Under Type-I censoring, optimum time points are obtained by using V-optimality criterion and D-optimality criterion.

The rest of this paper is organized as follows. In Section 2, we describe the model and some necessary assumptions. In Section 3, we obtain some important lemmas for later sections. In Section 4, the problem of choosing the optimal change times will be obtained by using two different optimization criteria. In Section 5, some numerical studies were conducted in order to investigate the existence of the optimal stress change points. Finally, some concluding remarks are made in Section 6.

2. Model description and assumptions

2.1. Marshall–Olkin multivariate exponential distribution

Before introducing the multivariate exponential distribution model, let us consider the simplest model that is bivariate Marshall–Olkin exponential distribution model. The model assumes that a two-component system fail after receiving a fatal shock, the occurrences of which are governed by the independent Poisson processes $Z_1(t; \lambda_1), Z_2(t; \lambda_2)$ and $Z_{12}(t; \lambda_{12})$ and make components 1, 2 or both fail, respectively. Thus if $X_1$ and $X_2$ denote the lifetime of the first and second component, then the distribution of residual life is

$$F(x_1, x_2) = P(X_1 > x_1, X_2 > x_2) = P(Z_1(x_1; \lambda_1) = 0, Z_2(x_2; \lambda_2) = 0, Z_{12}(\max(x_1, x_2); \lambda_{12}) = 0) = \exp(-\lambda_1x_1 - \lambda_2x_2 - \lambda_{12} \max(x_1, x_2)).$$

This distribution was called bivariate Marshall–Olkin exponential distribution (Marshall–Olkin, 1967).

We consider an extension of the fatal shock model for an m component system. Let the independent Poisson processes $Z_1(t; \lambda_1), Z_2(t; \lambda_2), \ldots, Z_m(t; \lambda_m)$ govern the occurrences of shocks for components 1, 2, ... , and m, respectively; $Z_{12}(t; \lambda_{12}),$
For example, for a component pair $(1,2)$, $(1,3)$, ..., and $(m-1,m)$, respectively, and $Z_{1,i}(t; \lambda_{1,i})$ govern the occurrences of shocks for the component pairs $(i,1)$; ..., $Z_{1,2,m}(t; \lambda_{1,2,m})$ govern the occurrences of shocks for the component pairs $(1,2,\ldots,m)$. Thus if $X_1, X_2, \ldots, X_m$ denote the lifetime of the first, second, and $m$th component, then similar to the bivariate Marshall–Olkin exponential model, the distribution of residual life is

$$
F(x_1,x_2,\ldots,x_m) = P(X_1 > x_1, \ldots, X_m > x_m) = \exp\left\{-\sum_{i=1}^{m} \lambda_{ij} \prod_{j \neq k \in S} \max(x_{ij},x_{jk}) - \sum_{i=1}^{m} \lambda_{ij} \max(x_{ij},x_{ik}) - \cdots - \lambda_{12\ldots m} \max(x_{1},x_{2},x_{3},\ldots,x_{m})\right\},
$$

(1)

Eq. (1) is called $m$-dimensional exponential distribution. To obtain a more compact notation for this distribution, let $S$ denote the set of vectors

$S = \{s = (s_1, \ldots, s_m); s_i = 0,1, s \neq (0,0,\ldots,0)\}.$

For example, for $s = (1,0,0,\ldots,0)$, $Z_1(t; \lambda_1)$ denotes $Z_{1,0,0,\ldots,0}(t; \lambda_{1,0,0,\ldots,0})$; and $s = (0,1,0,\ldots,0)$, $Z_2(t; \lambda_2)$ denotes $Z_{0,1,0,\ldots,0}(t; \lambda_{0,1,0,\ldots,0})$ and so on.

If $U_j$ denotes the time of the occurrence of shocks for the Poisson process $Z_j(t; \lambda_j)$, then $P(U_j > t) = \exp(-\lambda_j t)$ and $X_j = \min_{i \in S} (U_i)$, where $s_i = (s_1, s_{i-1}, 1, s_{i+1}, \ldots, s_m)$. Thus (1) can be written as

$$
F(x_1,x_2,\ldots,x_m) = P(X_1 > x_1, \ldots, X_m > x_m) = \exp\left\{-\sum_{i \in S} \lambda_{ij} \max(x_{i})\right\}.
$$

(2)

2.2. Assumptions

Let us first define $\varphi_0 < \varphi_1 < \cdots < \varphi_k$ as the ordered stress levels to be used in the test and $\varphi_0$ as the use stress. Consider the following k-step-stress ALT scheme with Type-I censoring: All $n$ test units are initially placed on stress level $\varphi_1$ and run until $t_1$ when the stress is changed to $\varphi_2$ for those units that have not failed. The test is continued until $t_2$ when the stress is changed to $\varphi_3$. Go on like this and continued until a pre-determined censoring time $t_k$. At each $\varphi_i$, the number of failures $n_i$ and failure time $t_i$ are observed. Note that the number of failures $n_i$ at stress $\varphi_i$ is random. The objective here is to choose the $t_1, \ldots, t_{k-1}$ according to some optimality criterion. The following assumptions A1–A5 are considered.

A1: A system functions if and only if all of its $m$ components are functioning. That is, the system is a series system. The components of the $m$-component system fail after receiving a fatal shock. This fatal shock model can be seen in Section 2.1.

A2: Failure time of a system T is the smallest failure time of its $m$ potential components, that is, $X = \min(X_i : i = 1, \ldots, m)$.

A3: For any level of stress, the time of the occurrence of shock $U_i$ follows an exponential distribution. In other words, at stress $\varphi_i$, the distribution of $U_i$ follows the exponential distribution with parameter $\lambda_i$, and $F_{U_i}(t) = P(U_i < t) = 1 - \exp(-\lambda_i t)$, $i = 0, 1, \ldots, k$.

A4: Stress variables (e.g., temperature, voltage, humidity, mechanical load, vibration, etc.) can be used for SSALT. Life times at low stress tend to be longer than those at high stress. The relationships between the mean (median, percentile) life time of products and stress are different for different products and stress variables according to engineering theory. In the following we list some life-stress relationships (see Nelson, 1990):

- Arhenius model: $\ln(\theta) = \gamma_0 + \gamma_1/v$, for example $v$ is the absolute temperature.
- Inverse power model: $\ln(\theta) = \gamma_0 + \gamma_1/(v)$, for example $v$ is the voltage.
- Exponential model: $\ln(\theta) = \gamma_0 + \gamma_1 v$, for example, $v$ is a weathering variable.

Thus the $\ln(\theta)$ is a linear function of the transformed stress $\varphi = 1/v, -\ln(v), v$ for the above three models. Life $\theta$ usually taken to be a specified percentile of the life distribution. The linear relationship is sometimes not adequately for use. Thus a polynomial relationship for the $\ln(\theta)$ as a function of (possible transformed) stress $\varphi$ is considered in this paper. That is, at stress level $\varphi_i$, the mean time of $U_i$, $\theta_{s_i}$, is a log-linear function with respect to stress, which is given by

$$
\ln(\theta_{s_i}) = \beta_{0s} + \beta_{1s} \varphi_i + \beta_{2s} \varphi_i^2 + \cdots + \beta_{l-1,s} \varphi_i^{l-1}, \quad l \geq 2, \ i = 0, \ldots, k, \ s \in S, \ \theta_{s_i} = 1/\lambda_{s_i},
$$

