A Robust Approach for Estimating Field Reliability Using Aggregate Failure Time Data

Samira Karimi, PhD Student, University of Arkansas
Haitao Liao, PhD, University of Arkansas
Ed Pohl, PhD, University of Arkansas

Key Words: Aggregate Data, Confidence Interval, Coxian Distribution, EM Algorithm, Fisher Information, Phase-type Distributions

SUMMARY & CONCLUSIONS

Failure time data of fielded systems are usually obtained from the actual users of the systems. Due to various operational preferences and/or technical obstacles, a large proportion of field data are collected as aggregate data instead of the exact failure times of individual units. The challenge of using such data is that the obtained information is more concise but less precise in comparison to using individual failure times. The most significant needs in modeling aggregate failure time data are the selection of an appropriate probability distribution and the development of a statistical inference procedure capable of handling data aggregation. Although some probability distributions, such as the Gamma and Inverse Gaussian distributions, have well-known closed-form expressions for the probability density function for aggregate data, the use of such distributions limits the applications in field reliability estimation. For reliability practitioners, it would be invaluable to use a robust approach to handle aggregate failure time data without being limited to a small number of probability distributions. This paper studies the application of phase-type (PH) distribution as a candidate for modeling aggregate failure time data. An expectation-maximization algorithm is developed to obtain the maximum likelihood estimates of model parameters, and the confidence interval for the reliability estimate is also obtained. The simulation and numerical studies show that the robust approach is quite powerful because of the high capability of PH distribution in mimicking a variety of probability distributions. In the area of reliability engineering, there is limited work on modeling aggregate data for field reliability estimation. The analytical and statistical inference methods described in this work provide a robust tool for analyzing aggregate failure time data for the first time.

1 INTRODUCTION

Failure time data of fielded systems are usually collected by actual users. Analyzing such data is very valuable because they reflect the impacts of actual environmental conditions on the products, which are difficult, if not impossible, to be imposed in laboratory [1]. However, this type of data is sometimes difficult to analyze due to the special structure of the data. In practice, a large number of data of fielded systems are not the exact failure times of individual units. Instead, the recorded data are usually the cumulative operating hours of each system, along with the number of failures of a certain component in the system, called aggregate data.

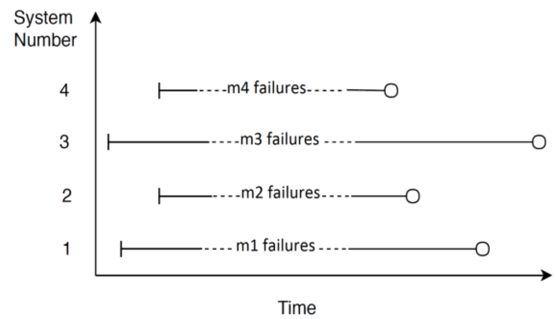

Figure 1. Aggregate data: The number of failures and the cumulative operating time of each system are known but the actual failure time of each component is unknown.

Figure 1 shows a schematic of an aggregate data set. Since this type of data collection mechanism can be accessible for many components that are difficult or expensive to collect, there is a need to develop statistical methods for practitioners to analyze such data for field reliability estimation. Due to the structure of aggregate data, for most of the probability distributions, it is intractable to find a closed-form expression for the distribution of aggregate data. In the literature, only a few distributions, such as exponential, Normal, Gamma and Inverse Gaussian (IG), have been utilized for the analysis of aggregate lifetime data. Indeed, a primary reason for practitioners to collect aggregate data is the assumption that the underlying failure time distribution is exponential. Coit and Dey [2] showed that the exponential distribution assumption is often invalid through a hypothesis test. Coit and Jin [1] presented a useful method for Gamma distribution for reliability analysis using aggregate data. They used a quasi-Newton method for parameter estimation. In the literature, IG distribution is another probability distribution that has been applied to aggregate data analysis. Chen and Ye [3] investigated aggregate data further to provide confidence intervals for the parameters of Gamma and IG distributions. They also compared the reliability estimates using Gamma, IG and Normal distributions, for which the Normal distribution assumption did not provide adequate results. Moreover, interval estimation has also been studied for Gamma distribution based on individual failure times [4], [5], [6].

