Likelihood Inference Based on Left Truncated and Right Censored Data From a Gamma Distribution

Narayanaswamy Balakrishnan and Debanjan Mitra

***Abstract*—The gamma distribution is used as a lifetime distribution widely in reliability analysis. Lifetime data are often left truncated, and right censored. The EM algorithm is developed here for the estimation of the scale and shape parameters of the gamma distribution based on left truncated and right censored data. The Newton-Raphson method is also used for the same purpose, and then these two methods of estimation are compared through an extensive Monte Carlo simulation study. The asymptotic variance-covariance matrix of the MLEs under the EM framework is obtained by using the missing information principle (Louis, 1982). Then, the asymptotic confidence intervals for the parameters are constructed. The confidence intervals based on the EM algorithm and the Newton-Raphson method are then compared empirically in terms of coverage probabilities. Finally, all the methods of inference discussed here are illustrated through a numerical example.**

***Index Terms*—Asymptotic variance-covariance matrix, coverage probability, expectation maximization algorithm, gamma distribution, information matrix, left truncation, lifetime data, maximum likelihood estimates, missing information principle, Monte Carlo simulation, right censoring.**

ACRONYMS

CDF	cumulative distribution function
EM	expectation maximization
MLE	maximum likelihood estimate
MSE	mean square error
NR	Newton-Raphson
PDF	probability density function
SF	survival function

NOTATION

X	gamma random variable
F_X, f_x	CDF, and PDF of X, respectively
$\Gamma(p)$	complete gamma function
$\gamma(p,x)$	lower incomplete gamma function, defined by $\int_0^x u^{p-1}e^{-u}du$
$\Gamma(p,x)$	upper incomplete gamma function, defined by $\int_x^\infty u^{p-1}e^{-u}du$
$\Psi(\cdot)$	digamma function
$\Psi'(\cdot)$	trigamma function
Y	observed lifetime variable
C	censoring time variable
δ_i	censoring indicator
ν_i	truncation indicator
τ_i^L	left-truncation time
S_1, S_2	index sets corresponding to the units which are not left truncated, and left truncated, respectively
$\mathrm{G}(\kappa,\theta)$	gamma distribution with shape parameter κ and scale parameter θ
$\boldsymbol{t}$	complete lifetime data vector
$\boldsymbol{\lambda}$	parameter vector (κ,θ)
$f_{T\|Y=y}$	conditional PDF of T, given $Y=y$
$f_{C\|Y}$	conditional PDF of C, given Y
$S^*(\cdot)$	SF of left truncated variable

Manuscript received May 25, 2012; revised November 24, 2012; accepted March 18, 2013. Date of publication July 19, 2013; date of current version August 28, 2013. Associate Editor: R. H. Yeh.

N. Balakrishnan is with the Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1 Canada, and also with the Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia.

D. Mitra is with the Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India.

Digital Object Identifier 10.1109/TR.2013.2273039

I. INTRODUCTION

THE gamma distribution has been used widely to model lifetime data in reliability and survival analyses; see, for example, [7], [9], [12], [13], [15], [19], and the references therein. Lifetime data are often truncated and censored. For detailed discussion on truncation and censoring, and inferential methods based on such samples, interested readers may refer to [1], [8], and [17]. In this paper, we discuss the fitting of a gamma distribution to left truncated and right censored lifetime data.

Hong, Meeker, and McCalley [11] carried out an analysis of left truncated and right censored lifetime data arising from power transformers from an energy company in the United States. They fitted a Weibull distribution to the data through a direct maximzation approach for obtaining the maximum likelihood estimates (MLEs). The expectation maximization (EM) algorithm is a very useful tool for analyzing incomplete statistical data [16]. Balakrishnan and Mitra [2], [3], and [4] have discussed in detail the EM algorithm for fitting the log-normal and Weibull distributions to left truncated and right censored data. In this paper, the necessary steps of the EM algorithm for fitting a gamma distribution to left truncated and right censored data are developed. The MLEs are also obtained by the Newton-Raphson (NR) method, and then the two methods are compared by means of an extensive Monte Carlo simulation study.

The asymptotic variance-covariance matrix of the MLEs within the EM framework are obtained by using the missing information principle [14]. Then, the asymptotic confidence intervals of the parameters are constructed, and the corresponding asymptotic confidence intervals based on the NR method are also constructed. Then, these confidence intervals are compared in terms of coverage probabilities through a simulation study.

The rest of this paper is organized as follows. In Section II, we describe the form of the data, and the likelihood function. In Section III, the EM algorithm and the NR method for estimating the model parameters are described, and the derivation of the asymptotic variance-covariance matrix and the asymptotic confidence intervals are presented. An application of this estimation method to prediction purposes is discussed in Section IV. The simulation results for the point and interval estimation of model parameters are presented in Section V. In Section VI, a numerical illustration is given. Finally, in Section VII, some concluding comments are made.

II. Form of Data and the Likelihood

In this paper, the performance of the methods of inference are studied through Monte Carlo simulation studies. For simulation, we mimic the dataset of Hong *et al.* [11], in the framework of lifetimes of power transformers.

Left truncation in the data arises when the lifetimes of units can be observed only when they exceed a threshold. In the study of Hong *et al.* [11], detailed record keeping of the lifetimes of power transformers started only in 1980, and so the failure of a transformer was observed only if it failed after 1980, thus making the data left truncated. Right censoring in the data occurs when the lifetimes of units are not followed until their complete failure, and that they are known only to exceed the censoring time. Hong *et al.* [11] observed the lifetimes of transformers until 2008, and this resulted in right censoring of the data. Here, we have followed the same setup as in their study.

Let X be a gamma random variable with scale parameter θ, and shape parameter κ. Then, the probability density function (PDF) of X is given by

$$f_X(x) = \frac{x^{\kappa-1}\exp(-x/\theta)}{\theta^{\kappa}\Gamma(\kappa)}, \quad x>0, \theta>0, \kappa>0.$$

The corresponding cumulative distribution function (CDF) of X is

$$F_X(x) = \frac{\gamma(\kappa, x/\theta)}{\Gamma(\kappa)}, \quad x>0, \theta>0, \kappa>0.$$

It is of interest to mention here that the above gamma distribution has been utilised to propose a gamma process for the modelling of the degradation of products; see for example [20] and [21] for some recent work in this direction.

δ_i is 0 if the i-th unit is censored, and 1 if it is not. Similarly, let ν_i be 0 if the i-th unit is truncated, and 1 if it is not.

The likelihood function for the left truncated and right censored data is then given by

$$L(\kappa,\theta) = \prod_{i\in S_1}\left\{f_Y(y_i)\right\}^{\delta_i}\left\{1-F_Y(y_i)\right\}^{1-\delta_i} \times \prod_{i\in S_2}\left\{\frac{f_Y(y_i)}{1-F_Y\left(\tau_i^L\right)}\right\}^{\delta_i}\left\{\frac{1-F_Y(y_i)}{1-F_Y\left(\tau_i^L\right)}\right\}^{1-\delta_i}.$$

When the lifetime distribution is gamma with scale parameter θ and shape parameter κ, upon using the truncation indicator ν_i, the likelihood function can be simplified to the form

$$\begin{aligned}\log L(\kappa,\theta) = &\sum_{i=1}^{n}\Big[\delta_i\left\{(\kappa-1)\log y_i - \frac{y_i}{\theta} - \kappa\log\theta\right\} \\ &+ (1-\delta_i)\log\Gamma\left(\kappa, \tau_i^L/\theta\right)\Big] \\ &- \sum_{i=1}^{n}\left[\nu_i\log\Gamma(\kappa) + (1-\nu_i)\log\Gamma\left(\kappa,\tau_i^L/\theta\right)\right].\end{aligned}$$

III. Methods of Estimation

A. The EM Algorithm

The EM algorithm is widely used to analyze incomplete data [16]. This algorithm involves two steps: the expectation step, and the maximization step. First, in the Expectation step (E-step), the conditional expectation of the complete data loglikelihood is obtained, given the observed data, and the current value of the parameter. Next, in the Maximization step (M-step), the expected loglikelihood is maximized with respect to the parameters. These two steps are then repeated till convergence to the required accuracy is achieved. The EM algorithm has many desirable properties, and provides additional insight in the maximum likelihood estimation of model parameters.

