A Generic Method for Modeling Accelerated Life Testing Data

Haitao Liao, PhD, University of Arizona
Huairui Guo, PhD, ReliaSoft Corporation

Key Words: accelerated life testing, Erlang-Coxian distribution, maximum likelihood estimation

SUMMARY & CONCLUSIONS

Accelerated life testing (ALT) is widely used to expedite failures of a product in a short time period for predicting the product’s reliability under normal operating conditions. The resulting ALT data are often characterized by a probability distribution, such as Weibull, Lognormal, Gamma distribution, along with a life-stress relationship. However, if the selected failure time distribution is not adequate in describing the ALT data, the resulting reliability prediction would be misleading. This paper proposes a generic method that assists engineers in modeling ALT data. The method uses Erlang-Coxian (EC) distributions, which belong to a particular subset of phase-type (PH) distributions, to approximate the underlying failure time distributions arbitrarily closely. To estimate the parameters of such an EC-based ALT model, two statistical inference approaches are proposed. First, the moment-matching approach (method of moments) is developed to simultaneously match the moments of the EC-based ALT model to the ALT data collected at all test stress levels. In addition, the maximum likelihood estimation (MLE) approach is proposed to handle ALT data with type-I censoring. A numerical example is provided to illustrate the capability of the generic method in modeling ALT data.

1 INTRODUCTION

As technology advances, new products can be made quite reliable with high-performance materials and advanced system structures through efficient design tools. As a result, it is difficult, if not impossible, to observe failures of such a product in a short time period under its normal operating conditions for reliability estimation. To cope with today’s fast paced technology innovation, accelerated life testing (ALT) has been widely used as a viable tool for estimating the long-term reliability of a new product. The basic idea of ALT is to expose some units of the product to harsher-than-normal operating conditions to expedite failures. Based on the resulting failure time data due to acceleration, a statistical model is developed and used to extrapolate the product’s long-term reliability under the normal operating conditions.

Non-parametric methods are practical choices for predicting the reliability of a product without the need of knowing the underlying failure time distribution. The most popular ones include the Kaplan-Meier estimator and Breslow estimator [1]. However, it is difficult to extend such methods to develop ALT models which require the inclusion of various life-stress relationships (distribution parameters as functions of a set of stresses, or called covariates) in order to perform extrapolation with respect to both time and stresses. Because of this, the most popular methods in modeling ALT data are to develop a parametric model in the form of a probability distribution with a set of stress-dependent parameters. Such methods, if properly applied, are relatively efficient in terms of statistical inference, and the related inference procedures, such as maximum likelihood estimation (MLE) and least squares estimation (LSE), have been extensively studied and made available to practitioners [1]-[4].

Regarding the use of parametric ALT models in practice, practitioners often face the challenge of developing or selecting a parametric ALT model that provides an adequate fit to the collected data. Technically, several probability distributions may be considered as candidates, some of which may offer comparably good fits for the same data, e.g., see [5]-[7]. To determine the best model, the likelihood values or residual plots (e.g., Cox-Snell residuals) of these models can be considered. For relevant goodness-of-fit tests of ALT models, readers are referred to [8]. Indeed, several methods have been proposed to help practitioners overcome model selection problems to some extent. For example, Elsayed et al. [9] proposed the extended linear hazard regression model that is capable of modeling various ALT data and includes many ALT models as special cases. Moreover, Yu and Chang [10] proposed a Bayesian model-averaging approach to estimating life time quantiles in ALT by combining the lognormal and Weibull log-location-scale regression models. However, all these methods require extensive experience in modeling ALT and statistical inference. To assist engineers in implementing ALT, a generic method for modeling ALT data using a more versatile distribution would be desirable for a wide range of engineering applications.