(3)

where the regression parameters $\beta_{0s}, \beta_{1s}, \ldots, \beta_{l-1,s}$ are unknown and needed to be estimated. When the values of $s$ are just taken as $(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,0,\ldots,0,1)$ and $k = l = 2$, this model was studied by Bai and Chun (1991).
A5: For each failure cause, a cumulative exposure model (Miller and Nelson, 1983) is assumed. That is, the cumulative distribution function of the failure cause \( s \) under \( k \)-step-stress test is

\[
G_s(t) = \begin{cases} 
F_{t_1}(t), & 0 \leq t < \tau_1, \\
F_{t_2}(\hat{\delta}_{11} + t - \tau_1), & \tau_1 \leq t < \tau_2, \\
\vdots \\
F_{t_k}(\hat{\delta}_{k-11} + t - \tau_{k-1}), & \tau_{k-1} \leq t < \tau_k,
\end{cases}
\]

where \( \hat{\delta}_{ij} \) is the solution of \( F_{t_i}(\hat{\delta}_{ij}) = G_s(\tau_{i-1}), F_{t_i}(t) = 1 - e^\left(-\lambda_{ij}t\right), i = 1, \ldots, k. \)

Because the cause of the product failure is due to random, we use the following \( m \)-dimensional random variable \( C = (c_1, \ldots, c_m) \) to denote the cause of product failure, where each \( c_i = 0 \) or \( 1 \) but \( (c_1, \ldots, c_m) \neq (0, \ldots, 0) \). For example, \( C = (1, 1, 0, \ldots, 0) \) means that the product failure is caused by the first and second components simultaneously. \( t_{ij} \) denotes the failure time of test unit \( j \) at stress level \( \phi_i \) and \( c_{ij} \) denotes the cause of the failure of test unit \( j \) at stress level \( \phi_i \). Thus, our failure data can be denoted as

\[
(t_{ij}, c_{ij}), \ i = 1, \ldots, k, j = 1, \ldots, n_i.
\]

3. Basic lemmas

**Lemma 1.** The probability density function (pdf) of the lifetime distribution at stress level \( \phi_i \) is

\[
F_i(t) = 1 - \exp\left(-\sum_{s \in S} \lambda_s t\right).
\]

**Proof.** Under the assumption of A2 and A3, we know

\[
F_i(t) = P_i(X < t) = P_i(\min(X_1, \ldots, X_m) < t) = P_i(\min(U_1) < t) = 1 - P_i(\min(U_i) > t) = 1 - \exp\left(-\sum_{s \in S} \lambda_s t\right). \quad \Box
\]

**Lemma 2** (Li, 2005). \( X = \min(X_1, \ldots, X_m) \) is independent of the following events:

\[
\{X_{i_0} = \min(X_1, \ldots, X_m)\}, \quad i_0 = 1, \ldots, m,
\]

\[
\{X_{i_0} = X_{i_2} = \min(X_1, \ldots, X_m)\}, \quad 1 \leq i_1 < i_2 \leq m,
\]

\[
\vdots
\]

\[
\{X_1 = X_2 = \cdots = X_m\}.
\]

**Proof.** See Theorem 6 of Li (2005). \( \Box \)

**Remark.** Lemma 2 shows that \( X \) and \( C \) are independent.

**Lemma 3.**

\[
P(C = s) = \frac{\lambda_s}{\lambda},
\]

where \( \lambda = \sum_{s \in S} \lambda_s \).

**Proof.** For \( s \in S \),

\[
P(C = s) = \int_0^\infty \int_0^\infty \int_0^\infty \text{d}U_s < u, y \text{d}U_s < u, y \text{d}U_s < u = \int_0^\infty \int_0^\infty \int_0^\infty e^{-u/\lambda} \lambda_s e^{-u/\lambda} \text{d}u = \frac{\lambda_s}{\lambda}. \quad \Box
\]

**Lemma 4.** The probability density function of \( (X, C) \) is

\[
f(t, c) = \prod_{s \in S} \lambda_s(t) \cdot \exp\left(-\sum_{s \in S} \lambda_s t\right),
\]

where \( \lambda_s(t) = I_{s = c} \) is an indicator function.
Based on Lemma 2, it is easy to obtain
\[ F(t) = \exp(-\sum_{s=3}^{t} \lambda_s t) \]
and
\[ f(t) = \sum_{s=3}^{t} \lambda_s \exp(-\sum_{s=3}^{t} \lambda_s t) = \lambda \exp(-\lambda t). \]

Based on Lemma 2, it is easy to obtain \( f(t, c) = f_T(t)P(C = c) \). Therefore
\[ f(t, c) = \prod_{s=3}^{t} \tilde{\xi}_s^{(c)} \cdot \exp\left(-\sum_{s=3}^{t} \lambda_s t\right). \]

From previous Assumptions A1–A5, and Lemmas 1–4, the cumulative distribution function of the lifetime of a test unit under k-step-stress test is
\[ G(t) = \begin{cases} 
F_1(t), & 0 \leq t < \tau_1, \\
F_2(\delta_1 + t - \tau_1), & \tau_1 \leq t < \tau_2, \\
\vdots \\
F_k(\delta_{k-1} + t - \tau_{k-1}), & \tau_{k-1} \leq t < \tau_k.
\end{cases} \]

where
\[ \delta_i = \frac{\sum_{j=1}^{k-i} \lambda_j (\tau_j - \tau_{j-1})}{\lambda_{i+1}}, \quad i = 1, 2, \ldots, k-1, \]
\[ \tau_0 = 0, \quad F_i(t) = 1 - e(-\lambda_i t), \quad \lambda_i = \sum_{s=3}^{t} \tilde{\lambda}_s, \quad i = 1, \ldots, k. \]

**Lemma 5.** The log-likelihood function of k-step-stress test is obtained as
\[ \ln L = \sum_{i=1}^{k} \sum_{s=3}^{t} \ln \lambda_i - \sum_{i=3}^{t} \lambda_i U_i, \]
where \( g_{ij} = \sum_{j=1}^{n} \tilde{\xi}_s (C_j), \quad U_i = \sum_{n=1}^{n} (t_i j - \tau_{i-1}) + (\tau_i - \tau_{i-1})(\sum_{j=1}^{k} n_j + n_k). \) Note that \( U_i \) is the total time at stress \( \varphi_i \).

**Proof.** By Lemma 4 and Eq. (5), we have the likelihood function
\[ L = \prod_{j=1}^{n} f_{U_j}(t_{ij}, C_j) \cdot \prod_{j=1}^{n} f_{C_j}(t_{ij} - \tau_1 + \delta_1, C_j) \cdot \prod_{j=1}^{n} f_{C_j}(t_{ij} - \tau_{k-1} + \delta_{k-1}, C_j) [1 - F_k(\tau_{k-1} + \delta_{k-1})]^{n_k} \]
\[ \prod_{j=1}^{n} \left( \prod_{s=3}^{x} \tilde{\xi}_s^{(c)} \cdot e^{-\tilde{\lambda}_s t_i} \right) \prod_{j=1}^{n} \left( \prod_{s=3}^{x} \tilde{\xi}_s^{(c)} \cdot e^{-\tilde{\lambda}_s (t_{ij} - \tau_{i-1} + \delta_{i-1})} \right) \]
\[ \prod_{j=1}^{n} \left( \prod_{s=3}^{x} \tilde{\xi}_s^{(c)} \cdot e^{-\tilde{\lambda}_s (t_{ij} - t_{j-1} + \delta_{j-1})} \right) \cdot [e^{-\tilde{\lambda}_s (t_{ij} - t_{j-1} + \delta_{j-1})}]^{n_k}. \]

Substituting for \( \delta_i \) in \( \ln L \), the log-likelihood function can be written as (6). \( \square \)

**Lemma 6.** For k-step-stress test under Type-I censoring, we have
\[ E(U_i) = n A_j / \tilde{\lambda}_i, \quad i = 1, 2, \ldots, k, \]
where \( A_i = p_i \sum_{j=1}^{n} (1 - p_j), \quad p_0 = 0, \quad p_i = 1 - \exp(-\lambda_i (\tau_i - \tau_{i-1})), \quad i = 1, 2, \ldots, k. \)

**Proof.** Let \( Y_j \) denote the failure time of test units \( j, j = 1, \ldots, n \).
\[ 1 \quad \tau_{i-1} \leq Y_j < \tau_i, \]
\[ 0 \quad \text{others.} \]