In this paper, we propose the use of PH distributions for reliability estimation based on aggregate failure time data. In practice, the use of PH distributions provides a robust approach for reliability estimation when the underlying failure time distribution is unknown. For example, Pohl and Dietrich [7] presented a multi-level environmental stress screening model for multi-component electronic systems using a PH distribution. Indeed, PH distributions are quite flexible in modeling the probability distributions of a variety of non-negative random variables. To estimate the parameters of PH distribution based on individual data, an Expectation-Maximization (EM) algorithm has been used. EM is originally a parameter estimation method for a statistical model based on data with missing values. Asmussen et al. [8] provided several extended formulas of EM algorithm for specific application of PH distributions. PH distributions have already been used for reliability estimation purposes. Liao and Guo [9], and Liao and Karimi [10] have considered ALT data analysis based on PH distributions. Riascos-Ochoa et al. [11], and Kharoufeh et al. [12] have studied the use of PH distributions in degradation-based models. In this work, PH distributions are employed, for the first time, in reliability estimation as well as confidence interval estimation using aggregate failure time data.

The remainder of this paper is organized as follows. In Section 2, PH distribution and the special case applied for our numerical study is explained. In Section 3, we present the EM algorithm used to estimate the parameters of PH distribution for individual lifetime data and extend the formulas to adapt to aggregate data. A simulation study as well as an analysis of a real data set is presented in Section 4. Conclusions are presented in Section 5.

2 PH DISTRIBUTION

In this section, we give a short introduction to PH distributions and a special case. Consider a continuous-time Markov Chain (CTMC) $\{Y(t)\}_{t>0}$ with finite states {1, 2, …, N, N+1}, where state N+1 is the absorbing state and the rest of the states are transient states. The infinitesimal generator of $\{Y(t)\}_{t>0}$ is:

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{S} & \boldsymbol{q} \\ \boldsymbol{0} & 0 \end{bmatrix}, \tag{1}$$

where $\boldsymbol{S}$ is called the transition rate matrix, $\boldsymbol{q} = -\boldsymbol{S}\boldsymbol{e}$ is the absorption rate matrix, and $\boldsymbol{e} = [1,1,...,1]'$. The PH distribution associated with the CTMC is the probability distribution of time from entering the CTMC until absorption. Specifically, the probability density function (pdf) of the PH distribution is:

$$f(t) = \boldsymbol{\pi} e^{\boldsymbol{S}t}\boldsymbol{q}, \tag{2}$$

where $\boldsymbol{\pi}$ is the matrix of initial probabilities, which denotes the probabilities that the process starts from any of those phases. The following is the transition rate matrix of an acyclic PH distribution that is appropriate for modeling lifetime data:

$$\boldsymbol{S} = \begin{bmatrix} -\lambda_1 & p_{12}\lambda_1 & p_{13}\lambda_1 & \cdots & p_{1n}\lambda_1 \\ 0 & -\lambda_2 & p_{23}\lambda_2 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\lambda_{n-1} & p_{(n-1)n}\lambda_{n-1} \\ 0 & \cdots & \cdots & 0 & -\lambda_n \end{bmatrix}. \tag{3}$$

An important special case of acyclic PH distribution is the Coxian distribution, which is very flexible case of PH distribution with a small number of parameters. This distribution will be used in this study. The Coxian distribution has the following transition rate and absorption rate matrices:

$$\boldsymbol{S} = \begin{bmatrix} -(\lambda_1+\mu_1) & \lambda_1 & \cdots & 0 \\ 0 & -(\lambda_2+\mu_2) & \lambda_2 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & -\mu_n \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}. \tag{4}$$

Figure 2 shows the CTMC of a three-phase Coxian distribution. It can be proved that a variety of PH distributions can be converted to an equivalent Coxian distribution making the Coxian distribution a robust candidate for failure time data analysis.

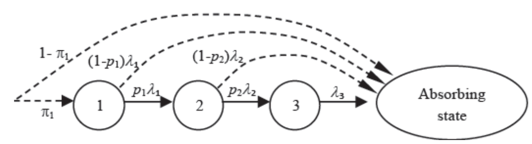

Figure 2. CTMC for a three-phase Coxian distribution.