When $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$ denotes the lifetime data, and $\boldsymbol{\lambda} = (\kappa, \theta)$ denotes the parameter vector, with all other notation remaining the same as before, had there been no censoring, the complete data likelihood function would be

$$L_c(\boldsymbol{t};\boldsymbol{\lambda}) = \prod_{i\in S_1}\left\{\frac{t_i^{\kappa-1}e^{-t_i/\theta}}{\theta^{\kappa}\Gamma(\kappa)}\right\} \times \prod_{i\in S_2}\left\{\frac{t_i^{\kappa-1}e^{-t_i/\theta}}{\theta^{\kappa}\Gamma(\kappa,\tau_i^L/\theta)}\right\}.$$

The loglikelihood function, using the truncation indicator ν_i, then becomes

$$\begin{aligned}\log L_c(\boldsymbol{t};\boldsymbol{\lambda}) = &\sum_{i=1}^{n}\left[(\kappa-1)\log t_i - \frac{t_i}{\theta} - \kappa\log\theta\right] \\ &- \sum_{i=1}^{n}\left[\nu_i\log\Gamma(\kappa) + (1-\nu_i)\log\Gamma\left(\kappa,\tau_i^L/\theta\right)\right].\end{aligned}$$

The E-step: Let $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)'$ denote the observed data vector, where $y_i = \min(t_i, c_i)$, as the data are right censored; $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_n)'$ denotes the vector of censoring indicators. Let the current value of the parameter vector, at the r-th step, be $\boldsymbol{\lambda}^{(r)}$. Our objective is to obtain

$$Q\left(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(r)}\right) = E_{\boldsymbol{\lambda}^{(r)}}\left[\log L_c(\boldsymbol{t};\boldsymbol{\lambda})|\boldsymbol{y},\boldsymbol{\delta}\right]. \tag{1}$$

With lifetime variables being denoted by $T_i, i = 1, 2, \ldots, n$, it is clear that the required expectations are $E_{1i}^{(r)} = E_{\boldsymbol{\lambda}^{(r)}}[\log T_i|T_i > y_i]$, and $E_{2i}^{(r)} = E_{\boldsymbol{\lambda}^{(r)}}[T_i|T_i > y_i]$. These expectations can be obtained from the conditional distribution of T_i, given $T_i > y_i$. This conditional density is given by

$$f_{T_i|Y_i=y_i}(t_i) = \frac{t_i^{\kappa-1}\exp(-t_i/\theta)}{\theta^\kappa \Gamma(\kappa, y_i/\theta)}, \quad t_i > y_i, \theta > 0, \kappa > 0. \tag{2}$$

Based on this conditional density, we easily find

$$E_{\boldsymbol{\lambda}^{(r)}}[T_i|T_i > y_i] = \frac{\theta^{(r)}\Gamma\left(\kappa^{(r)}+1, y_i/\theta^{(r)}\right)}{\Gamma\left(\kappa^{(r)}, y_i/\theta^{(r)}\right)}.$$

For the other expectation, it can be seen that

$$E_{\boldsymbol{\lambda}^{(r)}}[\log T_i|T_i > y_i] = \log\theta^{(r)} + \frac{1}{\Gamma\left(\kappa^{(r)}, y_i/\theta^{(r)}\right)} \times \int_{\frac{y_i}{\theta^{(r)}}}^{\infty} u_i^{\kappa^{(r)}-1}\log u_i e^{-u_i}\,du_i. \tag{3}$$

Following [10], we have

$$\int_{\frac{y_i}{\theta^{(r)}}}^{\infty} u_i^{\kappa^{(r)}-1}\log u_i e^{-u_i}\,du_i = \left[\frac{d}{da}\Gamma\left(a, y_i/\theta^{(r)}\right)\right]_{a=\kappa^{(r)}}. \tag{4}$$

Upon using (4) in (3), we get

$$\begin{aligned} &E_{\boldsymbol{\lambda}^{(r)}}[\log T_i|T_i>y_i] \\ &= \log\theta^{(r)} + \frac{1}{\Gamma\left(\kappa^{(r)}, y_i/\theta^{(r)}\right)} \\ &\times\left[\Psi\left(\kappa^{(r)}\right)\Gamma\left(\kappa^{(r)}\right)\left\{1-e^{-y_i/\theta^{(r)}}\sum_{p=0}^{\infty}\frac{\left(\frac{y_i}{\theta^{(r)}}\right)^{\kappa^{(r)}+p}}{\Gamma\left(\kappa^{(r)}+p+1\right)}\right\}\right. \\ &\quad -\Gamma\left(\kappa^{(r)}\right)e^{-y_i/\theta^{(r)}}\log\left(\frac{y_i}{\theta^{(r)}}\right)\sum_{p=0}^{\infty}\frac{\left(\frac{y_i}{\theta^{(r)}}\right)^{\kappa^{(r)}+p}}{\Gamma\left(\kappa^{(r)}+p+1\right)} \\ &\quad \left.+\Gamma\left(\kappa^{(r)}\right)e^{-y_i/\theta^{(r)}}\sum_{p=0}^{\infty}\frac{\left(\frac{y_i}{\theta^{(r)}}\right)^{\kappa^{(r)}+p}\Psi\left(\kappa^{(r)}+p+1\right)}{\Gamma\left(\kappa^{(r)}+p+1\right)}\right]. \end{aligned}$$

Substituting these two expectations in (1), we obtain

$$\begin{aligned} Q\left(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(r)}\right) = &\left\{\sum_{i:\delta_i=1}(\kappa-1)\log t_i + \sum_{i:\delta_i=0}(\kappa-1)E_{1i}^{(r)}\right\} \\ &-\left\{\sum_{i:\delta_i=1}\frac{t_i}{\theta}+\sum_{i:\delta_i=0}\frac{E_{2i}^{(r)}}{\theta}\right\} - n\kappa\log\theta \\ &-\sum_{i=1}^{n}\left\{\nu_i\log\Gamma(\kappa)+(1-\nu_i)\log\Gamma\left(\kappa, \tau_i^L/\theta\right)\right\}. \end{aligned}$$

The M-step: The objective now is to maximize $Q(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(r)})$ with respect to $\boldsymbol{\lambda}$ over the parametric space Λ to determine

$$\boldsymbol{\lambda}^{(r+1)} = \underset{\boldsymbol{\lambda}\in\Lambda}{\arg\max}\, Q\left(\boldsymbol{\lambda}, \boldsymbol{\lambda}^{(r)}\right).$$

The first-order derivatives of the Q-function with respect to θ, and κ are, repectively,

$$\frac{\partial Q}{\partial\theta} = \sum_{i:\delta_i=1}\frac{t_i}{\theta^2} + \sum_{i:\delta_i=0}\frac{E_{2i}^{(r)}}{\theta^2} - \frac{n\kappa}{\theta} - \sum_{i=1}^{n}(1-\nu_i)\frac{\frac{\partial}{\partial\theta}\Gamma\left(\kappa, \tau_i^L/\theta\right)}{\Gamma\left(\kappa, \tau_i^L/\theta\right)}, \tag{5}$$

$$\frac{\partial Q}{\partial\kappa} = \sum_{i:\delta_i=1}\log t_i + \sum_{i:\delta_i=0}E_{1i}^{(r)} - n\log\theta - \sum_{i=1}^{n}\left\{\nu_i\Psi(\kappa)+(1-\nu_i)\frac{\frac{\partial}{\partial\kappa}\Gamma\left(\kappa, \tau_i^L/\theta\right)}{\Gamma\left(\kappa, \tau_i^L/\theta\right)}\right\}. \tag{6}$$

Now, the upper incomplete gamma function can be easily differentiated with respect to θ using Leibnitz's rule, and we then obtain from (5) that

$$\theta = \frac{1}{n\kappa}\left[\sum_{i:\delta_i=1}t_i + \sum_{i:\delta_i=0}E_{2i}^{(r)} - \sum_{i=1}^{n}(1-\nu_i)\frac{\left(\tau_i^L\right)^\kappa e^{-\tau_i^L/\theta}}{\theta^{\kappa-1}\Gamma\left(\kappa, \tau_i^L/\theta\right)}\right]. \tag{7}$$

The right hand side of (7) can be evaluated at the current parameter value $\boldsymbol{\lambda}^{(r)}$ to obtain the updated parameter estimate $\theta^{(r+1)}$.