Phase-type (PH) distributions are a collection of stochastic models that represent the time to absorption of a continuous-time Markov chain (CTMC) defined on a finite-state space [11]. Figure 1 shows two examples, where the time until absorption to the respective absorbing state in each CTMC can be characterized by a PH distribution. Essentially, such a finite-state CTMC with the specific structure can be described by an infinitesimal generator matrix:

$$Q = \begin{bmatrix} 0 & 0 \\ \sigma & 0 \\ 0 & \delta \end{bmatrix}. \quad (1)$$

$$Q = \begin{bmatrix} 0 & 0 \\ \sigma & 0 \\ 0 & \delta \end{bmatrix}. \quad (1)$$
where $S$ is the subgenerator matrix consisting of the transition rates among transient states and the ones for transitions into the absorbing state(s), $S^0 = -S1$ and $1 = (1, 1, ..., 1)^t$.

**Figure 1 - CTMCs Whose Absorption Times Define the Specific Phase-type Distributions**

Let $T$ be the time to absorption of a $k$-phase CTMC and $\pi = (\pi_1, \pi_2, ..., \pi_k)$ be the initial distribution of the CTMC. The cumulative distribution function (cdf) of $T$ is:

$$F(t) = 1 - \pi \exp(tS)1,$$

where $\exp(tS)$ is matrix exponential defined as:

$$\exp(tS) = \sum_{k=0}^{\infty} \frac{1}{k!}(tS)^k.$$ (2)

The probability density function (pdf) of $T$ is given by:

$$f(t) = \pi \exp(tS)S^0,$$ (3)

and the hazard function can be expressed as:

$$h(t) = \pi \exp(tS)S^0/(\pi \exp(tS)1), \ t \geq 0.$$ (4)

In addition, the $l$th moment of the distribution is given by:

$$E[T^l] = (-1)^{l+1}! \pi S^{-l}1.$$ (5)

For examples, the CTMC in Figure 1(a) results in a three-phase Erlang distribution $E(3, \lambda)$ with:

$$\pi = (1, 0, 0) \text{ and } S = \begin{bmatrix} -\lambda & \lambda & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix},$$

and the one in Figure 1(b) gives a three-phase Coxian distribution $C(\lambda_1, \lambda_2, \lambda_3, p_1, p_2)$ with:

$$\pi = (1, 0, 0) \text{ and } S = \begin{bmatrix} -\lambda_1 & p_1\lambda_1 & 0 \\ 0 & -\lambda_2 & p_2\lambda_2 \\ 0 & 0 & -\lambda_3 \end{bmatrix},$$

where $0 < p_i \leq 1$, $i = 1, 2$.

One of the most attractive properties of PH distributions is that the set of PH distributions is dense in the set of nonnegative distributions [12]. In other words, in theory, any nonnegative distribution can be approximated arbitrarily closely by a PH distribution. The only limitation of PH distributions is that they are light-tailed, thus may not be used as effective models for heavy-tailed distributions. Among different subsets of PH distributions, Erlang distributions and Coxian distributions are probably the two most popular ones. They have been extensively studied as part of queueing theory and widely used in healthcare such as survival data analysis and modeling the length of stay of patients in hospital [13,14].

Motivated by the versatility of PH distributions that naturally meets the broad requirements for parametric modeling of failure time data, we propose a generic method using a specific and yet flexible subset of PH distributions, called Erlang-Coxian (EC) distributions [15], to model ALT data. Both moment-matching and MLE approaches are studied for statistical inference. This generic method contributes to the body of ALT literature in two ways. First, the use of EC distributions relaxes strong assumptions about the underlying failure time distributions in developing parametric ALT models. Moreover, for a specific set of ALT data, it is quite straightforward to adaptively adjust the number of phases of the associated CTMC to achieve the best fit based on statistical inference results.

## Development of EC-Based ALT Model

### 2.1 Basics of EC Distributions

Before we proceed with the proposed EC-based ALT model, we first introduce the following definitions [15].