Then \( n_i = \sum_{j=1}^{n} I_{ij} \), and \( \sum_{j=1}^{n} t_{ij} = \sum_{j=1}^{n} I_{ij} Y_j. \) From (5), \( Y_j \sim \tilde{\xi}_s G(t) \), we obtain
\[ E(n_i) = n E(I_{ij}) = n [G(\tau_i) - G(\tau_{i-1})] = n [\exp(-\tilde{\lambda}_i (\tau_i - \tau_{i-1})) - \exp(-\tilde{\lambda}_i (\tau_i - \tau_{i-1}))] = n p_i \sum_{j=0}^{n-1} (1 - p_j) \]
Based on Assumption A4, (3) can be written as

\[ \ln L = \sum_{j=1}^{n} \ln Y_j - \sum_{j=1}^{n} \ln \left( \lambda_j \exp \left( \frac{-\lambda_j (\delta_{i,j} - t) - \tau_j}{\lambda_j} \right) \right) - \sum_{j=1}^{n} \ln \left( \lambda_j \exp \left( \frac{-\lambda_j (\delta_{i,j} - \tau_j)}{\lambda_j} \right) \right) \]

Proof. From Lemma 5, by solving the equations \( \partial \ln L / \partial \lambda_i = 0 \), the maximum likelihood estimates of \( \lambda_i \) can be obtained as \( \hat{\lambda}_i = g_{i0}/U_i \). Thus, we have \( \hat{\lambda}_i = \sum_{s=1}^{S} \hat{\lambda}_{is} = \sum_{s=1}^{S} g_{is}/U_i \).

Lemma 7. The maximum likelihood estimates of \( \lambda_{is} \) and \( \lambda_i \) are

\[ \hat{\lambda}_{is} = \frac{g_{is}}{U_i}, \quad \hat{\lambda}_i = \sum_{s=1}^{S} \hat{\lambda}_{is} = \sum_{s=1}^{S} \frac{g_{is}}{U_i} \]

where \( g_{is} = \sum_{j=1}^{n} \xi_j(C_{ij}) \).

Proof. From Lemma 5, by solving the equations \( \partial \ln L / \partial \lambda_{is} = 0 \), the maximum likelihood estimates of \( \lambda_{is} \) can be obtained as \( \hat{\lambda}_{is} = g_{is}/U_i \). Thus, we have \( \hat{\lambda}_i = \sum_{s=1}^{S} \hat{\lambda}_{is} = \sum_{s=1}^{S} g_{is}/U_i \).

Lemma 8. The maximum likelihood estimation of parameter \( \beta_i = (\beta_{i0}, \beta_{1i}, \ldots, \beta_{li}) \) is \( \hat{\beta}_i = A^+ Q_i \), where

\[ A = \begin{pmatrix} 1 & \phi_1 & \cdots & \phi_{l1}^{-1} \\ 1 & \phi_2 & \cdots & \phi_{l2}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \phi_k & \cdots & \phi_{lk}^{-1} \end{pmatrix}, \quad Q_i = \begin{pmatrix} U_1 - \ln g_{i1} \\ U_2 - \ln g_{i2} \\ \vdots \\ U_k - \ln g_{ik} \end{pmatrix}, \quad \hat{\beta}_i = \begin{pmatrix} -\ln \hat{\lambda}_{11} \\ -\ln \hat{\lambda}_{21} \\ \vdots \\ -\ln \hat{\lambda}_{i1} \end{pmatrix} \]

Proof. Based on Assumption A4, (3) can be written as \( R_i = A \beta_i \). From Lemma 7, we obtain \( \hat{R}_i = Q_i \). Thus the maximum likelihood estimation of \( \beta_i \) can be obtained by solving the equation \( A \hat{\beta}_i = \hat{R}_i = Q_i \). We will discuss the solution of equation \( A \hat{\beta}_i = Q_i \), for the different circumstances of \( i \) and \( k \) in the following.

1. \( i > k \): rank(\( A \)) = \( k \). Therefore, \( A \beta_i = Q_i \) has solution \( A^+(A^+)^{-1} Q_i \). By Lemma 9, \( A^+(A^+)^{-1} Q_i \) can be written as \( A^+ Q_i \).

2. \( i = k \): rank(\( A \)) = rank(\( A Q_i \)) = \( k \). Therefore, \( A \beta_i = Q_i \) has unique solution \( A^{-1} Q_i \). By Lemma 9, \( A^{-1} Q_i \) can be written as \( A^{-1} Q_i \).

3. \( i < k \): We have rank(\( A \)) = \( i \). If rank(\( A Q_i \)) = \( i \), then \( A \beta_i = Q_i \) has unique solution \( (A^T A)^{-1} A^T Q_i \). If rank(\( A Q_i \)) = \( i+1 \), then \( A \beta_i = Q_i \) has least squares solution \( (A^T A)^{-1} A^T Q_i \). That is \( \hat{\beta}_i = (A^T A)^{-1} A^T Q_i = A^+ Q_i \).

Lemma 9 (Rao and Mitra, 1971). Generalized inverse matrix has the following properties:

1. \( (A^+)^+ = (A^+)^+ \);
2. \( (AA^T)^+ = (A^+)^+ A^+ (A^T A)^+ = A^+ (A^T)^+ \);
3. \( \text{rk}(A_{n,m}) = n \) case, \( A^+ = (A A^T)^{-1} \);
4. \( \text{rk}(A_{n,m}) = m \) case, \( A^+ = (A^T A)^{-1} A^T \).

4. Optimal step-stress test under Type-I censoring

4.1. Fisher’s information matrix

Since the Fisher information matrix \( G(\beta) \) is obtained by taking the expectation of the negative of second partial derivatives of \( \ln L \) with respect to \( \beta = (\beta_{i0}, \beta_{1i}, \ldots, \beta_{li}), s \in S \), the Fisher information matrix \( G \) is obtained as

\[ G = \text{Diag}(G_s, s \in S) \]
where

\[ G_s = n \left( \sum_{i=1}^{k} E_{i} \phi_{i} \cdots \sum_{i=1}^{k} E_{i} \phi_{i}^{l-1} \right) \left( \sum_{i=1}^{k} E_{i} \phi_{i} \cdots \sum_{i=1}^{k} E_{i} \phi_{i}^{l-1} \right) \cdots \left( \sum_{i=1}^{k} E_{i} \phi_{i}^{l-1} \cdots \sum_{i=1}^{k} E_{i} \phi_{i}^{l-2} \right) \]  

(8)

\[ E_i = A_i \lambda_i / \lambda_i \] and \( A_i \) are defined by Lemma 6.

**Proof.**

\[ \frac{\partial^2 \ln L}{\partial \beta_{0s}^2} = -\sum_{i=1}^{k} U_i \exp(-\beta_{0s} + \beta_{1s} \phi_i + \beta_{2s} \phi_i^2 + \cdots + \beta_{l-1s} \phi_i^{l-1}) = -\sum_{i=1}^{k} U_i \cdot \lambda_i. \]

Similarly we have

\[ \frac{\partial^2 \ln L}{\partial \beta_{ps}^2} = -\sum_{i=1}^{k} U_i \cdot \lambda_i \phi_i, \quad \frac{\partial^2 \ln L}{\partial \beta_{0s} \partial \beta_{js}} = -\sum_{i=1}^{k} U_i \cdot \lambda_i \phi_i, \quad j = 1, \ldots, l-1, \]

\[ \frac{\partial^2 \ln L}{\partial \beta_{ps} \partial \beta_{js}} = -\sum_{i=1}^{k} U_i \cdot \lambda_i \phi_i^{h+j}, \quad h \neq j, \quad \frac{\partial^2 \ln L}{\partial \beta_{ps} \partial \beta_{js}} = 0, \quad s' \neq s''. \]