Upon expanding the incomplete gamma function as an infinite series, then differentiating and simplifying the expression, (6) can be expressed as

$$\begin{aligned} \frac{\partial Q}{\partial\kappa} = &\sum_{i:\delta_i=1}\log t_i + \sum_{i:\delta_i=0}E_{1i}^{(r)} - n\log\theta - n\Psi(\kappa) - \sum_{i=1}^{n}(1-\nu_i) \\ &\times\left[\log\left(\frac{\tau_i^L}{\theta}\right) - \frac{\log\left(\frac{\tau_i^L}{\theta}\right)}{\left\{1-e^{-\tau_i^L/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{\tau_i^L}{\theta}\right)^{\kappa+p}}{\Gamma(\kappa+p+1)}\right\}}\right. \\ &\quad\left. + e^{-\tau_i^L/\theta}\frac{\sum_{p=0}^{\infty}\frac{\left(\frac{\tau_i^L}{\theta}\right)^{\kappa+p}\Psi(\kappa+p+1)}{\Gamma(\kappa+p+1)}}{\left\{1-e^{-\tau_i^L/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{\tau_i^L}{\theta}\right)^{\kappa+p}}{\Gamma(\kappa+p+1)}\right\}}\right]. \end{aligned}$$

Equating $(\partial Q)/(\partial\kappa)$ to zero, the equation can be solved numerically for κ to obtain the current estimate $\kappa^{(r+1)}$ by using, $\theta^{(r+1)}$ for θ. The E-step and M-step are then repeated till convergence is achieved to the desired level of accuracy.

B. Asymptotic Variances and Covariance of the MLEs

For obtaining the asymptotic variance-covariance matrix of the MLEs under the EM framework, we employ the missing information principle (see [14]), which is given by

$$\text{Observed inform} = \text{Complete inform} - \text{Missing inform}.$$

Then, the inverse of the observed information matrix evaluated at the estimated parameter values gives the estimated asymptotic variance-covariance matrix of the MLEs.

Let $I_T(\boldsymbol{\lambda})$, $I_Y(\boldsymbol{\lambda})$, and $I_{C|Y}(\boldsymbol{\lambda})$ denote the complete information matrix, observed information matrix, and the missing information matrix, respectively. The complete information matrix is obtained as

$$I_T(\boldsymbol{\lambda}) = -E\left[\frac{\partial^2}{\partial \boldsymbol{\lambda}^2}\log L_c(\boldsymbol{t};\boldsymbol{\lambda})\right]. \tag{8}$$

The Fisher information matrix in the i-th observation that is censored is given by

$$I_{C|Y}^{(i)}(\boldsymbol{\lambda}) = -E\left[\frac{\partial^2}{\partial \boldsymbol{\lambda}^2}\log f_{C_i}(c_i|C_i > y_i, \boldsymbol{\lambda})\right], \tag{9}$$

Using (9), the expected missing information matrix can then be obtained as

$$I_{C|Y}(\boldsymbol{\lambda}) = \sum_{i:\delta_i=0} I_{C|Y}^{(i)}(\boldsymbol{\lambda}). \tag{10}$$

Now, by employing the missing information principle, the observed information matrix can be obtained as

$$I_Y(\boldsymbol{\lambda}) = I_T(\boldsymbol{\lambda}) - I_{C|Y}(\boldsymbol{\lambda}). \tag{11}$$

Finally, the inverse of $I_Y(\hat{\boldsymbol{\lambda}})$ yields the estimated asymptotic variance-covariance matrix of the MLEs.

The elements of the complete information matrix are given by

$$\begin{aligned}
&-E\left[\frac{\partial^2}{\partial\theta^2}\log L_c(\boldsymbol{t};\boldsymbol{\lambda})\right]\\
&=\frac{n\kappa}{\theta^2}+\sum_{i=1}^{n}(1-\nu_i)\times\left[\frac{{\tau_i^L}^{\kappa}e^{-\tau_i^L/\theta}}{\theta^{\kappa+2}\Gamma\left(\kappa,\tau_i^L/\theta\right)}\left\{\frac{\tau_i^L}{\theta}-\kappa+1\right\}\right.\\
&\qquad\left.-\left\{\frac{\left(\frac{{\tau_i^L}^{\kappa}}{\theta^{\kappa+1}}\right)e^{-\tau_i^L/\theta}}{\Gamma\left(\kappa,\tau_i^L/\theta\right)}\right\}^2\right],\\
&-E\left[\frac{\partial^2}{\partial\kappa^2}\log L_c(\boldsymbol{t};\boldsymbol{\lambda})\right]\\
&=\sum_{i=1}^{n}\left[\nu_i\Psi'(\kappa)+(1-\nu_i)\right.\\
&\qquad\left.\times\left\{\frac{\frac{\partial^2}{\partial\kappa^2}\Gamma\left(\kappa,\tau_i^L/\theta\right)}{\Gamma\left(\kappa,\tau_i^L/\theta\right)}-\left\{\frac{\frac{\partial}{\partial\kappa}\Gamma\left(\kappa,\tau_i^L/\theta\right)}{\Gamma\left(\kappa,\tau_i^L/\theta\right)}\right\}^2\right\}\right],\\
&-E\left[\frac{\partial^2}{\partial\kappa\partial\theta}\log L_c(\boldsymbol{t};\boldsymbol{\lambda})\right]\\
&=\frac{n}{\theta}+\sum_{i=1}^{n}(1-\nu_i)\left(\frac{{\tau_i^L}^{\kappa}}{\theta^{\kappa+1}}\right)\\
&\qquad\times e^{-\tau_i^L/\theta}\left[\frac{\log\left(\frac{\tau_i^L}{\theta}\right)}{\Gamma\left(\kappa,\tau_i^L/\theta\right)}-\frac{\frac{\partial}{\partial\kappa}\Gamma\left(\kappa,\tau_i^L/\theta\right)}{\Gamma\left(\kappa,\tau_i^L/\theta\right)^2}\right].
\end{aligned}$$

The conditional PDF, required for the derivation of the missing information matrix, can be derived to be (see [18])

$$f_{C_i|Y_i}(c_i|C_i>y_i,\theta,\kappa)=\frac{c_i^{\kappa-1}e^{-c_i/\theta}}{\theta^{\kappa}\Gamma(\kappa,y_i/\theta)}, c_i>y_i, \theta>0, \kappa>0.$$

Using this conditional PDF, we can easily derive the following expressions.

$$\begin{aligned}
&-\frac{\partial^2}{\partial\theta^2}\log f_{C_i|Y_i}\\
&=\frac{2c_i}{\theta^3}-\frac{\kappa}{\theta^2}+\frac{e^{-y_i/\theta}y_i^{\kappa}}{\Gamma(\kappa,y_i/\theta)\theta^{\kappa+2}}\\
&\quad\times\left\{\frac{y_i}{\theta}-\kappa-1\right\}-\left\{\frac{e^{-y_i/\theta}\left(\frac{y_i^{\kappa}}{\theta^{\kappa+1}}\right)}{\Gamma(\kappa,y_i/\theta)}\right\}^2, \qquad (12)\\
&-\frac{\partial^2}{\partial\kappa^2}\log f_{C_i|Y_i}\\
&=\frac{\frac{\partial^2}{\partial\kappa^2}\Gamma(\kappa,y_i/\theta)}{\Gamma(\kappa,y_i/\theta)}-\left\{\frac{\frac{\partial}{\partial\kappa}\Gamma(\kappa,y_i/\theta)}{\Gamma(\kappa,y_i/\theta)}\right\}^2, \qquad (13)\\
&-\frac{\partial^2}{\partial\kappa\partial\theta}\log f_{C_i|Y_i}\\
&=\frac{1}{\theta}+\frac{e^{-y_i/\theta}y_i^{\kappa}\log\left(\frac{y_i}{\theta}\right)}{\Gamma(\kappa,y_i/\theta)\theta^{\kappa+1}}\\
&\quad-\frac{e^{-y_i/\theta}\left(\frac{y_i^{\kappa}}{\theta^{\kappa+1}}\right)\frac{\partial}{\partial\kappa}\Gamma(\kappa,y_i/\theta)}{\Gamma(\kappa,y_i/\theta)^2}. \qquad (14)
\end{aligned}$$