**Definition 1.** Let $E[X^l]$ be the $l$th moment of random variable $X$ with distribution $G$. The normalized $l$th moment $m^G_l$ of $X$ for $l = 2, 3$ is defined as: $m^G_2 = \frac{E[X^2]}{\sigma^2}$ and $m^G_3 = \frac{E[X^3]}{\sigma^3}$.

**Definition 2.** A distribution $G$ is well represented by a distribution $F$ if $F$ and $G$ agree on their first three moments. $PH_k$ refers to the set of distributions that are well represented by a PH distribution.

It is known that a distribution $G$ is in $PH_k$ if and only if its normalized moments satisfy $m^G_2 > m^G_3 > 1$ [16]. Since any nonnegative distribution $G$ satisfies $m^G_2 \geq m^G_3 \geq 1$ [17], almost all the nonnegative distributions are in $PH_3$. Osogami and Harchol-Balter [15] introduced the EC distributions, which are quite efficient for approximating $PH_k$ distributions. Figure 2 shows the CTMC for a $k$-phase EC distribution ($k \geq 3$), which consists of an Erlang with $k-2$ phases: $E(k - 2, \lambda_1)$ and a two-phase Coxian: $C(\lambda_2, \lambda_3, p_2)$, for which:

$$\pi = (1, 0, ..., 0).$$

**Figure 2 - CTMC for a k-phase EC Distribution**
The idea of creating an EC distribution is that a two-phase Coxian distribution can well represent any distribution that has high second and third moments while an Erlang distribution has only two free parameters and has the least normalized second moment among all the PH distributions with a fixed number of phases. Therefore, such a k-phase EC distribution can represent probability distributions with all ranges of variability using only a small number of phases. This is important in developing a PH-based ALT model, which needs additional parameters for the life-stress relationship.

2.2 Formulation of EC-based ALT Model

Accelerated failure time (AFT) models are probably the most widely used parametric ALT models. Unlike a proportional hazards model, an AFT model assumes that the effect of a covariate is to multiply the failure time, instead of the hazard function, by some constant [9]. Mathematically, an AFT model for the cdf $F(t; Z)$ of failure time under a constant stress $Z$ can be expressed as [8]:

$$F(t; Z) = F_0(r(Z; \theta)t), \quad (8)$$

where $F_0(\cdot)$ is the baseline cdf and $r(Z; \theta)$ is a deterministic function of $Z$. Equivalently, this model can be expressed in terms of the corresponding hazard rate:

$$h(t; Z) = h_0(r(Z; \theta)t)r(Z; \theta), \quad (9)$$

where $h_0(\cdot)$ is the corresponding baseline hazard rate. An important aspect of implementing an AFT model is to assume the underlying failure time distribution, i.e., $F_0(\cdot)$.

Following Eqs. (2) and (8), the EC-based AFT model for a product’s failure time under stress level $Z$ is given by:

$$F(t; Z) = F_0(r(Z; \theta)t) = 1 - \pi \exp\{r(Z; \theta)tS\}1, \quad (10)$$

where $\pi$ and $S$ are given in Eq. (9). The corresponding pdf is:

$$f(t; Z) = r(Z; \theta)\pi \exp\{r(Z; \theta)tS\} S^\alpha, \quad (11)$$

and the hazard function is:

$$h(t; Z) = r(Z; \theta)\pi \exp\{r(Z; \theta)tS\} S^\alpha/ (\pi \exp\{r(Z; \theta)tS\}1). \quad (12)$$

The corresponding $l$th moments can be expressed as:

$$E[T^2_i] = (-1)^l(r(Z; \theta))^l\pi S^\alpha1, \quad (13)$$

3 STATISTICAL INFERENCE

This section addresses two statistical methods for estimating the parameters of an EC-based ALT model.