From Lemma 6, we obtain

\[ E \frac{\partial^2 \ln L}{\partial \beta_{ps}^2} = -n \sum_{i=1}^{k} \lambda_i A_i / \lambda_i, \quad E \frac{\partial^2 \ln L}{\partial \beta_{0s} \partial \beta_{js}} = -n \sum_{i=1}^{k} \lambda_i \phi_i^{h+j} A_i / \lambda_i, \quad E \frac{\partial^2 \ln L}{\partial \beta_{ps} \partial \beta_{js}} = 0. \]

Under the mild regularity conditions, the maximum likelihood properties can be used to make inference about \( \beta \). Hence, the asymptotic normality of \( \beta \) is obtained. That is, for any \( s \in S \)

\[ (\beta_{0s}, \beta_{1s}, \ldots, \beta_{l-1s}) \sim \mathcal{N}(0, \Sigma_s^{-1}), \]

and the approximate confidence intervals for \( \beta_s \) or the asymptotic joint confidence region for \( \beta \) can easily obtained. For example, the two sided 100(1-\( \alpha \)% approximate confidence interval for the parameter \( \beta_{0s} \) is given by

\[ \beta_{0s} \pm z_{1-\alpha/2} \sqrt{\text{var}(\hat{\beta}_{0s})}, \]

where \( z_{1-\alpha/2} \) is the \( (1-\alpha/2) \)th quantile of the standard normal distribution and \( \sqrt{\text{var}(\hat{\beta}_{0s})} \) is obtained by taking square root of the first diagonal element of \( \Sigma_s^{-1} \).

### 4.2. D-optimality

The main purpose of this paper is to study the choice of \( \tau_1, \tau_2, \ldots, \tau_{k-1} \), in a \( k \)-step-stress ALT with Type-I censoring. D-optimality criterion, often used in planning ALT, is based on the determinant of the Fisher information matrix, which is the same as the reciprocal of the determinant of the asymptotic variance covariance matrix. Note that the overall volume of the asymptotic joint confidence region of \( (\beta_{0s}, \beta_{1s}, \ldots, \beta_{l-1s}) \) is proportional to \( |G_s|^{-1/2} \) at a fixed confidence level. Consequently, a larger value of \( |G_s| \) would correspond to a smaller asymptotic joint confidence ellipsoid of \( (\beta_{0s}, \beta_{1s}, \ldots, \beta_{l-1s}) \) and thus a higher joint precision of the estimators of \( (\beta_{0s}, \beta_{1s}, \ldots, \beta_{l-1s}) \). Motivated by this, our objective is to select the optimal change times \( \tau_1, \tau_2, \ldots, \tau_{k-1} \), to maximum \( |G| = \prod_{s \in S} |G_s| \).

**Theorem 1.** For \( k \)-step-stress test under Type-I censoring, we have

\[ |G| = \begin{cases} 0, & k < l, \\ \prod_{s \in S} \sum_{1 \leq i < j < \cdots < k \leq k} \left[ E_{i1} E_{i2} \cdots E_{i(k-1)} \prod_{1 \leq i < j \leq k} (\phi_{ij} - \phi_{ij}^*)^2 \right], \end{cases} \]

\[ k \geq l. \]
Proof.

For $k < l$, because all of $E_{i1}, E_{i2}, \ldots, E_{il}$ are taken from $E_{i1}, E_{i2}, \ldots, E_{i8}$, there must exist two equivalent values. Therefore $|G_s| = 0$.

For $k \geq l$,

$$|G_s| = n! \sum_{i_1 = 1}^{k} \sum_{i_2 = 1}^{k} \ldots \sum_{i_l = 1}^{k} 
\begin{bmatrix}
\phi_{i_1} & \phi_{i_2} & \ldots & \phi_{i_l} \\
\phi_{i_1}^2 & \phi_{i_2}^2 & \ldots & \phi_{i_l}^2 \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{i_1}^{l-1} & \phi_{i_2}^{l-1} & \ldots & \phi_{i_l}^{l-1} 
\end{bmatrix}
\begin{bmatrix}
1 & 1 & \ldots & 1 
\end{bmatrix}
\begin{bmatrix}
E_{i_1} & E_{i_2} & \ldots & E_{i_l} \\
E_{i_1} & E_{i_2} & \ldots & E_{i_l} \\
\vdots & \vdots & \ddots & \vdots \\
E_{i_1} & E_{i_2} & \ldots & E_{i_l} 
\end{bmatrix}
$$

$$= n! \sum_{i_1 = 1}^{k} \sum_{i_2 = 1}^{k} \ldots \sum_{i_l = 1}^{k} 
\begin{bmatrix}
E_{i_1} & E_{i_2} & \ldots & E_{i_l} \\
E_{i_1} & E_{i_2} & \ldots & E_{i_l} \\
\vdots & \vdots & \ddots & \vdots \\
E_{i_1} & E_{i_2} & \ldots & E_{i_l} 
\end{bmatrix}
= n! \prod_{1 \leq t < r \leq l} (\varphi_{i_t} - \varphi_{i_r})^2.
\square
$$

Remark. Under the condition of $k < l$, from Theorem 1, we know that no matter how the arrangement of $\tau_1, \tau_2, \ldots, \tau_{k-1}$ is, the value of $|G|$ is always zero. By the D-optimality criterion, the smaller the value of $|G|$, the worse the estimator of $\beta$. So we should try to avoid such situation occurs ($k < l$). In other words, it is strongly recommended that ALT test under the condition of $k \geq l$.

Theorem 2. In the case of $k = l$, in order to maximum the value of $|G| = \prod_{s \in S} |G_s|$, the D-optimal stress change times $\tau_1, \tau_2, \ldots, \tau_{k-1}$ should satisfy the following equations:

$$\lambda_i^k [k+1-i-\frac{1}{p_i}] = \lambda_k^k [1-\frac{1}{p_k}], \quad i = 1, \ldots, k-1,$n

where $\sum_{i=1}^{k} \tau_i = T_1, T_i = \tau_i - \tau_{i-1}, \tau_0 = 0, \tau_1 = T_1 + \cdots + T_i, i = 1, \ldots, k$.

Proof. The maximum of $|G|$ is equivalent to the maximum of $\ln |G|$, so we consider the value of $\ln |G|$.

$$\ln |G| = \sum_{s \in S} \ln \left[ \sum_{1 \leq j < r \leq k} (\varphi_{i_j} - \varphi_{i_r})^2 \right]
$$

In the case of $k = l$, the above equation can be written as

$$\sum_{s \in S} \ln \left[ E_{i_1} E_{i_2} \cdots E_{i_l} \prod_{1 \leq j < r \leq l} (\varphi_{i_j} - \varphi_{i_r})^2 \right]
= \sum_{s \in S} \ln E_{i_1} + 2 \sum_{1 \leq j < r \leq l} \ln |\varphi_{i_j} - \varphi_{i_r}|
= \sum_{s \in S} \sum_{i = 1}^{k} \ln A_i + \sum_{s \in S} \sum_{i = 1}^{k} \ln A_{i,i} + \sum_{s \in S} \sum_{i = 1}^{k} \ln A_{i,i} + 2 \sum_{1 \leq j < r \leq k} \ln |\varphi_{i_j} - \varphi_{i_r}|.
$$

Form (10), as long as the value of $\sum_{i=1}^{k} \ln A_i$ reaches the maximum, the value of $|G|$ will reach the maximum. Let $g(T_1, \ldots, T_k) = -\sum_{i=1}^{k} \ln A_i$. If the value of $\sum_{i=1}^{k} \ln A_i$ reaches the maximum, then the value of $g(T_1, \ldots, T_k)$ reaches the minimum. In the minimum, we will seek $T_1, \ldots, T_k$ to make $g(T_1, \ldots, T_k)$ reaches minimum.