For aggregate data, suppose that the failure time of each component follows a Coxian distribution with transition rate matrix $\boldsymbol{S}$ given by Eq. (4). Then the transition rate matrix for the sum of m such random variables (i.e., an aggregate data point consisting of m component failures) can be expressed as:

$$\boldsymbol{S^{new}} = \begin{bmatrix} \boldsymbol{S} & \boldsymbol{q\pi} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S} & \boldsymbol{q\pi} & \boldsymbol{0} \\ \vdots & \boldsymbol{0} & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{S} \end{bmatrix}. \tag{5}$$

Note that the size of the matrix depends on the number of failures (i.e., m) for each aggregate data point.

3 STATISTICAL INFERENCE

3.1 Maximum likelihood method

Asmussen et al. [8] developed an EM algorithm to estimate the parameters of a PH distribution. This method regards the number of times the corresponding Markov process starts in phase i, the number of jumps from phase i to phase j, and the total time spent in phase i as the missing values. These quantities are denoted by B_i, N_{ij} and Z_i, respectively. The likelihood function for a data set can be expressed as:

$$l((\boldsymbol{\pi},\boldsymbol{S})|\boldsymbol{\tau}) = \prod_{i=1}^{n}\pi(i)^{B_i}\prod_{i=1}^{n}e^{Z_i(i,i)}\prod_{i=1}^{n}\prod_{j=1}^{n+1}S(i,j)^{N_{ij}}. \quad (6)$$

The EM algorithm consists of two main steps. The expectation step, where the missing variables are estimated based on the current assumed parameter values, and the maximization step, where the parameters of the distribution are updated based on the current estimates of the missing values, are performed iteratively until the convergence criterion is met.

Technically, the expectation step formulas for a data set consisting of exact failure times of individual units are:

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[B_i] = \frac{1}{M}\sum_{k=1}^{M}\frac{\boldsymbol{\pi}(i)\boldsymbol{b}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}(i)}{\boldsymbol{\pi}\boldsymbol{b}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}} \quad (7)$$

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[Z_i] = \frac{1}{M}\sum_{k=1}^{M}\frac{\boldsymbol{G}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}(i,i)}{\boldsymbol{\pi}\boldsymbol{b}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}} \quad (8)$$

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[N_{ij}] = \frac{1}{M}\sum_{k=1}^{M}\frac{\boldsymbol{G}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}(i,j)\boldsymbol{S}(i,j)}{\boldsymbol{\pi}\boldsymbol{b}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}} \quad (9)$$

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[N_{in+1}] = \frac{1}{M}\sum_{k=1}^{M}\frac{\boldsymbol{g}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}(i)\boldsymbol{q}(i)}{\boldsymbol{\pi}\boldsymbol{b}_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{t_k}}} \quad (10)$$

where M is the number of data points, n is the number of phases of the PH distribution, m_k is the number of failures for data point k, and the other EM statistics are defined as:

$$\boldsymbol{g}_{(\boldsymbol{\pi},\boldsymbol{S}),t} = \boldsymbol{\pi}e^{\boldsymbol{S}t}, \boldsymbol{b}_{(\boldsymbol{\pi},\boldsymbol{S}),t} = e^{\boldsymbol{S}t}\boldsymbol{q},$$
$$\boldsymbol{G}_{(\boldsymbol{\pi},\boldsymbol{S}),t} = \int_0^t (\boldsymbol{g}_{(\boldsymbol{\pi},\boldsymbol{S}),t-u})^T(\boldsymbol{b}_{(\boldsymbol{\pi},\boldsymbol{S}),u})^T du. \quad (11)$$

Note that the regular EM algorithm for PH distribution is developed based on data points that follow the same probability distribution. Nonetheless, considering a similar PH distribution for the lifetime of individual components, aggregate data points have PH distributions with different numbers of phases (i.e., convolution of different numbers of non-negative random variables). To handle aggregate failure time data, the formulas in the expectation step need to be modified for complying with the structure of each aggregate data point. As the formulas of the previous EM algorithm cannot accept these differences, new formulas should be developed for the expectation step. Based on the pattern that can be seen in the transition rate matrix (see Eq. (5)) for aggregate data, the trends in other statistics of EM algorithm are also discoverable, leading us to the idea of decomposing all the related matrices to their submatrices, each referring to one hidden component. As a result, the following updated formulas for the expectation step are developed:

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[B_i] = \frac{1}{M}\sum_{k=1}^{p}\sum_{l=0}^{m_k}\frac{\boldsymbol{\pi}^{(k)}(i+ln)\boldsymbol{b}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}(i+ln)}{\boldsymbol{\pi}^{(k)}\boldsymbol{b}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}} \quad (12)$$

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[Z_i] = \frac{1}{M}\sum_{k=1}^{p}\sum_{l=0}^{m_k}\frac{\boldsymbol{G}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}(i+ln,i+ln)}{\boldsymbol{\pi}^{(k)}\boldsymbol{b}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}} \quad (13)$$

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[N_{ij}] = \frac{1}{M}\sum_{k=1}^{p}\sum_{l=0}^{m_k}\frac{\boldsymbol{G}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}(i+ln,i+ln)\boldsymbol{S}^{(k)}(i+ln,i+ln)}{\boldsymbol{\pi}^{(k)}\boldsymbol{b}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}} \quad (14)$$

$$E_{(\boldsymbol{\pi},\boldsymbol{S}),\tau}[N_{in+1}] = \frac{1}{M}\sum_{k=1}^{p}\sum_{l=0}^{m_k}\frac{\boldsymbol{g}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}(i+ln)\boldsymbol{d}^{(k)}(i+ln)}{\boldsymbol{\pi}^{(k)}\boldsymbol{b}_{(\boldsymbol{\pi}^{(k)},\boldsymbol{S}^{(k)}),\boldsymbol{t_k}}} \quad (15)$$

where p is the number of data points, and $\boldsymbol{d}^{(k)}$ is a different version of absorption rate matrix for data point k. Considering the absorption rate of each individual component as $d_i = \sum_{j=1}^{n}\boldsymbol{q}(i)\boldsymbol{\pi}(j)$ leads to:

$$\boldsymbol{d}^{(k)} = [d_1^{(k)}, \dots, d_N^{(k)}, \dots, d_1^{(k)}, \dots, d_N^{(k)}]_{1\times nm_k}. \quad (16)$$

For the maximization step, the related formulas remain unchanged as long as the matrices for the corresponding aggregate data point are used:

$$\hat{\pi}(i) = E_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{\tau}}[B_i], \widehat{\boldsymbol{S}}(i,j) = \frac{E_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{\tau}}[N_{ij}]}{E_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{\tau}}[Z_i]},$$
$$\widehat{\boldsymbol{q}}(i) = \frac{E_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{\tau}}[N_{in+1}]}{E_{(\boldsymbol{\pi},\boldsymbol{S}),\boldsymbol{\tau}}[Z_i]}, \quad (17)$$
$$\widehat{\boldsymbol{S}}(i,i) = -\left(\widehat{\boldsymbol{q}}(i) + \sum_{i\neq j}^{n}\widehat{\boldsymbol{S}}(i,j)\right).$$

3.2 *ML confidence interval estimation*

Bladt et al. [13] came up with interval estimation for PH distributions using Fisher information matrix with an EM algorithm and a Newton-Raphson method. Their Newton-Raphson method is extended for aggregate data in this paper. To calculate the contribution of each data point to the second derivative of likelihood function, the pdf should be determined based on the number of failures of each data point, as stated previously. Moreover, the derivative of the transition rate matrix with respect to each element has the same pattern for one data point but is different for different data points.

In this paper, the confidence intervals are calculated using Wald's method. Specifically, Wald's statistic is defined as:

$$W = (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'[\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}]^{-1}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}), \quad (18)$$

where $\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\theta}}}$ is the estimated variance-covariance matrix, $W \sim \chi_v^2$ and v is the number of parameters. Note that $W \leq \chi_v^2$ shows the confidence region as an ellipsoid for the model parameters.