3.1 Moment-Matching (Method of Moments)

For complete ALT data, the $l$th sample moment of failure times $t_{ij}$ at stress level $Z_i$ can be obtained as:

$$\frac{1}{n_i} \sum_{j=1}^{n_i} t_{ij}^l.$$  

The mathematical formulation for determining the matching EC-based AFT model with the least number of phases (i.e., $k$) can be expressed as:

$$\begin{align*}
\text{Min} & \quad k \\
\text{Subject to} & \quad (-1)^i(r(Z_i; \theta))^{-1}!S^{-1}1 - \frac{1}{n_i} \sum_{j=1}^{n_i} t_{ij}^l \leq \epsilon_{i,1}, \\
& \quad \frac{1}{n_i} \sum_{j=1}^{n_i} t_{ij}^l \leq \epsilon_{i,2}, \\
& \quad (-1)^i(r(Z_i; \theta))^{-2}!S^{-2}1 - \frac{1}{n_i} \sum_{j=1}^{n_i} t_{ij}^l \leq \epsilon_{i,3}, \quad i = 1, ..., M, \\
& \quad \frac{1}{n_i} \sum_{j=1}^{n_i} t_{ij}^l \leq \epsilon_{i,3}, \quad i = 1, ..., M, \\
\end{align*}$$

where $\epsilon_{i,l}$ are the pre-specified levels of tolerance for each stress level $Z_i$. Note that the size of matrix $S$ increases as the value of $k$ increases, which only increases the number phases in Erlang $E(k-2, \lambda_i)$. The limitation of this approach is that extensive computational effort must be taken to find the solution that satisfies these nonlinear constraints, which may not be an easy task.

3.2 MLE method

We now consider a more complex case of constant-stress ALT experiment with Type-I censoring. Again, let $t_{ij}$ be the recorded failure/censoring time of unit $j$ tested under stress level $Z_i$, $i = 1, 2, ..., M$, and $n_i$ be the total number of units tested under $Z_i$.

The MLE method is widely used, which can handle different types of censored data quite efficiently. For ALT data with type-I censoring, the log-likelihood function $\ln L$ can be expressed as:

$$\ln L(k, \lambda_1, \lambda_2, \lambda_3, p_c, \theta) = \sum_{i=1}^{M} \sum_{j=1}^{n_i} [\delta_{ij}\log f(t_{ij}; Z_i)]$$

$$+(1 - \delta_{ij})\log (1 - F(t_{ij}; Z_i))$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n_i} \delta_{ij}\log [r(Z_i; \theta)\pi \exp\{r(Z_i; \theta)t_{ij}S\} S^\alpha]$$

$$+(1 - \delta_{ij})\log (\pi \exp\{r(Z_i; \theta)t_{ij}S\}1), \quad (15)$$

where $\delta_{ij} = \{1, \text{if } t_{ij} \text{ is a failure time}; 0, \text{otherwise}\}$. The MLEs of the model parameters can be obtained by maximizing the log-likelihood function. In practice, different optimization algorithms, such as Quasi-Newton and Nelder-Mead algorithms, can be used.

3.3 Determination of the Number of Phases (Value of $k$)

A practical issue in developing an EC-based ALT model based on likelihood is to determine the number of phases, i.e., the value of $k$. For such model selection problems, the Akaike information criterion (AIC) can be utilized, which is expressed as:

$$AIC = 2q - 2\ln L(k, \lambda_1, \lambda_2, \lambda_3, p_c, \theta), \quad (16)$$

where $q$ is the number of parameters, which is the same for EC-based ALT models when the number of parameters in $\theta$ is
fixed. As a result, it is straightforward to compare the likelihood values when comparing candidate EC-based ALT models with different numbers of phases in order to determine the best one.

Note that the likelihood-ratio test has been widely used to test against nested models, which may not be appropriate for determining the number of phases in the EC-based ALT model.

4 NUMERICAL EXAMPLE

The ALT data reported by Liao and Elsayed [18] is used to illustrate the use of the proposed method in practice.