Because $\frac{\partial g}{\partial T_j} = \lambda_j^k \exp(-\lambda_j T_j) / p_j^2$, we get $\frac{\partial g}{\partial T_j} = \lambda_j^k \exp(-\lambda_j T_j) / p_j^2 + \lambda_j^k \exp(-\lambda_j T_j) / p_j^2 = \lambda_j^k + h_i, i = 1, \ldots, k-1$. The Hessen matrix of $g(T_1, \ldots, T_k)$ is obtained as

$$H(g) = \begin{bmatrix} h_i + h_k & h_k & \ldots & h_k & h_k - 1 & h_k \\ h_k & h_k - 1 & \ldots & h_k & h_k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_k & h_k & \ldots & h_k & h_k \\ \end{bmatrix} = \begin{bmatrix} h_i & 0 & 0 & \ldots & 0 \\ 0 & h_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & h_k & 0 \\ \end{bmatrix} \pm H_1 + H_2,
$$

where $A_i = p_i \prod_{j=0}^{i-1} (1-p_j), p_0 = 0, p_i = 1 - \exp(-\lambda_i (\tau_i - \tau_{i-1})) = 1 - \exp(-\lambda_i T_i), \tau_0 = 0, i = 1, 2, \ldots, k$. Because $\lambda_j^k \exp(-\lambda_j T_j) / p_j^2 > 0$, $H_1$ is a positive definite matrix and $H_2$ is a positive semidefinite matrix. Therefore, $g(T_1, \ldots, T_k)$ is a strictly convex
function. Hence, the D-optimal stress time length $T_1, \ldots, T_{k-1}$ are the solution of
\[
\frac{\partial \ell}{\partial T_i} = -\frac{\lambda_i}{p_i} \exp(-\lambda_i T_i) + \frac{\lambda_k}{p_k} \exp(-\lambda_k T_k) + (k-i) \lambda_i = 0, \quad i = 1, \ldots, k-1.
\]
By simplifying the above equations, and replaced $T_i$ by $\tau_i - \tau_{i-1}$, the D-optimal stress change points $\tau_1, \tau_2, \ldots, \tau_{k-1}$ satisfy the following equations:
\[
\lambda_i \left[ k+1-i - \frac{1}{p_i} \right] = \lambda_k \left[ 1 - \frac{1}{p_k} \right], \quad i = 1, \ldots, k-1. \quad \square
\]

4.3. Variance-optimality

The mean of the failure time distribution is an important characteristic and indispensable in reliability analysis. In step-stress setting, we need to estimate the mean lifetime at the use stress with maximum precision. We can use the asymptotic variance of the logarithm of mean lifetime at use stress as the criterion for selecting the optimal time points. For this purpose, we consider an objective function from (8) as
\[
g_v(\tau_1, \tau_2, \ldots, \tau_{k-1}) = \sum_{s=3}^{n} \text{AsVar}(\ln \theta_{n0}) = \sum_{s=3}^{n} \left[ \text{AsVar}(\beta_{01} + \beta_{11} \phi_{01} + \cdots + \beta_{1s-1,1} \phi_{1s-1}^{(s-1)}) \right] = \sum_{s=3}^{n} \left[ nX_g^2 G_s^{-1} X_0 \right] = \sum_{s=3}^{n} H_s,
\]
where $H_s = nX_g^2 G_s^{-1} X_0$ and $X_0^2 = (1, \phi_{01}, \ldots, \phi_{1s-1})$. AsVar stands for the asymptotic variance and $\phi_{01}$ is the use stress. The Variance-optimality criterion is to find $\tau_1, \tau_2, \ldots, \tau_{k-1}$ which minimize $g_v(\tau_1, \tau_2, \ldots, \tau_{k-1})$.

**Theorem 3.** For $k \leq l$,
\[
g_v(\tau_1, \tau_2, \ldots, \tau_{k-1}) = \frac{\phi_2^2}{E_1} + \frac{\phi_2^2}{E_2} + \cdots + \frac{\phi_2^2}{E_k} = \frac{\phi_2^2}{A_1} + \frac{\phi_2^2}{A_2} + \cdots + \frac{\phi_2^2}{A_k},
\]
and for $k > l$,
\[
g_v(\tau_1, \tau_2, \ldots, \tau_{k-1}) = \frac{\eta_1^2}{E_1} + \frac{\eta_1^2}{E_2} + \cdots + \frac{\eta_1^2}{E_k} = \frac{\eta_1^2}{A_1} + \frac{\eta_1^2}{A_2} + \cdots + \frac{\eta_1^2}{A_k},
\]
where
\[
\phi_2^2 = \lambda_0 \theta_1 \phi_2^2, \quad \eta_1^2 = \lambda_0 \theta_1 \eta_1^2, \quad \frac{1}{E_i} = \sum_{s=3}^{n} \frac{1}{E_{is}}, \quad \theta_k = \sum_{s=3}^{n} \frac{A_i}{A_k} E_{is} = \frac{A_i}{A_k} E_{is},
\]
and
\[
A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \phi_1 & \phi_2 & \cdots & \phi_k \\ \vdots & \vdots & \ddots & \vdots \\ \phi_k^{-1} & \phi_k^{-1} & \cdots & \phi_k^{-1} \end{pmatrix}, \quad E_i = \begin{pmatrix} E_{1s} & 0 & \cdots & 0 \\ 0 & E_{2s} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{ks} \end{pmatrix}, \quad D = E^{1/2} = \begin{pmatrix} \sqrt{E_{1s}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{E_{ks}} \end{pmatrix}
\]
and $A_i$ is given in Lemma 6.

**Proof.** $g_v(\tau_1, \tau_2, \ldots, \tau_{k-1}) = \sum_{s=3}^{n} [nX_g^2 G_s^{-1} X_0] = \sum_{s=3}^{n} H_s$. $G_s = nA^2 E_s A^T$ is obtained from (8), therefore, $H_s = nX_g^2 G_s^{-1} X_0 = nX_g^2 (nA E_s A^T)^{-1} X_0 = nX_g^2 (nA E_s A^T)^{-1} X_0$. In order to calculate $H_s$, we need to compute $(A E_s A^T)^+$. In the following two cases, we discuss how to compute $(A E_s A^T)^+$.

1. In the case of $k \leq l$, $\text{rank}(A) = \text{rank}(E_s) = \text{rank}(E_s) = \text{rank}(G_s) = k$. From Lemma 9, we have
\[
(A E_s A^T)^+ = A (A^T A)^{-1} (A E_s A^T)^T = A (A^T A)^{-1} \left( E_s A^T \right)^T = A (A^T A)^{-1} \left( E_s A^T \right)^T = A (A^T A)^{-1} E_s A^T = A (A^T A)^{-1} E_s A^T = (A^T A)^{-1} E_s A^T.
\]

$H_s = nX_g^2 G_s^{-1} X_0 = nX_g^2 (nA E_s A^T)^{-1} X_0 = X_0^2 A^T E_s A^T X_0 = \left( a_1^T + \phi_2^0 a_2 + \cdots + \phi_2^{k-1} a_k \right)^T D^{-1} \left( a_1^T + \phi_2^0 a_2 + \cdots + \phi_2^{k-1} a_k \right) = \left( a_1^T + \phi_2^0 a_2 + \cdots + \phi_2^{k-1} a_k \right)^T D^{-1} \left( a_1^T + \phi_2^0 a_2 + \cdots + \phi_2^{k-1} a_k \right) = \frac{\phi_2^2}{E_1} + \frac{\phi_2^2}{E_2} + \cdots + \frac{\phi_2^2}{E_k},
\]
where $d_{is}$ is the $i$th column of $D^{-1}$. 