4 NUMERICAL STUDY

4.1 Simulation study

This section illustrates the flexibility of the PH distribution in analyzing aggregate data with an arbitrary underlying distribution. To do this, a set of 6 aggregate data points and a set of 12 aggregate data points are generated from the Gamma distribution $\Gamma(2.5,3.5)$ for a total of 36 and 56 component failures, respectively.

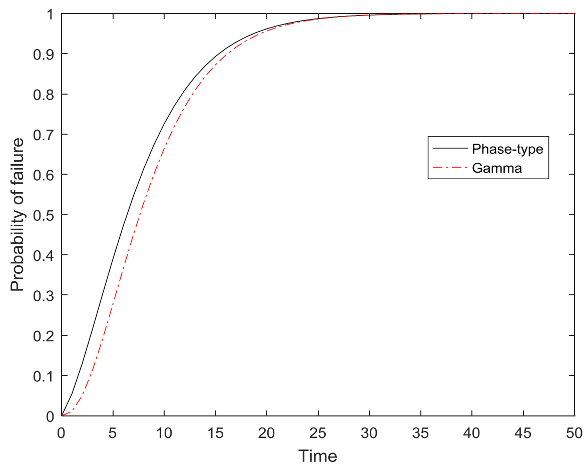

Figure 3. A three-phase Coxian distribution CDF estimated from the data generated from a $\Gamma(2.5,3.5)$; 6 aggregate data points for a total of 36 failures.

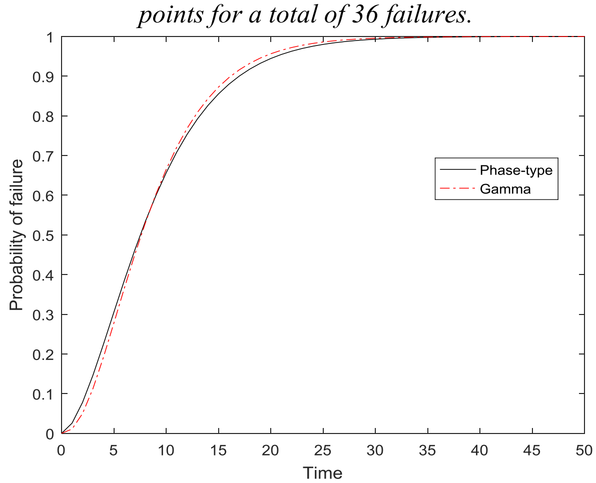

Figure 4. A three-phase Coxian distribution CDF estimated from the data generated from a $\Gamma(2.5,3.5)$; 12 aggregate data points for a total of 56 failures.

Figure 3 shows how well a three-phase Coxian distribution is capable of estimating the underlying distribution with 6 data points. Figure 4 shows that as the number of data points increases, the difference between true distribution (i.e., the Gamma distribution) and the estimated distribution (Coxian in this example) can be hardly recognized.

4.2 Real world example

We performed our proposed method on a real data set of aircraft indicator lights (see Table 1) from the Reliability Information Analysis Center (RIAC), formerly known as Reliability Analysis Center (RAC). Previously, researchers [1] [3] have considered the same set of data in their studies. For our numerical study, a Coxian distribution is used. Note that our model is capable of dealing with general PH distributions, however, to reduce the computational time of our algorithm, a Coxian is selected. Here, our estimated component failure time distribution is depicted and compared with the three distributions studied by others. Figure 5 illustrates the estimated cumulative distribution functions (CDF) of the PH distribution against the Gamma, IG and Normal distributions obtained in [3].

System number	Number of failures	Cumulative operating hours
1	2	51000
2	9	194900
3	8	45300
4	8	112400
5	6	104000
6	5	44800

