SciencePG
Science Publishing Group

Maximum Likelihood Estimation for a Progressively Type II Censored Generalized Inverted Exponential Distribution via EM Algorithm

Karuoya Grace Njeri[*], Edward Gachangi Njenga

Department of Mathematics and Actuarial Science, Kenyatta University (KU), Nairobi, Kenya

Email address:
Gnmungai@gmail.com (K. G. Njeri)
[*]Corresponding author

Abstract: The Generalized Inverted Exponential (GIE) distribution is a mixed lifetime model used in a number of fields such as queuing theory, testing of products or components and modelling the speed of winds. The study aims to focus on the determination of maximum likelihood estimates of GIE distribution when the test units are progressively (type II) censored. The scheme permits the withdrawal of units from the life test at stages during failure. This may be due to cost and time constraints. Both Expectation- Maximization (EM) and Newton-Raphson (NR) methods have been used to obtain the maximum likelihood estimates of the GIE parameters. Also, the variance-covariance matrix of the obtained estimators has been derived. The performance of the obtained MLEs via EM method is compared with those obtained using NR method in terms of bias and root mean squared errors and confidence interval widths for different progressive type II censoring schemes at fixed parameter values of λ and θ Simulation results reveal that estimates obtained via EM approach are more robust compared to those obtained via NR algorithm. It's also noted that the bias, root mean squared errors and confidence interval widths decrease with an increase in the sample size for a fixed number of failures. A similar trend in results is observed with increase in number of failures for a fixed sample size. The results of the obtained estimators are finally illustrated on two real data sets.

Keywords: Generalized Inverted Exponential Distribution, Progressive type II Censoring, EM Algorithm,
Newton Raphson Algorithm

1. Introduction

Censoring is a common feature in survival analysis and reliability experiments. According to Horst [9], a censored sample is one in which either by design or accidentally the event time of some items in the experiment are unobserved. There exist many different forms of censoring that have been studied by many authors in the past. Progressive type II censoring permits withdrawal of units at various levels of testing other than end point of the experiment. In so doing, this type of censoring provides saving in terms of cost associated with testing.

A number of studies based on progressive type II censoring have been carried out in the past using different univariate models like Gamma, log-normal, Exponential, Weibull and log-logistic distributions. However, due to the limitations associated with some of these models such as assumption of a constant hazard rate function for the exponential distribution, flexible mixed lifetime models have been adopted. For example, Inverted Exponential (IE) distribution can model data with inverse bath curve failure rate. A modification of inverse exponential distribution termed as Generalized inverted exponential (GIE) distribution is said to give a better fit to lifetime data compared to Inverted exponential (IE) distribution. The distribution was introduced by Abouammoh and Alshingiti [2] as model for business failure data.

This study, assumes that the lifetimes have GIE distribution. This distribution has a non-constant hazard rate function which is unimodal and positively skewed and increases or decreases as per the value of the shape parameter. Due to these properties, the distribution is able to model

different shapes of failure rates of ageing criteria.

The two parameter GIE distribution PDF and CDF are represented by:

$$f(x, \lambda, \theta) = \lambda\theta \exp\left(-\frac{\lambda}{x}\right) x^{-2} \left(1 - \exp\left(-\frac{\lambda}{x}\right)\right)^{\theta-1} \quad (1)$$

$$F(x, \lambda, \theta) = 1 - \left(1 - \exp\left(-\frac{\lambda}{x}\right)\right)^{\theta}, (x, \lambda, \theta) > 0 \quad (2)$$

Where λ and θ are scale and shape parameters respectively

Inferences for the GIE distribution in reference to censored data have been undertaken in a number of studies in the past.

Bakoban [4], did a study on Estimation in step stress partially accelerated life tests for GIE distribution using type I censoring. The, maximum likelihood estimators for the parameters, accelerated factor and asymptotic variance covariance matrix were derived. The performance of the estimators was compared using the mean squared error, absolute relative bias and relative error via Monte Carlo simulations.

Singh and Kumar [14] carried out a study on parameter estimation of GIE by computing the MLES and the Bayes estimates using a progressively censored sample with binomial removals. Using simulation, they compared their performance in terms of their risks.

Dey and Pradhan [8] considered estimations of GIE distribution under hybrid censoring scheme and derived MLEs and Fishers information matrix. Using Lindley approximation, the Bayes estimators under squared loss functions were derived.

Krishna and Kumar [10] scrutinized reliability estimation of progressive type II censored data with GIE distribution. The MLEs together with their reliability and failure rates were derived by means of least squares method and assessed using both real and simulated data sets.

In this paper we will focus on the inference for GIED via EM and carry out a comparative study of the estimators under NR algorithm.