Clearly, the only required expectation here is $E[C_i|C_i > y_i]$, which can be easily derived from the above conditional PDF to be

$$E[C_i|C_i > y_i] = \frac{\theta\Gamma(\kappa+1, y_i/\theta)}{\Gamma(\kappa, y_i/\theta)}. \tag{15}$$

Using (9), (10), (12)–(15), the expected missing information matrix $I_{C|Y}(\boldsymbol{\lambda})$ can be easily obtained. Using that result, the observed information matrix can be obtained without any difficulty from (11). Finally, the asymptotic variance-covariance matrix of the MLEs can be obtained by inverting $I_Y(\hat{\boldsymbol{\lambda}})$.

C. Confidence Intervals

Once the MLEs and their asymptotic variance-covariance matrix are obtained, we can construct asymptotic confidence intervals for the parameters θ and κ by using the asymptotic normality of the MLEs. Here, we construct the asymptotic confidence intervals for the parameters corresponding to both the EM algorithm and the NR method, which will be clearly different. The confidence intervals are then compared in terms of coverage probabilities through an extensive Monte Carlo simulation study.

One can also construct parametric bootstrap confidence intervals for θ and κ in the following manner. First of all, based on the data of size n, the MLE $\hat{\boldsymbol{\lambda}}$ of $\boldsymbol{\lambda} = (\kappa, \theta)'$ is obtained. Then, using $\hat{\boldsymbol{\lambda}}$ as the true value of the parameter, a sample of size n in the same sampling framework with left truncation and right censoring is produced. This process is repeated for 1000 Monte Carlo simulation runs, and the MLEs are obtained for each of

TABLE I
BIAS (B) AND MSE FOR THE EM ALGORITHM AND THE NR METHOD

n = 50							n = 100						
(κ, θ)	Trunc.	Method	B($\hat{\kappa}$)	B($\hat{\theta}$)	MSE($\hat{\kappa}$)	MSE($\hat{\theta}$)	(κ, θ)	Trunc.	Method	B($\hat{\kappa}$)	B($\hat{\theta}$)	MSE($\hat{\kappa}$)	MSE($\hat{\theta}$)
(4, 8)	10%	EM	0.584	-0.160	2.727	11.277	(4, 8)	10%	EM	0.394	-0.374	1.178	4.903
		NR	0.584	-0.159	2.730	11.287			NR	0.394	-0.373	1.180	4.910
	15%	EM	0.670	-0.399	2.936	9.135		15%	EM	0.336	-0.313	1.112	4.560
		NR	0.670	-0.399	2.939	9.141			NR	0.337	-0.313	1.113	4.565
(5, 5)	10%	EM	0.429	-0.271	1.336	1.153	(5, 5)	10%	EM	0.375	-0.195	1.309	1.261
		NR	0.430	-0.272	1.338	1.153			NR	0.376	-0.196	1.311	1.262
	15%	EM	0.604	-0.242	2.838	2.223		15%	EM	0.434	-0.279	1.353	1.160
		NR	0.604	-0.242	2.841	2.227			NR	0.435	-0.280	1.355	1.161
n = 200							n = 300						
(κ, θ)	Trunc.	Method	B($\hat{\kappa}$)	B($\hat{\theta}$)	MSE($\hat{\kappa}$)	MSE($\hat{\theta}$)	(κ, θ)	Trunc.	Method	B($\hat{\kappa}$)	B($\hat{\theta}$)	MSE($\hat{\kappa}$)	MSE($\hat{\theta}$)
(4, 8)	10%	EM	0.256	-0.397	0.539	2.478	(4, 8)	10%	EM	0.217	-0.427	0.348	1.708
		NR	0.256	-0.396	0.540	2.480			NR	0.217	-0.426	0.348	1.709
	15%	EM	0.254	-0.400	0.509	2.284		15%	EM	0.195	-0.368	0.336	1.580
		NR	0.254	-0.399	0.510	2.287			NR	0.195	-0.368	0.336	1.582
(5, 5)	10%	EM	0.288	-0.229	0.662	0.653	(5, 5)	10%	EM	0.204	-0.203	0.386	0.428
		NR	0.289	-0.231	0.662	0.653			NR	0.205	-0.204	0.387	0.428
	15%	EM	0.286	-0.250	0.577	0.588		15%	EM	0.195	-0.194	0.362	0.396
		NR	0.287	-0.251	0.578	0.588			NR	0.196	-0.195	0.362	0.396

TABLE II
COVERAGE PROBABILITIES FOR THE TWO ASYMPTOTIC CONFIDENCE INTERVALS FOR θ FOR DIFFERENT NOMINAL CONFIDENCE LEVELS

n=50					n=100				
θ	Truncation	Nominal CL	Coverage Probability		θ	Truncation	Nominal CL	Coverage Probability	
			EM	NR				EM	NR
8	10%	90%	0.814*	0.814*	8	10%	90%	0.833*	0.833*
		95%	0.863*	0.863*			95%	0.872*	0.872*
	15%	90%	0.806*	0.805*		15%	90%	0.837*	0.835*
		95%	0.845*	0.844*			95%	0.872*	0.872*
5	10%	90%	0.846*	0.845*	5	10%	90%	0.830*	0.830*
		95%	0.895*	0.894*			95%	0.883*	0.883*
	15%	90%	0.837*	0.837*		15%	90%	0.836*	0.836*
		95%	0.896*	0.895*			95%	0.882*	0.882*
n=200					n=300				
θ	Truncation	Nominal CL	Coverage Probability		θ	Truncation	Nominal CL	Coverage Probability	
			EM	NR				EM	NR
8	10%	90%	0.834*	0.834*	8	10%	90%	0.815*	0.815*
		95%	0.881*	0.880*			95%	0.864*	0.864*
	15%	90%	0.828*	0.828*		15%	90%	0.841*	0.840*
		95%	0.871*	0.871*			95%	0.896*	0.895*
5	10%	90%	0.836*	0.836*	5	10%	90%	0.840*	0.839*
		95%	0.890*	0.890*			95%	0.901*	0.901*
	15%	90%	0.839*	0.838*		15%	90%	0.843*	0.843*
		95%	0.905*	0.905*			95%	0.898*	0.898*

Note: The $*$ values are significantly different from the nominal level with 95% confidence.

these samples. Then, based on these 1000 estimates, the bootstrap bias and variance for the estimates of θ and κ are obtained. In the final step, a $100(1-\alpha)\%$ parametric bootstrap confidence interval for θ is obtained as

$$\mathrm{LCL} = \hat{\theta} - b_\theta - z_{\frac{\alpha}{2}}\sqrt{v_\theta}, \quad \mathrm{UCL} = \hat{\theta} - b_\theta + z_{\frac{\alpha}{2}}\sqrt{v_\theta},$$

where b_θ, and v_θ are the bootstrap bias, and variance for θ, respectively; and z_α is the upper α-percentage point of the standard normal distribution. The confidence interval for κ can be constructed in a similar manner.