Hence

\[ g_*(\tau_1, \tau_2, \ldots, \tau_{k-1}) = \sum_{s \leq s} H_s = \frac{\sigma^2_s}{E_1} + \frac{\sigma^2_s}{E_2} + \cdots + \frac{\sigma^2_s}{E_k} = \frac{\sigma^2_s}{A_1} + \frac{\sigma^2_s}{A_2} + \cdots + \frac{\sigma^2_s}{A_k}. \]

(2) In the case of \( k > l \), it is easy to know that \( \text{rank}(A) = \text{rank}(AE_s) = \text{rank}(G_s) = l \) and \( \text{rank}(E_s) = k \). Denote \( C = AE_s^{-1/2} \), that is \( C = AD \). Then we obtain \( (AE_sA^+) = (C^+)^T = (C^+)^T + (D^+) = (A^+) - (D^+) - (D^+) = (A)^T - A^T \).

\[ H_s = nX_0 G_s^{-1} X_0 = nX_0 (nAB A^+)^{-1} X_0 = X_0 (A)^T B^{-1} A^T X_0 \text{ similar to (1)} \ldots \ldots = \frac{\eta^2_s}{E_1} + \cdots + \frac{\eta^2_{k-1}}{E_{k-1}} + \frac{\eta^2_k}{E_k}. \]

\[ \square \]

**Theorem 4.** \( g_*(\tau_1, \tau_2, \ldots, \tau_{k-1}) = n \sum_{i=1}^{k} \text{AsVar}(\ln \theta_i) \) is a convex function under Type-I censoring in a step-stress test.

**Proof.** Let \( B_0 = 1, B_i = \prod_{j=1}^{i-1}(1-p_j) \), \( K_i = \exp(-(-\tau_i - \tau_{i-1})/\Theta_i), i = 1, \ldots, k \). Then \( B_i = K_i B_{i-1}, A_i = p_i \prod_{j=1}^{i-1}(1-p_j) = B_{i-1} - B_i, i = 1, \ldots, k \).

According to **Theorem 3**, we know that \( g_*(\tau_1, \tau_2, \ldots, \tau_{k-1}) \) is a positive linear combination of \( 1/A_i \). In order to prove \( g_*(\tau_1, \tau_2, \ldots, \tau_{k-1}) \) is a convex function, we only prove \( 1/A_i \) is a convex function. Because a positive linear combination of convex function is convex.

\( H(1/A_i), H(A_i) \) and \( H(B_i) \) denote Hession matrix of \( 1/A_i A_i \) and \( B_i, VB_i \) and \( V A_i \) denote gradient vector of \( B_i \) and \( A_i \); we have \( H(B_i) = \epsilon_i e_i B_i, \nabla V = -\epsilon_i B_i, \) where \( \epsilon_i \) is a \( k \times 1 \) dimension of vector, \( \epsilon^T = (0, \ldots, 0), \epsilon^T = (1/\Theta_1, 0, \ldots, 0) \), \( \epsilon^T = (1/\Theta_1 - 1/\Theta_2, 1/\Theta_2, 0, \ldots, 0) \), \( \epsilon^T = (1/\Theta_1 - 1/\Theta_2, 1/\Theta_2 - 1/\Theta_3, 0, \ldots, 0) \), \( \epsilon^T = (1/\Theta_1 - 1/\Theta_2, 1/\Theta_2 - 1/\Theta_3, 1/\Theta_3 - 1/\Theta_4, 0, \ldots, 0) \), \( \epsilon^T = (1/\Theta_1 - 1/\Theta_2, 1/\Theta_2 - 1/\Theta_3, 1/\Theta_3 - 1/\Theta_4, 1/\Theta_4 - 1/\Theta_5, 0, \ldots, 0) \). We obtain

\[ A_i^2 H^{-1}_{A_i} = -|H(A_i)| \cdot A_i + 2(\nabla B_i - \nabla V_i)(\nabla V_i - \nabla B_i) = -(H(B_i) - H(A_i)) \cdot A_i + 2(\nabla B_i - \nabla V_i)(\nabla V_i - \nabla B_i) = -(K_i e_i - e_i - e_i B_i) \cdot A_i + 2(e_i - e_i - e_i - e_i B_i) \]

\[ = (K_i e_i - e_i - e_i B_i)(K_i e_i - e_i B_i) + K_i e_i - e_i - e_i B_i. \]

Let \( X_i = K_i e_i - e_i, Y_i = e_i - e_i, \) then

\[ A_i^2 H^{-1}_{A_i} = X_i Y_i + Y_i X_i. \]

Therefore, \( A_i^2 H(1/A_i) B_i^{-1} \) is a semi-definite matrix. In other words, \( H(1/A_i) \) is a semi-definite matrix. Moreover, for \( i = k - 1, k \), \( X_i Y_i \) are definite matrices. Therefore, \( H(1/A_{k-1}) \) and \( H(1/A_k) \) are definite matrices. Thus \( g_*(\tau_1, \tau_2, \ldots, \tau_{k-1}) \) is a convex function. \( \square \)

**Remark.** Since \( g_*(\tau_1, \tau_2, \ldots, \tau_{k-1}) = n \sum_{i} \text{AsVar}(\ln \theta_i) \) is a convex function, so the points of minima exist.

**Theorem 5.** In the case of a step-stress test under Type-I censoring,

1. For \( k \leq l \), the V-optimal stress change points \( \tau^*_1, \tau^*_2, \ldots, \tau^*_k \) are the solution of

\[
\lambda_1 \left[ 1 + \frac{\sigma^2_1}{\lambda_1 \theta_1 E_1} \right] - \sum_{j=1}^{k-1} \frac{\sigma^2_j}{\lambda_j \theta_j E_j} A_j = \lambda_2 \left[ 1 + \frac{\sigma^2_2}{\lambda_2 \theta_2 E_2} \right] - \sum_{j=1}^{k-1} \frac{\sigma^2_j}{\lambda_j \theta_j E_j} A_j
\]

\[ \cdots = \lambda_k \left[ 1 + \frac{\sigma^2_k}{\lambda_k \theta_k E_k} \right] - \sum_{j=1}^{k-1} \frac{\sigma^2_j}{\lambda_j \theta_j E_j} A_j \]

(13)

2. For \( k > l \), the V-optimal stress change points \( \tau^*_1, \tau^*_2, \ldots, \tau^*_k \) are the solution of a new formula which replaces \( \sigma^2_i \) by \( \eta^2_i \) in (13).

**Proof.** (1) For \( k \leq l \), \( c = ([\tau_1, \tau_2, \ldots, \tau_k] : \tau_1 < \tau_2 < \cdots < \tau_k, 0 < \tau_i < +\infty, i = 1, 2, \ldots, k \) is the definition domain of \( g_*(\tau_1, \tau_2, \ldots, \tau_{k-1}) \). For any \( 0 < \lambda < 1, u = (u_1, u_2, \ldots, u_k) \in c \), and \( V = (v_1, v_2, \ldots, v_k) \in c \), we have \( 0 < u_i < +\infty, 0 < u_i < +\infty, i = 1, 2, \ldots, k \), \( u_i < u_2 < \cdots < u_k, v_i < v_2 < \cdots < v_k \). Hence, \( 0 < u_i + v_i < +\infty, v_i = 1, 2, \ldots, k, u_i + v_i < u_2 + v_2 < \cdots < u_k + v_k \). That is \( c \) is a convex domain. Moreover, it is easy to know that \( \partial g_*(\tau_1, \ldots, \tau_k) / \partial \tau_i = -\sum_{j=1}^{k} (\sigma^2_j/A_j) \partial A_j / \partial \tau_i \) is a continuous function with respect to \( c \). Therefore, \( g_*(\tau_1, \ldots, \tau_{k-1}) \) is a differentiable function with respect to \( c \). In summary, from the above arguments and **Theorem 4**, the optimal stress change points are the solution of \( \partial g_*(\tau_1, \ldots, \tau_k) / \partial \tau_i = 0, i = 1, 2, \ldots, k \). After solving these equations, we obtained \( \tau^*_i \) which satisfy the formula (13). \( \square \)
From Theorems 2 and 5, we know that the optimal change times $\tau_1, \ldots, \tau_k$ are no close form. We can use Newton–Raphson method to obtain the optimal change times. In the following section, we shall use an example to illustrate the proposed procedure.