4.1 Experiment

The purpose of this ALT experiment is to estimate the reliability of a type of miniature lamps under the use condition: 2 volts. The highest operating voltage of the lamp is 5 volts. It is well known that the coil temperature of an incandescent lamp during operation is mainly due to the electric current. Three constant voltage levels were utilized in the experiment: 5 volts, 3.5 volts, and 2 volts. Table 1 gives the observed failure times and censoring times under the three stress levels.

<table>
<thead>
<tr>
<th>Stress level</th>
<th>Failure times (&quot;+&quot;: the unit is censored)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 volts</td>
<td>20.5 22.3 23.2 24.7 26 34.1 39.6 41.8 43.6 44.9 47.7 61.6 62.1 65.5 70.8 87.8 118.3 120.1 145.4 157.4 180.9 187.7 204 206.7 213.9 215.2 218.7 254.1 262.6 293 304 313.7 314.1 317.9 337.7 430.2</td>
</tr>
<tr>
<td>3.5 volts</td>
<td>37.8 43.6 51.1 58.6 65.5 65.9 75.6 82.5 88.1 89 106.6 113.1 121.1 121.5 128.3 151.8 171.7 181 202.7 211.7 230.7 249.9 275.6 285 296.2 358.5 379.8 434.5 493.1 506.4 561.1 570 577.7 876.3 922 890+ 890+ 941+ 941+</td>
</tr>
<tr>
<td>2 volts</td>
<td>223.1 254 316.7 560.2 679 737 894.4 930.5+ 930.5+ 930.5+ 930.5+ 930.5+ 930.5+ 930.5+ 930.5+ 930.5+ 930.5+</td>
</tr>
</tbody>
</table>

Table 1 - ALT Data of Miniature Lamps

Figure 3 shows the empirical cdf (Kaplan-Meier) of the lamp under each different voltage level. To avoid making an assumption on the underlying distribution, such as Weibull and Lognormal, we use the proposed EC-based ALT model to predict the reliability of this type of miniature lamps.

4.2 EC-based ALT Model and Estimation Results

To facilitate data analysis, we standardize the stress levels by defining $Z_i = [V_i - V_0]_+/ [V_H - V_0]$, where $V_0 = 2$ volts and $V_H = 5$ volts. As a result, we have: $Z_1 = 1$, $Z_2 = 0.5$, and $Z_3 = 0$. The life-stress relationship is assumed to be characterized by:

$$r(Z_i; \theta) = \exp (\alpha Z_i^{\beta})$$

Because the data set contains censoring times, we use the MLE method introduced in section 3.2 for statistical inference. Table 2 shows the MLEs of model parameters for different EC-based ALT models with different numbers of phases. By comparing the log-likelihood values, the EC-based ALT model with $k = 5$ phases is selected after balancing the prediction accuracy and the complexity of the models.

| Values of $k$ | MLEs of parameters | Log-likelihood
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3: $E(1, \lambda_i)$ &amp; $C(\lambda_2, \lambda_3, p_c)$</td>
<td>$\alpha_0 = 2.9091$; $\alpha_1 = 0.5762$; $\lambda_1 = 0.0026$; $\lambda_2 = 0.0026$; $\lambda_3 = 0.0003$; $p_c = 0.6730$</td>
<td>-518.5038</td>
</tr>
<tr>
<td>4: $E(2, \lambda_i)$ &amp; $C(\lambda_2, \lambda_3, p_c)$</td>
<td>$\alpha_0 = 2.8807$; $\alpha_1 = 0.5730$; $\lambda_1 = 0.0045$; $\lambda_2 = 0.0045$; $\lambda_3 = 0.0003$; $p_c = 0.6980$</td>
<td>-516.4058</td>
</tr>
<tr>
<td>5: $E(3, \lambda_i)$ &amp; $C(\lambda_2, \lambda_3, p_c)$</td>
<td>$\alpha_0 = 2.8182$; $\alpha_1 = 0.5693$; $\lambda_1 = 0.0068$; $\lambda_2 = 0.0068$; $\lambda_3 = 0.0004$; $p_c = 0.7269$</td>
<td>-515.4942</td>
</tr>
<tr>
<td>6: $E(4, \lambda_i)$ &amp; $C(\lambda_2, \lambda_3, p_c)$</td>
<td>$\alpha_0 = 2.7463$; $\alpha_1 = 0.5747$; $\lambda_1 = 0.0097$</td>
<td>-515.0856</td>
</tr>
</tbody>
</table>
Table 2 - MLEs of Parameters for Different EC-based ALT Models with Different Numbers of Phases (k)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_2 )</td>
<td>0.0097</td>
</tr>
<tr>
<td>( \lambda_3 )</td>
<td>0.0004</td>
</tr>
<tr>
<td>( p_k )</td>
<td>0.7545</td>
</tr>
</tbody>
</table>