One can also construct nonparametric bootstrap confidence intervals for the model parameters. However, in the presence of both truncation and censoring, the nonparametric bootstrap is expected to show poor performance compared to the parametric

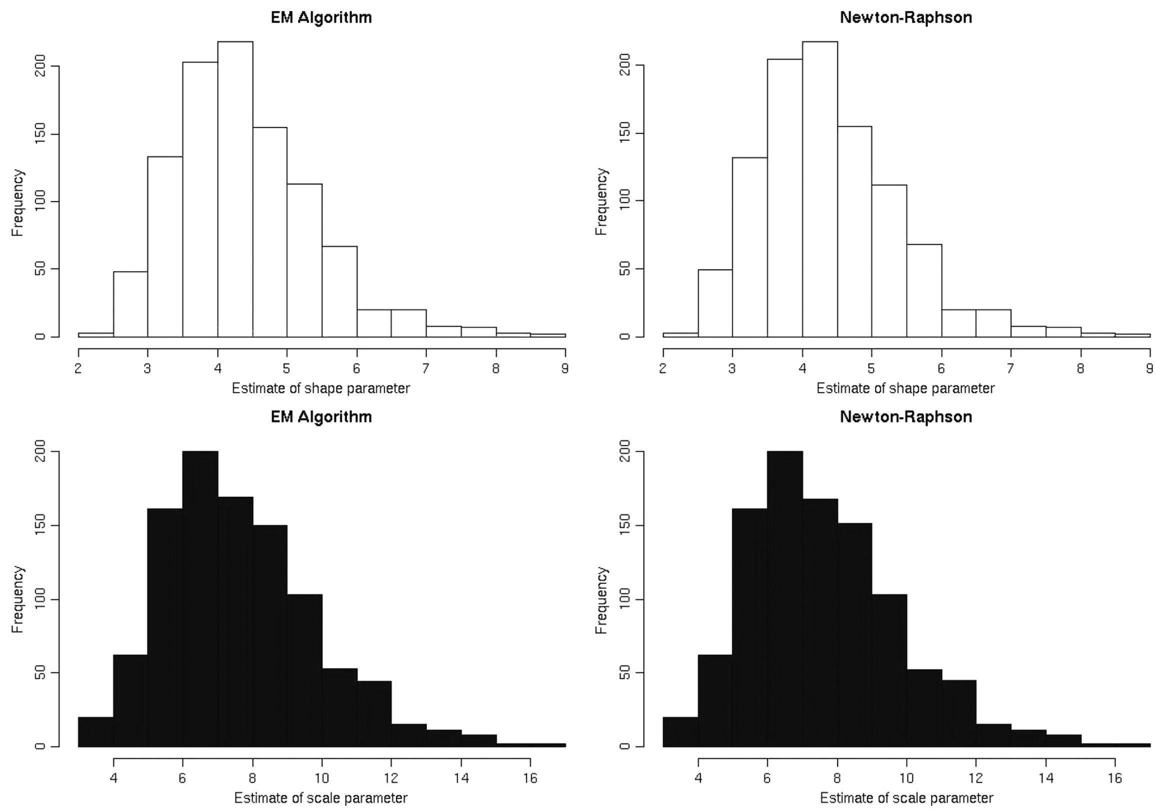

Fig. 1. Histograms for the estimates of κ (top), and θ (bottom), for the EM algorithm, and the NR method, when $n = 100$, $(\kappa, \theta) = (4, 8)$, and the amount of truncation is 10%.

bootstrap; for example, Balakrishnan *et al.* [5] observed that parametric bootstrap confidence intervals performed better than their nonparametric counterpart for censored data in the context of step-stress experiments.

D. The NR Method

The NR method is a direct approach for maximizing the likelihood function; and in this paper, we employ it for obtaining the MLEs, for comparison purposes. An R-library function, maxNR, is used for this purpose, and it is observed in the simulation study that the two methods, EM and NR, give quite close results, as can be seen in Sections V and VI.

IV. An Application to Prediction

With the estimated parameters θ and κ, one can obtain the probability of a censored unit working till a future year, given that it has worked till Y_{cen} (the right censoring point). Suppose a unit is installed in year Y_{ins}, before 1980, i.e., the unit is left truncated. Then, the left truncation time for the unit is $\tau^L = 1980 - Y_{ins}$. Then, the probability that this unit will be working till a future year Y_{fut}, given that it is right censored at Y_{cen}, will be given by

$$\begin{aligned}\pi &= P\left(T > (Y_{fut} - Y_{ins}) | T > (Y_{cen} - Y_{ins})\right) \\ &= \frac{S^*\left((Y_{fut} - Y_{ins})\right)}{S^*\left((Y_{cen} - Y_{ins})\right)},\end{aligned}$$

Clearly, the above probability reduces to

$$\pi = \frac{S\left((Y_{fut} - Y_{ins})\right)}{S\left((Y_{cen} - Y_{ins})\right)} = g(\boldsymbol{\lambda}),$$

where $S(\cdot)$ is the survival function (SF) of the untruncated lifetime variable, and $g(\cdot)$ is a function of $\boldsymbol{\lambda}$. Incidentally, this is also the probability of the same event for a unit which is not left truncated. One can obtain an estimate $\hat{\pi}$ by using the MLE $\hat{\boldsymbol{\lambda}}$ as

$$\hat{\pi} = \frac{\Gamma(\hat{\kappa}, a/\hat{\theta})}{\Gamma(\hat{\kappa}, b/\hat{\theta})} = g(\hat{\boldsymbol{\lambda}}), \tag{16}$$

where $a = (Y_{fut} - Y_{ins})$, and $b = (Y_{cen} - Y_{ins})$.

Using the delta-method, and the asymptotic variance-covariance matrix of the MLE $\hat{\boldsymbol{\lambda}}$, one can also estimate the variance of the above estimate $\hat{\pi}$. A straightforward application of the delta-method yields

$$\hat{\pi} \sim \mathrm{N}\left(\pi, \mathrm{Var}(\hat{\pi})\right),$$

where $\mathrm{Var}(\hat{\pi})$ can be estimated as

$$\begin{aligned}\widehat{\mathrm{Var}(\hat{\pi})} = \Bigg(\left(\frac{\partial g}{\partial \kappa}\right)^2 \mathrm{Var}(\hat{\kappa}) + 2\left(\frac{\partial g}{\partial \kappa}\right)\left(\frac{\partial g}{\partial \theta}\right)\mathrm{Cov}(\hat{\kappa}, \hat{\theta}) \\ + \left(\frac{\partial g}{\partial \theta}\right)^2 \mathrm{Var}(\hat{\theta})\Bigg)\Bigg|_{\boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}}.\end{aligned} \tag{17}$$

Examples of this technique are demonstrated with the illustrative example in Section VI.

V. Simulation Study

All the computational work in this paper is performed using the R software. The simulation steps are as follows. First, a truncation percentage was fixed. The set of installation years was split into two parts: (1960–1979), and (1980–1995). Unequal