5. Example

In this section, we illustrate the proposed procedure with an example based on the data set from Nelson (1990). The data in Table 1 are times of failure of a Class-H insulation in motors tested at high temperatures of 190, 220, 240, 260 °C. A test purpose was to estimate the median life of such insulation at its design temperature of 180 °C. Nelson (1990) discussed the main causes of failure at the design temperature and consider phase ($X_1$) and ground ($X_2$) as the main causes. In use, the first failure from any cause determines the motor life. Motors 11, 13 and 40 have the same failure times for both causes ($X_1$) and ($X_2$). That is $\Pr(X_1 = X_2) \neq 0$, the Marshall–Olkin model is appropriate for analyzing this data set. Lu (1992) gave the Bayesian analysis for this data with Marshall–Olkin model. In order to apply our method for this data, let $s_1 = (1, 0)$ and $s_2 = (0, 1)$ denote the failures of motor caused by $X_1$, and $X_2$ respectively, and $s_3 = (1, 1)$ denotes the failure of motor caused simultaneously by $X_1$ and $X_2$. Then the data $(X_1, X_2)$ transforms into the failure data $(t, c)$, where $t$ is the failure time, $c$ is the cause of failure. In addition, based on the failure-causing chemical reaction mechanism that causes failure,

<table>
<thead>
<tr>
<th>Phase ($X_1$)</th>
<th>Ground ($X_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>190 °C motor</td>
<td></td>
</tr>
<tr>
<td>1 10 511</td>
<td>10 511+</td>
</tr>
<tr>
<td>2 11 855</td>
<td>11 855+</td>
</tr>
<tr>
<td>3 11 855</td>
<td>11 855+</td>
</tr>
<tr>
<td>4 11 855</td>
<td>11 855+</td>
</tr>
<tr>
<td>5 12 191+</td>
<td>12 191+</td>
</tr>
<tr>
<td>6 12 191+</td>
<td>12 191+</td>
</tr>
<tr>
<td>7 12 191+</td>
<td>12 191+</td>
</tr>
<tr>
<td>8 12 191+</td>
<td>12 191+</td>
</tr>
<tr>
<td>9 12 191+</td>
<td>12 191+</td>
</tr>
<tr>
<td>10 12 191+</td>
<td>12 191+</td>
</tr>
<tr>
<td>220 °C motor</td>
<td></td>
</tr>
<tr>
<td>11 2436</td>
<td>2436</td>
</tr>
<tr>
<td>12 2436</td>
<td>2490</td>
</tr>
<tr>
<td>13 2436</td>
<td>2436</td>
</tr>
<tr>
<td>14 2772+</td>
<td>2772</td>
</tr>
<tr>
<td>15 2436+</td>
<td>2436</td>
</tr>
<tr>
<td>16 4116+</td>
<td>4116+</td>
</tr>
<tr>
<td>17 4116+</td>
<td>4116+</td>
</tr>
<tr>
<td>18 4116+</td>
<td>4116+</td>
</tr>
<tr>
<td>19 3108+</td>
<td>3108+</td>
</tr>
<tr>
<td>20 4116+</td>
<td>4116+</td>
</tr>
<tr>
<td>240 °C motor</td>
<td></td>
</tr>
<tr>
<td>21 1175+</td>
<td>1175</td>
</tr>
<tr>
<td>22 1881+</td>
<td>1175</td>
</tr>
<tr>
<td>23 1881+</td>
<td>1881+</td>
</tr>
<tr>
<td>24 1761</td>
<td>1761+</td>
</tr>
<tr>
<td>25 1881+</td>
<td>1881+</td>
</tr>
<tr>
<td>26 1881+</td>
<td>1881+</td>
</tr>
<tr>
<td>27 1881+</td>
<td>1881+</td>
</tr>
<tr>
<td>28 1881+</td>
<td>1881+</td>
</tr>
<tr>
<td>29 1881+</td>
<td>1881+</td>
</tr>
<tr>
<td>30 1953+</td>
<td>1953+</td>
</tr>
<tr>
<td>260 °C motor</td>
<td></td>
</tr>
<tr>
<td>31 1632+</td>
<td>600</td>
</tr>
<tr>
<td>32 1632+</td>
<td>744</td>
</tr>
<tr>
<td>33 1632+</td>
<td>744</td>
</tr>
<tr>
<td>34 1632+</td>
<td>744</td>
</tr>
<tr>
<td>35 1632+</td>
<td>912</td>
</tr>
<tr>
<td>36 1128+</td>
<td>1128</td>
</tr>
<tr>
<td>37 1512+</td>
<td>1320</td>
</tr>
<tr>
<td>38 1632+</td>
<td>1632+</td>
</tr>
<tr>
<td>39 1608+</td>
<td>1608</td>
</tr>
<tr>
<td>40 1896</td>
<td>1896</td>
</tr>
</tbody>
</table>
the Arrhenius relationship is expected to be adequate to describe the temperature acceleration relationship up to 260 °C. Thus, the acceleration model can be expressed as
\[
\begin{align*}
\ln \frac{g_0}{g_i} &= (\gamma_0 + \gamma_1)S_i, \\
S_i &= -1000/(\text{temp}^iC + 273), \\
i &= 0, 1, 2, 3, 4.
\end{align*}
\]

In this example, temp$_0$ = 180 °C, temp$_1$ = 190 °C, temp$_2$ = 220 °C, temp$_3$ = 240 °C, temp$_4$ = 260 °C. Let \( \phi_i = 1 + 4 \cdot (S_i - S_0)/(S_4 - S_0) \) then \( \phi_0 = 1, \phi_1 = 1.576, \phi_2 = 3.162, \phi_3 = 4.117 \) and \( \phi_4 = 5 \), and the above model becomes
\[
\ln \frac{g_0}{g_i} = \beta_0 + \beta_1 \phi_i.
\]

Use Lemmas 7 and 8, we obtain the MLE of the model parameters \( \hat{\theta}_1 = 177.4 \) weeks, \( \hat{\theta}_2 = 31.8 \) weeks, \( \hat{\theta}_3 = 25.8 \) weeks, and \( \hat{\theta}_4 = 7.5 \) weeks, 1 week = 7 \times 24 hours, and \( \hat{\beta}_0 = 6.491, \hat{\beta}_1 = -0.871 \). Moreover, we can obtain the MLE of \( \theta_{i1}, \theta_{i2}, \theta_{i3} \), \( i = 1, 2, 3, 4 \). For example, \( \hat{\theta}_{21} = \hat{\theta}_{23} = \hat{\theta}_{25} = 95.5 \) weeks.

5.1. Optimal SSALT plan

We first consider the case of \( k = 2, m = 2, l = 2 \) to illustrate the proposed method in Section 4. We construct an optimal SSALT plan with two higher stress levels, \( \phi_1 = 1.576(190 \text{°C}), \phi_2 = 3.162(220 \text{°C}) \). In practical applications, the true parameters \( \theta_1, \theta_2 \) or \( \theta_{i1}, \theta_{i2}, \theta_{i3} \), \( i = 1, 2, 3 \) will be unknown. The above \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) will consider a pre-estimate for parameters. For illustrative purpose, we assume that \( \hat{\theta}_{i1} = \hat{\theta}_{i2} = \hat{\theta}_{i3}, i = 1, 2, 3, 4 \). We consider the optimal plans under D-optimality and V-optimality criteria respectively.

1. D-optimal plan: For fixed \( m \) and \( n \), from formula (10), it is easy to know that the value of \( \Delta \) is only dependent on the value of \( \sum_{i=1}^l \ln A_i \). Therefore, to maximize \( \Delta \) is equivalent to minimize \( g(T_1, T_2) \), where \( g(T_1, T_2) = -\sum_{i=1}^l \ln A_i \). When preassigned censoring time \( T \) is given, \( g(T_1, T_2) = g(T_1, T - T_1) \) is a function of \( T_1 \). Let \( g(T) = g(T_1, T - T_1) \). The relation of \( g(T_1) \) and \( T_1 \) for different preassigned censoring times \( T \) is determined by Theorems 1 and 2 and that is showed in Fig. 1. For example, for \( T = 100 \), Fig. 1 shows that the optimal time length at stress \( \phi_1, T_1 \) is 54.7. From Fig. 1, we know that \( T_1 \) increases and the minimum of \( g(T_1) \) decreases with \( T \) increases.