Table 1. Aggregate failure data of aircraft indicator lights

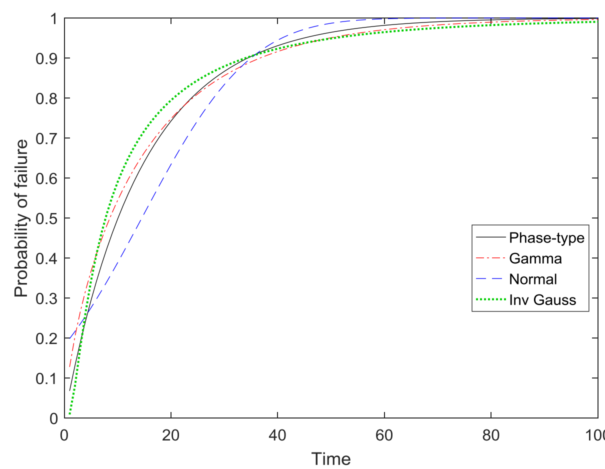

Figure 5. Gamma, IG, Normal and three-phase Coxian distributions estimated failure CDF's for the data in Table 1.

As seen in Figure 5, the estimated CDF of the PH distribution is close to Gamma and IG alternatives. Both of these distributions have been recognized proper for dealing with aggregate data. However, the results from the Normal distribution has a high coefficient of variation [3]. Based on these results, the PH distribution can be introduced as a more flexible alternative for estimating the failure distributions based on aggregate data.

Moreover, using the same data set and the Newton-Raphson method and the fisher information, individual and simultaneous confidence regions of the model parameters are estimated. A three-phase Coxian distribution has 5 unknown parameters, for which the variance-covariance matrix can be calculated by taking the inverse of the Fisher information matrix as:

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\theta}}} = \begin{bmatrix} 0.00018 & -0.00022 & 0.003515 & -0.00848 & 0.000154 \\ -0.00022 & 0.000203 & -0.00443 & 0.010321 & -0.00023 \\ 0.003515 & -0.00443 & 0.120267 & -0.2326 & -0.01246 \\ -0.00848 & 0.010321 & -0.2326 & 0.517358 & -0.01074 \\ 0.000154 & -0.00023 & -0.01246 & -0.01074 & 0.035756 \end{bmatrix}, \quad (19)$$

where $\boldsymbol{\theta} = (\mu_1, \lambda_1, \mu_2, \lambda_2, \mu_3)$. Individual confidence intervals (C.I.'s) are estimated using Wald's statistic considering other parameters as known and constant (see Table 2). In addition, simultaneous confidence regions are demonstrated for $\mu's$, in Figure 6 and $\lambda's$, in Figure 7.

Parameter	ML Estimate	90% MLE C.I.
μ_1	0.0702	(0,4.783)
μ_2	0.0431	(0,0.225)
μ_3	0.0823	(0,0.417)
λ_1	0.0121	(0, 4.456)
λ_2	0.0392	(0,0.127)

Table 2. MLEs of three-phase Coxian distribution parameters and confidence intervals based on the data from Table 1.

Based on the estimated model parameters and their variance-covariance matrix, the variance of the estimated CDF of failure time can be calculated using the delta method as:

$$\widehat{Var}\big(F(\boldsymbol{t};\boldsymbol{\theta})\big) = \left[\frac{\partial F}{\partial \mu_1}, \dots, \frac{\partial F}{\partial \lambda_2}\right] \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\theta}}} \left[\frac{\partial F}{\partial \mu_1}, \dots, \frac{\partial F}{\partial \lambda_2}\right]^T. \quad (20)$$

Figure 8 shows the 90% C.I. for the CDF.

The computation was performed in Matlab 2017b on a desktop computer with Core™ i5-6300HQ CPU and 8 GB RAM. Fifty-seven seconds were used for parameter estimation and twenty-two seconds were spent on the estimation of the variance-covariance matrix of model parameters.

5 CONCLUSIONS

Collecting reliability data from fielded systems may result in aggregate lifetime data. Analyzing aggregate data relies on the special characteristics in the distribution as well as statistical inference method. Because of these complexities, only a few distributions, such as exponential, Gamma and IG, have been introduced for aggregate data analysis. To augment the applications of aggregate data, in this work, a PH distribution is studied for the first time to analyze aggregate lifetime data. An important characteristic of the PH distribution is its high flexibility for fitting data with an arbitrary underlying distribution by controlling the number of phases. For parameter estimation, since the existing EM algorithm cannot be directly applied for aggregate data, new EM formulas are developed in this paper for aggregate data analysis. In addition to point estimation of model parameters, ML-based confidence intervals are obtained with the assistance of the Fisher information matrix. In our simulation study, aggregate data was generated from a Gamma distribution, and the PH distribution was able to estimate the underlying distribution meticulously. Moreover, performing our method on a set of aggregate reliability data from RIAC also depicted the strength of PH distribution in comparison to other alternatives previously studied.