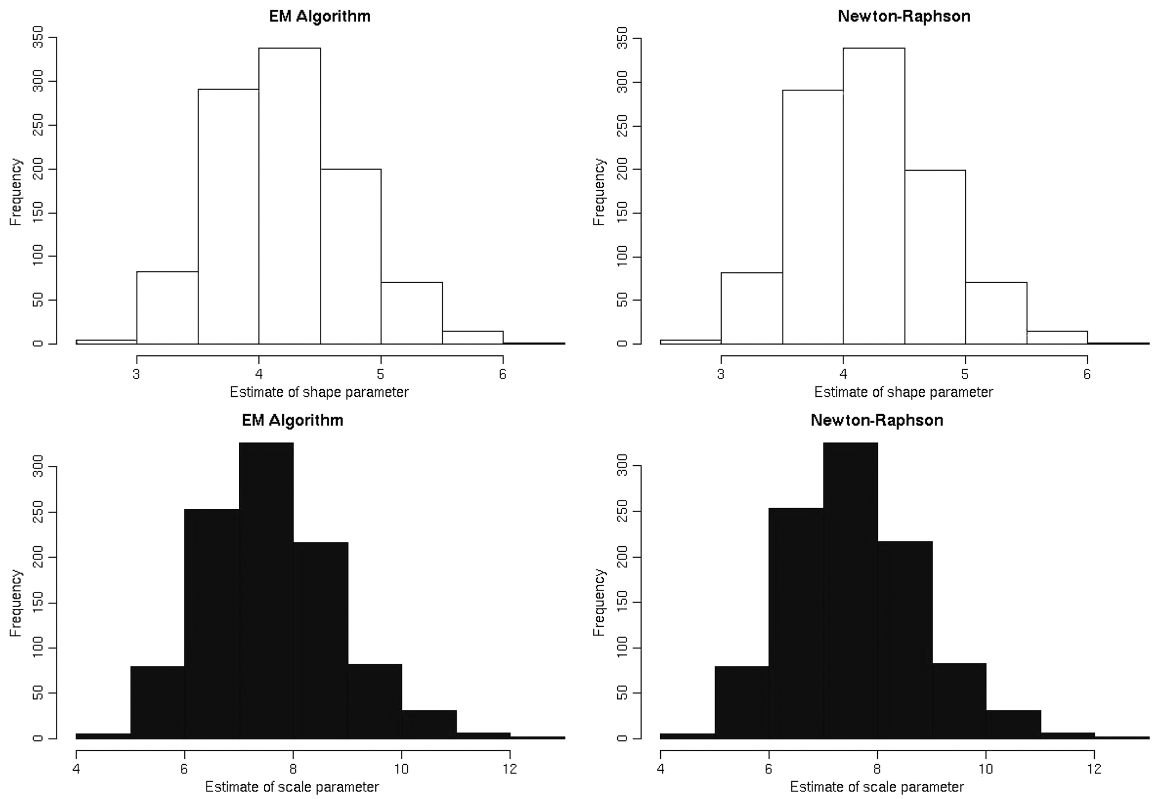

Fig. 2. Histograms for the estimates of κ (top), and θ (bottom) for the EM algorithm, and the NR method, when $n = 300$, $(\kappa, \theta) = (4, 8)$, and the amount of truncation is 10%.

probabilities were assigned to the different years as follows. For the period 1980–1995, a probability of 0.1 was attached to each of the first six years, and a probability of 0.04 was attached to each of the rest of the years of this period. For the period 1960–1979, a probability of 0.15 was attached to each of the first five years, and the remaining probability was distributed equally over the rest of the years of this period. The lifetimes of the machines, in years, are sampled from a gamma distribution with scale parameter θ, and shape parameter κ. The years of failure of the machines are obtained by adding these lifetimes with the corresponding installation years. The year of failure of a machine decides whether it is a censored observation or not. As mentioned earlier, the year of truncation has been fixed to 1980, and the year of censoring has been fixed to 2008, in our study, along the lines of [11]. As the data are right censored, the lifetime of a censored unit is taken as the minimum of the lifetime and the censoring time. Because the data are left truncated, no information on the lifetime of a machine is available if the year of failure is before 1980. Therefore, if the year of failure for a machine is obtained to be a year before 1980, that observation is discarded, and a new installation year and lifetime are simulated for that unit. This setup produced, along with the desired proportion of truncated observations, sufficiently many censored observations.

The sample sizes used in this study are 100, 200, and 300. The truncation percentages are fixed at 10, and 15. As the Gamma distribution is right skewed, if the left truncation percentage is at a higher level, then we would expect the estimates to be highly biased. The two choices of the gamma parameter vector (κ, θ) are made as (4, 8), and (5, 5). In this regard, it is useful to mention that Balakrishnan and Cohen [1] observed that the standard errors of the gamma parameter estimates based on complete data were not reliable when the shape parameter value was less than or equal to 2.5, as the corresponding information matrix becomes nearly singular. Here, for left truncated and right censored data, empirically we observed the same as well. All the simulation results are based on 1000 Monte Carlo runs.

We know that, when T follows $\mathrm{G}(\kappa, \theta)$,

$$E(T) = \theta\kappa, \quad \mathrm{Var}(T) = \kappa\theta^2.$$

From these expressions, the moment estimates of θ and κ can be easily derived. However, these estimates are crude as the data are incomplete. Using these crude moment estimates, we obtain the expected lifetimes of the censored observations. These expected lifetimes of the censored observations, along with the observed lifetimes, can be considered to form a pseudo-complete data set. Then, the moment estimates for the parameters are obtained from these pseudo-complete data, and are used as initial values. This method is used in the iterative algorithm throughout this paper. The bias, and mean square error (MSE) of the parameter estimates for the EM and NR methods are reported in Table I.

From Table I, see that the two methods are in quite close agreement, in terms of bias, and MSE of the parameter estimates.

Table II gives the coverage probabilities for θ, corresponding to the two asymptotic confidence intervals based on the EM algorithm, and the NR method. As the time required for the construction of parametric bootstrap confidence intervals is very

TABLE III
COVERAGE PROBABILITIES FOR THE TWO ASYMPTOTIC CONFIDENCE INTERVALS FOR κ FOR DIFFERENT NOMINAL CONFIDENCE LEVELS

n=50					n=100				
κ	Truncation	Nominal CL	Coverage Probability		κ	Truncation	Nominal CL	Coverage Probability	
			EM	NR				EM	NR
4	10%	90%	0.915	0.915	4	10%	90%	0.908	0.909
		95%	0.951	0.951			95%	0.951	0.951
	15%	90%	0.915	0.914		15%	90%	0.898	0.898
		95%	0.970*	0.969*			95%	0.954	0.954
5	10%	90%	0.905	0.905	5	10%	90%	0.893	0.893
		95%	0.960	0.960			95%	0.951	0.951
	15%	90%	0.934*	0.934*		15%	90%	0.890	0.890
		95%	0.962	0.962			95%	0.951	0.951
n=200					n=300				
κ	Truncation	Nominal CL	Coverage Probability		κ	Truncation	Nominal CL	Coverage Probability	
			EM	NR				EM	NR
4	10%	90%	0.894	0.893	4	10%	90%	0.879*	0.879*
		95%	0.951	0.951			95%	0.944	0.944
	15%	90%	0.886	0.885		15%	90%	0.880*	0.880*
		95%	0.95	0.95			95%	0.948	0.948
5	10%	90%	0.884	0.884	5	10%	90%	0.889	0.889
		95%	0.932*	0.932*			95%	0.939	0.939
	15%	90%	0.899	0.899		15%	90%	0.887	0.887
		95%	0.942	0.941			95%	0.942	0.942

Note: The $*$ values are significantly different from the nominal level with 95% confidence.

high, it was skipped in the simulation study. But, for illustrative purposes, the parametric bootstrap confidence intervals are constructed in the next section.

Notice from Table II that the coverage probabilities corresponding to the two methods are quite close, but they are consistently below the nominal level. The smaller coverage probabilities compared to the nominal level seem to occur due to small values of the standard error of the estimates. The small values of the standard error result in confidence intervals not wide enough, and as a result, the true parameter value does not belong to these intervals quite a number of times. From the histograms of the parameter estimates shown in Figs. 1 and 2, it is apparent that the small coverage probabilities are not entirely due to non-normality.

Analogous to Table II, Table III gives the coverage probabilities for κ. Note from Table III that the coverage probabilities corresponding to the two methods are again quite close, and they are very close to the nominal level also.

VI. ILLUSTRATIVE EXAMPLE

The sample size used for the illustration is 100, the truncation percentage is 15, and the true parameter value is $(\kappa, \theta) = (5, 5)$. The simulated data that are used for this illustration are presented in the Appendix. The initial value, calculated by the method of moments explained in the preceding section, is found to be (3.293, 9.895). The tolerance level was set at 0.0001. The EM algorithm took 114 steps to give the final estimates of (κ, θ) to be (4.961, 4.897), while the NR method took 6 steps to converge to (4.962, 4.896). However, the time taken by the two methods in terms of CPU usage is almost the same. The asymptotic confidence intervals based on the EM algorithm, the NR method, and the parametric bootstrap technique are presented in Table IV.