2. V-optimal plan: By using the V-criterion in Section 4.3, the optimal plans were determined and are showed in Fig. 2. For example, for \( T = 100 \), Fig. 2 shows that the optimal time length at stress \( \phi_1, T_1 \) is 86.9. From Fig. 2, we know that \( T_1 \) increases and the minimum of \( g_T \) decreases with \( T \) increases, where \( g_T = n \sum_{i=1}^l \ln \text{Var(} \ln \theta_{i0} \text{)} \).

Remark. Under different optimality criteria, Figs. 1 and 2 demonstrate that the optimal stress change points are different for the same \( T \). From a practitioner’s point of view, the choice will certainly be guided by the objective of the experiment. In cases where the planner is more interested in estimating the MTTF \( \theta_0 \) at the use stress with high precision, certainly V-optimality will be the criterion of choice. On the other hand, if one is more interested in the estimation of the mean function (3) with high precision, a more reasonable criterion should be D-optimality.

![Fig. 1. The relation of g(T1) and T1 for different preassigned censoring times T.](image-url)
5.2. Sensitivity analysis

In practice, the pre-estimated parameter $c_y^1 = 177.4$ weeks, $c_y^2 = 31.8$ weeks would depart from the true parameters $y_1$ and $y_2$. Hence, it is important to investigate the effects of these unknown parameters on the optimal test plan. Without loss of generality, we assume that $e_1$ and $e_2$ denote the predicted errors for $y_1$ and $y_2$ respectively. Tables 2 and 3 present the optimal plans under various combinations of $(1 + e_1)c_1$, $(1 + e_2)c_2$. From these results, it shows that the test plan is quite robust for a moderate departure from the assumed values of these parameters. In addition, the relative bias on the $g$ or $g_v$ are presented in Tables 2 and 3 as well.

5.3. Numerical analysis

In order to further investigate the existence of the optimal stress change points and to evaluate them as a function of varying parameters (MTTF, the number of stress levels, and the degree of censoring), we conducted a small numerical study for D-criterion and V-criterion respectively.

1. D-optimal numerical result: From Eq. (3), we know that $\theta_i / \theta_{i-1}$ is determined by giving stresses $\phi_1, \ldots, \phi_k$. For the purpose of illustration, we consider $\theta_1, \theta_2$ and $T$, the values of $T_1, \ldots, T_k$ and $g(T_1, \ldots, T_k)$ are determined from Theorem 2 and Formula (9) in Section 4.2. The results are listed in Tables 4 and 5.

We obtain from Tables 4 and 5 the following conclusions:

- The time length at stress $\phi_i$, $T_i$ increases as $T$ increases as expected. However, for fixed $\theta_1$, the last column of Tables 4 and 5 demonstrates that $T_i$ becomes a constant and is almost the same for different $\rho$ as the $T$ increase to very large. This is quite interesting. For example, when $\theta_1 = 100$, the last column of Table 4 demonstrates that $T_i = 69.2$ and $T_1 = 69.3$ for $\rho = 0.5$ and $\rho = 0.1$. The results reveal that $\rho$ has very little effect in determining the optimal $T_i$ when $T$ is allowed very large in practice. Intuitive judgment, if the pre-determined censoring time $T$ can be allowed very large,
then the allocation time $T_1$ in low stress can be allowed enough to be a constant and not worry about no time being allocated in high stress. Thus these results are not so surprising in a sense. Unfortunately, in practice, the value of $T_1$ compared with the value of $y_1$ cannot be very large.

- $g$ decreases as $T$ increases, and when $T$ becomes large enough, $g$ will go down to a constant. It is interesting to note that for a fixed $k$, the smallest values of $g$ are the same for different $\rho$ and $\theta_1$. Moreover, the smallest values of $g$ increase as $k$ increases. If $T$ is allowed to be large in practice, it seems that $k = 2$ is more optimal than $k = 3$ under the D-optimality criterion.

- The behavior of $T_i$ as a function of the MTTF values is quite interesting. For given $k$ and $\rho$, if we want $g$ keep the same value for different $\theta_1$, then the corresponding $T$ and $T_i$ behave in a manner such that either of the ratios $T_i/\theta_1$ or $T/\theta_1$ is constant for different $\theta_1$. From a practical point of view, if $\theta_1 = a$ and the optimal change time $T_1 = b$ for a
pre-determined censoring time $T = c$, then for another $\theta_1 = \gamma \times a$ the optimal change time would be $T_1 = \gamma \times b$ for $T = \gamma \times c$. Thus the behavior of the ratios $T_i / \theta_1$ or $T / \theta_1$ coincide with the practical view.

- $T_i$ increases as $\rho$ decreases for fixed $\theta_1, k$, and $T$. Intuitively, it is more likely to observe failures in a short interval of time at the more severe condition (smaller the $\rho$), thus more test time ($T_i$ increase) will be allocated in lower stress level. Moreover, $g$ decreases as $\rho$ decreases for fixed $\theta_1, k$, and $T$. This means that the smaller $\rho$ is more optimal than the bigger $\rho$ under the D-optimality criterion.

2. V-optimal numerical result: For illustrative, we consider $\beta_{0s} = 7.09, \beta_{1s} = -0.6931$ for any $s$, and $\phi_0 = 1, \phi_1 = 2, \phi_2 = 3, \phi_3 = 4$. With different choices of $m$ and termination time $T$, the values of $T_i$ and $g_r$ are determined from Theorem 4 and Formula (13) in Section 4.3. The results are list in Tables 6 and 7.

We obtain from Tables 4 and 5 the following conclusions:

- For fixed $k$ and $m$, the optimal values $T_1$ and $T_2$ increase as the termination time $T$ increases. Furthermore, we find that with the increase of $T$, the value of $g_r$ decreases. In other words, $T$ is the bigger the better under the V-optimality criterion.

- For fixed $k$ and $T$, $g_r$ increases rapidly as $m$ (the dimension of multivariate exponential distribution) increases. Intuitively, this means that the more complicated the system (bigger $m$), the worse the estimation of parameters.

- For fixed $m$ and $T$, $g_r$ decreases as $k$ increases. For example, when $m = 2$ and $T = 100$, the values of $g_r$ are 77.4, and 58.2 for $k = 2$ and $k = 3$ respectively. It is strongly recommended that the more stresses ($k$) is the better for the life test.

6. Conclusions and areas for further research

In this paper, we obtain an optimum step-stress ALT for the multivariate exponential distribution with Type-I censoring. We also obtain the MLEs and confidence intervals for the parameters of the model. Furthermore, we show that the number of stresses should be equal or greater than the number of unknown parameters for step-stress experiments. By some simulation studies, we provide a deep insight into the way the optimal change points changes as a function of the relevant parameters. The sensitivity analysis carried out reveals that the optimal test plan is quite robust to moderate departures from the assumed value of the model parameters.

In practice, the model parameters may not be estimated precisely in a pilot study. The optimal designs depend on the preestimates of its parameters (called planning values). (See Elsayed and Zhang, 2007). To deal with the uncertainty in the planning values, we can consider some prior to describe this uncertainty. An attractive future research is to follow a Bayesian approach to planning by assigning prior distributions to experimental conditions and deriving the corresponding optimal designs. Consider the optimal design of SSALT subject to a budget constraint is also interesting areas for future research.

Acknowledgments

The authors gratefully acknowledge the Associate Editor and two anonymous referees for their valuable comments and suggestions that improve the quality of this paper.
References