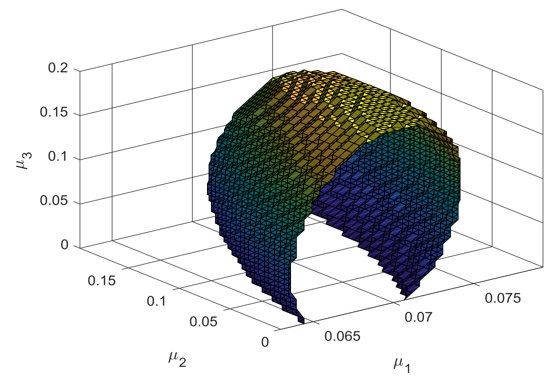

Figure 6. Wald's simultaneous confidence regions for μ of a three-phase Coxian in the form of eq. (4).

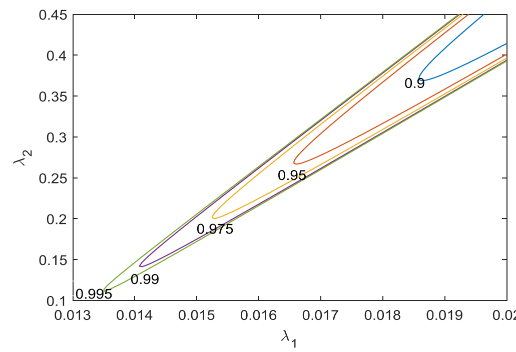

Figure 7. Wald's simultaneous confidence regions for λ of a three-phase Coxian in the form of eq. (4).

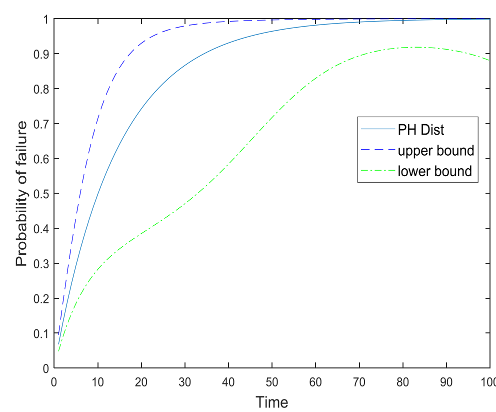

Figure 8. 90% confidence interval for the failure CDF.

AKNOWLEDGMENTS

This research is sponsored by the U.S. National Science Foundation under the grant #1635379.

REFERENCES

[1] D. W. Coit and T. Jin, "Gamma distribution parameter estimation for field reliability data with missing failure times," *IIE Transactions,* vol. 32, no. 12, pp. 1161-1166, 2000.

[2] D. W. Coit and K. A. Dey, "Analysis of grouped data from field-failure reporting systems," *Reliability Engineering & System Safety,* vol. 65, no. 2, pp. 95-101, 1999.

[3] P. Chen and Z. S. Ye, "Estimation of field reliability based on aggregate lifetime data," *Technometrics,* vol. 59, no. 1, pp. 115-125, 2017.

[4] L. J. Bain and M. Engelhardt, "A Two-Moment Chi-Square Approximation for the Statistic Log," *Journal of the American Statistical Association,* vol. 70, no. 352, pp. 948-950, 1975.

[5] M. Engelhardt and L. J. Bain, "Simplified statistical procedures for the Weibull or extreme-value distribution," *Technometrics,* vol. 19, no. 3, pp. 323-331, 1977.

[6] D. K. Bhaumik, K. Kapur and R. D. Gibbons, "Testing parameters of a gamma distribution for small samples," *Technometrics,* vol. 51, no. 3, pp. 326-334, 2009.

[7] E. A. Pohl and D. L. Dietrich, "Optimal Stress Screening Strategies for Multi-Component Systems Under Warranty: The Case of Phase-Type Lifetimes," *Annals of Operations Research Special Volume on Reliability and Maintainability in Production Control,* vol. 91, no. 0, pp. 137-161, 1999.