TABLE IV
CONFIDENCE INTERVALS OBTAINED BY DIFFERENT METHODS

Parameter	Nominal CL	EM	NR	Bootstrap
θ = 5	90%	(3.212, 6.581)	(3.211, 6.580)	(3.453, 6.748)
	95%	(2.889, 6.904)	(2.888, 6.903)	(3.138, 7.064)
κ = 5	90%	(3.460, 6.461)	(3.461, 6.462)	(2.918, 6.286)
	95%	(3.173, 6.749)	(3.173, 6.750)	(2.596, 6.609)

From Table IV, we observe that confidence intervals based on the EM algorithm and the NR method are quite close. It is also observed that the parametric bootstrap confidence intervals are shorter compared to other asymptotic confidence intervals for θ, while they are slightly wider than the other two for κ.

By employing the missing information principle, the complete information matrix for these data is obtained to be

$$\begin{pmatrix} 20.6016 & 19.2816 \\ 19.2816 & 20.0112 \end{pmatrix},$$

while the observed information is obtained to be

$$\begin{pmatrix} 17.7724 & 15.2849 \\ 15.2849 & 14.0989 \end{pmatrix}.$$

Thus, the ratio of the determinants of the observed information matrix to the complete information matrix is 0.4185, from which we could provide 58.15% as the proportion of information lost due to censoring. Alternatively, we can compute the corresponding variance-covariance matrices as

$$\begin{pmatrix} 0.4943 & -0.4763 \\ -0.4763 & 0.5089 \end{pmatrix} \text{ and } \begin{pmatrix} 0.8322 & -0.9022 \\ -0.9022 & 1.0490 \end{pmatrix}$$

TABLE V
A SIMULATED DATASET, FOR SAMPLE SIZE 100, TRUNCATION PERCENTAGE 15, AND THE TRUE PARAMETER VALUE OF (κ, θ) AS (5, 5)

Serial No.	Installation year	Truncation indicator	Truncation time	Failure year	Censoring indicator	Lifetime	Censoring time
1	1986	1	*	*	0	22	22
2	1987	1	*	2008	1	21	*
3	1980	1	*	1999	1	19	*
4	1995	1	*	*	0	13	*
5	1987	1	*	1997	1	10	*
6	1993	1	*	2006	1	13	*
7	1992	1	*	*	0	16	16
8	1993	1	*	*	0	15	15
9	1980	1	*	2006	1	26	*
10	1985	1	*	2007	1	22	*
11	1982	1	*	1995	1	13	*
12	1990	1	*	1998	1	8	*
13	1982	1	*	*	0	26	26
14	1983	1	*	2005	1	22	*
15	1984	1	*	*	0	24	24
16	1980	1	*	1998	1	18	*
17	1992	1	*	1998	1	6	*
18	1982	1	*	1996	1	14	*
19	1987	1	*	*	0	21	21
20	1982	1	*	*	0	26	26
21	1984	1	*	2007	1	23	*
22	1988	1	*	*	0	20	20
23	1980	1	*	2006	1	26	*
24	1993	1	*	*	0	15	15
25	1989	1	*	*	0	19	19
26	1987	1	*	*	0	21	21
27	1985	1	*	2002	1	17	*
28	1984	1	*	1989	1	5	*
29	1980	1	*	*	0	28	28
30	1980	1	*	*	0	28	28
31	1985	1	*	1999	1	14	*
32	1985	1	*	*	0	23	23
33	1982	1	*	*	0	26	26
34	1982	1	*	1996	1	14	*
35	1983	1	*	*	0	25	25
36	1980	1	*	2003	1	23	*
37	1980	1	*	1999	1	19	*
38	1982	1	*	*	0	26	26
39	1987	1	*	*	0	21	21
40	1988	1	*	*	0	20	20
41	1984	1	*	*	0	24	24
42	1986	1	*	*	0	22	22
43	1982	1	*	2008	1	26	*
44	1983	1	*	1999	1	16	*
45	1981	1	*	1994	1	13	*
46	1990	1	*	*	0	18	18
47	1983	1	*	2001	1	18	*
48	1983	1	*	1997	1	14	*
49	1993	1	*	*	0	15	15
50	1984	1	*	1999	1	15	*
51	1981	1	*	1999	1	18	*
51	1982	1	*	2000	1	18	*
52	1984	1	*	*	0	24	24
53	1989	1	*	2005	1	16	*
54	1986	1	*	2002	1	16	*
55	1981	1	*	*	0	27	27
56	1982	1	*	1998	1	16	*
57	1981	1	*	1990	1	9	*
58	1984	1	*	*	0	24	24
59	1982	1	*	2008	1	26	*
60	1983	1	*	*	0	25	25
61	1990	1	*	2003	1	13	*
62	1986	1	*	*	0	22	22
63	1985	1	*	*	0	23	23
64	1991	1	*	*	0	17	17
65	1984	1	*	*	0	24	24
66	1982	1	*	1991	1	9	*

TABLE V *(Continued)*

Serial No.	Installation year	Truncation indicator	Truncation time	Failure year	Censoring indicator	Lifetime	Censoring time
67	1984	1	*	2002	1	18	*
68	1984	1	*	2008	1	24	*
69	1982	1	*	*	0	26	26
70	1983	1	*	*	0	25	25
71	1983	1	*	1998	1	15	*
72	1985	1	*	*	0	23	23
73	1984	1	*	1999	1	15	*
74	1994	1	*	*	0	14	14
75	1985	1	*	*	0	23	23
76	1989	1	*	*	0	19	19
77	1982	1	*	1995	1	13	*
78	1990	1	*	2005	1	15	*
79	1980	1	*	*	0	28	28
80	1993	1	*	*	0	15	15
81	1984	1	*	2008	1	24	*
82	1982	1	*	1995	1	13	*
83	1985	1	*	*	0	23	23
84	1994	1	*	*	0	14	14
85	1981	1	*	2006	1	25	*
86	1964	0	16	1983	1	19	*
87	1963	0	17	*	0	45	45
88	1962	0	18	1991	1	29	*
89	1963	0	17	1981	1	18	*
90	1964	0	16	2000	1	36	*
91	1962	0	18	1989	1	27	*
92	1961	0	19	2002	1	41	*
93	1974	0	6	1992	1	18	*
94	1961	0	19	1985	1	24	*
95	1977	0	3	*	0	31	31
96	1967	0	13	1989	1	22	*
97	1965	0	15	1990	1	25	*
98	1964	0	16	1987	1	23	*
99	1963	0	17	1986	1	23	*
100	1963	0	17	1980	1	17	*

from which we can compute the trace-efficiency of the estimates based on censored data as compared to complete data to be 53.33%, and the determinant-efficiency to be 41.85%.

Example to the Prediction Problem: Refer to the 95-th unit in Table V, for example. For this unit, Y_{ins} is 1977, i.e., the unit is left truncated; also, it is right censored, with the censoring year being 2008. The probability that this unit will be working till 2016, by using (16), and the estimated parameters κ and θ, is obtained to be 0.415. The standard error of this probability estimate, obtained by using (17), and the estimated variance-covariance matrix of the MLEs as

$$\begin{pmatrix} 0.8322 & -0.9022 \\ -0.9022 & 1.0490 \end{pmatrix},$$

turns out to be 0.066. In fact, an approximate 95% confidence interval for this probability is (0.286, 0.544). Similarly, for the 15-th unit for which the installation year is 1984 (i.e., not left truncated and also right censored), the probability that the unit will be working till 2016 is estimated to be 0.476, with a standard error of 0.060. An approximate 95% confidence interval for this probability is then (0.358, 0.594). Note that the second unit (installed in 1984) has a higher probability of working till 2016 than the first unit (installed in 1977), as one would expect.

VII. CONCLUDING REMARKS

In this paper, the EM algorithm is developed for fitting a gamma distribution based on left truncated and right censored data. The Newton-Raphson method is also applied for the same purpose, and it is observed that both methods converge accurately to the true parameter values, and yield quite close results. Asymptotic confidence intervals for the parameters are constructed under both methods, and it is observed that the coverage probabilities for these two methods are quite close as well. While the coverage probabilities for the parameter κ are close to the nominal level, those for θ are found to be always less than the nominal level.