Figure 4 illustrates the statistical fittings of the resulting EC-based ALT model for the three test stress levels. Compared to the corresponding empirical cdf’s, this model exhibits satisfactory prediction capability. Figure 5 shows the predicted reliability function, pdf and hazard rate of the lamp under the normal operating condition. The mean-time-to-failure can be easily obtained as 2405.5 hours.

5 CONCLUSIONS

Although the applications of PH distributions have been studied in other areas, such as queueing systems and healthcare, they have not been used in modeling ALT. This paper introduces a generic method for modeling ALT data using EC distributions, which belong to an important subset of PH distributions. Without assuming other particular probability distributions for failure times, such as extreme value distributions (Gumbel, Weibull, and Fréchet distributions), lognormal distribution, and mixture of distributions, this method leads to an EC-based ALT model which can well represents the underlying failure time distribution that may be unknown and/or difficult to verify. Both moment-matching approach and MLE approach are developed for parameter estimation. The technical contribution of this paper is to demonstrate, for the first time, the potential of using PH distributions in developing ALT models. The numerical example demonstrates that the method indeed provides practitioners with a powerful tool for modeling ALT data.

ACKNOWLEDGMENT

This work is supported in part by the U.S. National Science Foundation under grants CMMI-1238301 and CMMI-1238304.

REFERENCES


**BIOGRAPHIES**

Haitao Liao, PhD
Systems and Industrial Engineering Department
The University of Arizona
1127 E. James E. Rogers Way
Tucson, Arizona 85721, USA

e-mail: hliao@email.arizona.edu

Dr. Haitao Liao is an Associate Professor in the Systems and Industrial Engineering Department at the University of Arizona (UofA), Tucson, Arizona. He is also the Director of Reliability & Intelligent Systems Engineering (RISE) Laboratory (http://sie.engr.arizona.edu/reliability-intelligent-systems-engineering-laboratory) at UofA. He received his Ph.D. in Industrial and Systems Engineering from Rutgers University, New Jersey. His research interests focus on modeling of accelerated testing, probabilistic risk assessment, maintenance models and optimization, service part inventory control, and prognostics. His current research is sponsored by the U.S. National Science Foundation, Department of Energy, and industry. He was a recipient of the National Science Foundation CAREER Award in 2010. He is a member of IIE and INFORMS.

Huairui Guo, PhD
ReliaSoft Corporation
1450 S. Eastside Loop
Tucson, AZ 85710, USA

e-mail: Harry.Guo@ReliaSoft.com

Dr. Huairui (Harry) Guo is the Director of the Theoretical Development Department at ReliaSoft Corporation. He received his Ph.D. in Systems & Industrial Engineering and M.S. in Reliability & Quality engineering; both from the University of Arizona. His research and publications cover reliability areas, such as life data analysis, repairable system modeling and reliability test planning, and quality areas, such as process monitoring, analysis of variance and design of experiments. In addition to research and product development, he is also part of the training and consulting arm and has been involved in various projects from automobile, medical device, oil and gas, and aerospace industry. He is a certified reliability professional (C.R.P), ASQ certified CQE. He is a member of IIE, SRE and ASQ.