[8] S. Asmussen, O. Nerman and M. Olsson, "Fitting phase-type distributions via the EM algorithm," *Scandinavian Journal of Statistics,* pp. 419-441, 1996.

[9] H. Liao and H. Guo, "A generic method for modeling accelerated life testing data," in *Reliability and Maintainability Symposium (RAMS)*, Orlando, FL, 2013.

[10] H. Liao and S. Karimi, "Comparison study on general methods for modeling lifetime data with covariates," in *Prognostics and System Health Management Conference (PHM)*, Harbin, 2017.

[11] J. Riascos-Ochoa, M. Sanchez-Silva and R. Akhavan-Tabatabaei, "Reliability analysis of shock-based deterioration using phase-type distributions," *Probabilistic Engineering Mechanics,* vol. 38, pp. 88-101, 2014.

[12] J. P. Kharoufeh, C. J. Solo and M. Y. Ulukus, "Semi-Markov models for degradation-based reliability," *IIE Transactions,* vol. 42, no. 8, pp. 599-612, 2010.

[13] M. Bladt, L. J. R. Esparza and B. F. Nielsen, "Fisher information and statistical inference for phase-type distributions," *Journal of Applied Probability,* vol. 48, no. A, pp. 277-293, 2011.

BIOGRAPHIES

Samira Karimi, PhD Student
Department of Industrial Engineering
University of Arkansas
Fayetteville, AR, 72701, USA

Email: sakarimi@email.uark.edu

Samira is a Ph.D. student in the Industrial Engineering Department at the University of Arkansas, Fayetteville, Arkansas. She received her B.Sc. (2014) and M.Sc. (2016) in Industrial Engineering from Sharif University of Technology, Tehran, Iran. Her research interests include reliability modeling and statistical analysis. She received 2017 PHM-Harbin Best Paper Award. She is a student member of IISE and INFORMS.

Haitao Liao, PhD
Department of Industrial Engineering
University of Arkansas
Fayetteville, AR, 72701, USA

Email: liao@uark.edu

Dr. Haitao Liao is a Professor and John and Mary Lib White Systems Integration Chair in the Industrial Engineering Department at the University of Arkansas, Fayetteville, Arkansas. He received his Ph.D. degree in 2004 from the Department of Industrial and Systems Engineering at Rutgers University. His research interests focus on modeling of accelerated testing, probabilistic risk assessment, maintenance models and optimization, spare part inventory control, and prognostics. He serves as an Area Editor for the IISE Transactions on Quality and Reliability Engineering. He received a National Science Foundation CAREER Award in 2010, IISE William A. J. Golomski Award in 2010, 2013 and 2018, 2015 SRE Stan Ofsthun Best Paper Award, and 2017 Alan O. Plait Award for Tutorial Excellence. He is a member of IISE, INFORMS, and SRE.

Edward Pohl, PhD
Department of Industrial Engineering
University of Arkansas
Fayetteville, AR, 72701, USA

Email: epohl@uark.edu

Dr. Ed Pohl is a Professor and Head of the Department of Industrial Engineering at the University of Arkansas. He is the current holder of the 21st Century Professorship. Ed spent twenty years in the United States Air Force where he served in a variety of engineering, analysis and academic positions during his career. Ed received his Ph.D. in systems and industrial engineering from the University of Arizona, an M.S. in reliability engineering from the University of Arizona, an M.S. in systems engineering from AFIT, an M.S. in engineering management from the University of Dayton, and a B.S.E.E. from Boston University. His primary research interests are in risk, reliability, engineering optimization, healthcare and supply chain risk analysis, decision making, and quality. Ed is a Fellow of IISE, a Fellow of the Society of Reliability Engineers (SRE), a senior member of ASQ, and a senior member of IEEE. Ed serves as an Associate Editor for the Journal of Military Operations Research, the Journal of Risk and Reliability, and the IEEE Transactions on Reliability. He is a two-time winner of the Alan Plait award for Outstanding Tutorial at RAMS and the William A. J. Golomski Award.