The EM algorithm for the parameter estimation of lognormal and Weibull distributions based on left truncated and right censored data have been developed earlier by Balakrishnan and Mitra ([2], [3], and [4]). The lognormal, exponential, Weibull, and gamma distributions all belong to the generalized gamma family of distributions. It will, therefore, be of interest to consider the parsimonious family of generalized gamma distribution as in [6], develop the EM algorithm for the estimates of the parameters of the distribution based on left truncated and right censored data, and apply it to do model discrimination within this family by using likelihood-ratio and information-based methods. We are currently working on this problem, and hope to report the findings in a future paper.

APPENDIX

The first, and second derivatives of the upper incomplete gamma function $\Gamma(\kappa, y/\theta)$ with respect to θ can be easily obtained by using Leibnitz's rule. For obtaining the derivatives with respect to κ, we use the following infinite series expansion of incomplete gamma function.

$$\Gamma(\kappa, y/\theta) = \Gamma(\kappa)\left[1 - e^{-y/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}}{\Gamma(\kappa+p+1)}\right].$$

The first, and second derivatives with respect to κ are obtained to be

$$\begin{aligned}
&\frac{\partial}{\partial\kappa}\Gamma(\kappa, y/\theta)\\
&\quad = \left\{\Psi(\kappa)\Gamma(\kappa) + \Gamma(\kappa)\log\left(\frac{y}{\theta}\right)\right\}\\
&\qquad \times\left[1 - e^{-y/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}}{\Gamma(\kappa+p+1)}\right] - \Gamma(\kappa)\log\left(\frac{y}{\theta}\right)\\
&\qquad + \Gamma(\kappa)e^{-y/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}\Psi(\kappa+p+1)}{\Gamma(\kappa+p+1)},\\
&\frac{\partial^2}{\partial\kappa^2}\Gamma(\kappa, y/\theta)\\
&\quad = \left\{\Psi'(\kappa)\Gamma(\kappa) + \Psi^2(\kappa)\Gamma(\kappa)\right\}\\
&\qquad \times\left[1 - e^{-y/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}}{\Gamma(\kappa+p+1)}\right]\\
&\qquad - \left\{2\Psi(\kappa)\Gamma(\kappa)\log\left(\frac{y}{\theta}\right) + \Gamma(\kappa)\left\{\log\left(\frac{y}{\theta}\right)\right\}^2\right\}\\
&\qquad \times e^{-y/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}}{\Gamma(\kappa+p+1)}\\
&\qquad + 2\left\{\Gamma(\kappa)\Psi(\kappa) + \Gamma(\kappa)\log\left(\frac{y}{\theta}\right)\right\}\\
&\qquad \times e^{-y/\theta}\sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}\Psi(\kappa+p+1)}{\Gamma(\kappa+p+1)}\\
&\qquad + \Gamma(\kappa)e^{-y/\theta}\left[\sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}\Psi'(\kappa+p+1)}{\Gamma(\kappa+p+1)}\right.\\
&\qquad\qquad\left. - \sum_{p=0}^{\infty}\frac{\left(\frac{y}{\theta}\right)^{\kappa+p}\Psi^2(\kappa+p+1)}{\Gamma(\kappa+p+1)}\right].
\end{aligned}$$

Table V presents the left truncated and right censored data that are analyzed in Section VI.

REFERENCES

[1] N. Balakrishnan and A. C. Cohen, *Order Statistics and Inference: Estimation Methods*. Boston, MA, USA: Academic Press, 1991.

[2] N. Balakrishnan and D. Mitra, "Likelihood inference for lognormal data with left truncation and right censoring with an illustration," *J. Statist. Plan. Inference*, vol. 141, pp. 3536–3553, 2011.

[3] N. Balakrishnan and D. Mitra, "Left truncated and right censored Weibull data and likelihood inference with an illustration," *Comput. Statist. Data Ana.*, vol. 56, pp. 4011–4025, 2012, (to appear).

[4] N. Balakrishnan and D. Mitra, "Some further issues concerning likelihood inference for left truncated and right censored lognormal," *Commun. Statist.-Simulation Comput.*, 2013, to be published.

[5] N. Balakrishnan, D. Kundu, H. K. T. Ng, and N. Kannan, "Point and interval estimation for a simple step-stress model with Type-II censoring," *J. Quality Technol.*, vol. 39, pp. 35–47, 2007.

[6] N. Balakrishnan and Y. Peng, "Generalized gamma frailty model," *Statist. Med.*, vol. 25, pp. 2797–2816, 2006.

[7] K. O. Bowman and L. R. Shenton, *Properties of Estimators for the Gamma Distribution*. New York: Marcel Dekker, 1988.

[8] A. C. Cohen, *Truncated and Censored Samples*. New York, NY, USA: Marcel Dekker, 1991.

[9] A. J. Fernandez, "Optimal reliability demonstration test plans for k-out-of-n systems of gamma distributed components," *IEEE Trans. Rel.*, vol. 60, no. 4, pp. 833–844, Dec. 2011.

[10] K. O. Geddes, M. L. Glasser, R. A. Moore, and T. C. Scott, "Evaluation of classes of definite integrals involving elementary functions via differentiation of special functions," *Applicable Algebra Eng., Commun. Comput.*, vol. 1, pp. 149–165, 1990.

[11] Y. Hong, W. Q. Meeker, and J. D. McCalley, "Prediction of remaining life of power transformers based on left truncated and right censored lifetime data," *Annal. Appl. Statist.*, vol. 3, pp. 857–879, 2009.

[12] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions—Vol. 1*. Hoboken, NJ, USA: Wiley, 1994.

[13] K. Krishnamoorthy, T. Mathew, and S. Mukherjee, "Normal-based methods for a gamma distribution: Prediction and tolerance intervals and stress-strength reliability," *Technometrics*, vol. 50, pp. 69–78, 2008.

[14] T. A. Louis, "Finding the observed information matrix when using the EM algorithm," *J. Roy. Statist. Soc., Series B*, vol. 44, pp. 226–233, 1982.

[15] W. Lu and T. R. Tsai, "Interval censored sampling plans for the gamma lifetime model," *Eur. J. Operational Res.*, vol. 192, pp. 116–124, 2009.

[16] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. Hoboken, NJ, USA: Wiley, 2008.

[17] W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*. Hoboken, NJ, USA: Wiley, 1998.

[18] H. K. T. Ng, P. S. Chan, and N. Balakrishnan, "Estimation of parameters from progressively censored data using EM algorithm," *Comput. Statist. Data Anal.*, vol. 39, pp. 371–386, 2002.

[19] Y. Takagi, "On the estimation of the shape parameter of the gamma distribution in second-order asymptotics," *Statist. Probability Lett.*, vol. 82, pp. 15–21, 2012.

[20] C. C. Tsai, S. T. Tseng, and N. Balakrishnan, "Optimal burn-in policy for highly reliable products using gamma degradation process," *IEEE Trans. Rel.*, vol. 60, no. 1, pp. 234–245, Mar. 2011.

[21] C. C. Tsai, S. T. Tseng, and N. Balakrishnan, "Optimal design for degradation tests based on gamma processes with random effects," *IEEE Trans. Rel.*, vol. 61, no. 2, pp. 604–613, Jun. 2012.

Narayanaswamy Balakrishnan is a Professor of Statistics at McMaster University, Hamilton, Ontario, Canada, and is a Visiting Professor in the Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia. He received his B.Sc., and M.Sc. degrees in statistics from the University of Madras, India, in 1976, and 1978, respectively. He finished his Ph.D. in statistics from the Indian Institute of Technology, Kanpur, India, in 1981. He is a *Fellow of the American Statistical Association, and a Fellow of the Institute of Mathematical Statistics*. He is currently the Editor-in-Chief of *Communications in Statistics*. His research interests include distribution theory, ordered data analysis, censoring methodology, reliability, survival analysis, nonparametric inference, and statistical quality control.

Debanjan Mitra is an Assistant Professor in the Department of Mathematics at Indian Institute of Technology Guwahati, India. He completed his Ph.D. from the Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada. He received his B.Sc., and M.Sc. degrees in Statistics from the University of Calcutta, India, in 2006, and 2008, respectively. His research interests include reliability and survival analysis, statistical inference, and censoring methodology. The work of this paper was completed when he was a Ph.D. student at McMaster University.