2. Parameter Estimation

2.1. Progressive Type II Censoring

In this scheme of censoring from the total of n units placed simultaneously on an experiment only m is observed until the end i.e. failure time. Let $X_{1:m:n}, \ldots X_{m:m:n}$ be an order statistic exhibiting a progressively type II censoring scheme.

Given a censoring plan $R_1, R_2, \ldots R_m$ at the time $X_{1:m:n}$ the first failure time R_1 of the $n-1$ surviving units are randomly removed from the life test experiment. $X_{2:m:n}$ the second failure time, R_2 from the surviving units $n-2-R_1$ are removed.

At $X_{m:m:n}$ the mth failure all the remaining $n-2-R_1-R_2, \ldots, R_{m-1}$ surviving units are removed. Censoring takes place in progressively in m stages. This is referred to as progressive Type II right censoring with $R = R_1, R_2, \ldots, R_m$ scheme, as per Childs and Balakrishnan [6].

2.2. Maximum Likelihood Estimation

Let $X_{1:n}, X_{2:n} \ldots \ldots X_{m:n}$ denote a progressive Type II censored sample. Then, as per Child and Balakrishnan [6] the likelihood function for the sample is represented by

$$L(\theta, \lambda; x_{1:m}, \ldots. x_{m:n}) = k \prod_{i=1}^{m} f(x_i; \theta, \lambda)[1 - F(x_i; \theta, \lambda)]^{R_i}$$

Where $k = n(n - R_1 - 1) \ldots (n - R_1 - R_2 - R_{m-1} - m + 1)$ $\qquad (3)$

Using equations (1) and (2), the log likelihood function on θ and λ subject to progressive Type II censored sample can be written as follow:

$$l(\theta, \lambda; x_{1:m}, \ldots. x_{m;n}) \alpha \lambda^m \theta^m \prod_{i=1}^{m} x_i^{-2} e^{\frac{-\lambda}{x_i}} \left(1 - e^{\frac{-\lambda}{x_i}}\right)^{\theta-1} \left(\left(1 - e^{\frac{-\lambda}{x_i}}\right)^{\theta}\right)^{R_i} \qquad (4)$$

Which can then be represented as:

$$l(\theta, \lambda; x_{1:m}, \ldots. x_{m;n}) \alpha m ln\lambda + m ln\theta + \sum_{j=1}^{m} ln e_j^{-2} + \sum_{j1=}^{m} ln e^{\frac{-\lambda}{x_i}} + (\theta - 1) \sum_{j=1}^{m} ln\left(1 - e^{\frac{-\lambda}{x_i}}\right) + \theta \sum_{j=1}^{m} R_i ln\left(1 - e^{\frac{-\lambda}{x_i}}\right) \qquad (5)$$

2.3. Expectation Maximization Method

We suggest the expectation maximization method proposed by Dempster et al [7].

Let $Z = (Z_1, Z_2, \ldots, Z_M)$ with $Z_{j1}, Z_{j2}, \ldots Z_{jR_i}, j = 1,2, \ldots, m$.

be the censored data, where the censored data is regarded as missing data. The component (X, Z) = W forms the complete data set.

The log-likelihood function under W can be represented as:

$$H(W, \theta, \lambda) \alpha m ln\lambda + m ln\theta + \sum_{j=1}^{m} ln e_j^{-2} + \sum_{j1=}^{m} ln e^{\frac{-\lambda}{x_i}} + (\theta - 1) \sum_{j=1}^{m} ln\left(1 - e^{\frac{-\lambda}{x_i}}\right) + \sum_{j=1}^{m} \sum_{k=1}^{m} ln Z_{jk}^{-2}$$

$$+ \sum_{j=1}^{m} \sum_{k=1}^{R_j} ln e^{\frac{-\lambda}{Z_{jk}}} + (\theta - 1) \sum_{j=1}^{m} \sum_{k=1}^{R_j} ln\left(1 - e^{\frac{-\lambda}{Z_{jk}}}\right) \qquad (6)$$

In the E-step, the pseudo-likelihood function is obtained by replacing any function of Z_{jk} say, $h(z_{jk})$ by $E\left(h(Z_{jk})/_{Z_{jk}>x_j}\right)$ in equation (6) giving

$$H^*(w;\theta,\lambda)\alpha m ln\lambda + mln\theta + \sum_{j=1}^m lnx_j^{-2} + \sum_{j=1}^m lne^{\frac{-\lambda}{x_j}} + (\theta-1)\sum_{j=1}^m ln\left(1-e^{\frac{-\lambda}{x_j}}\right) + \sum_{j=1}^m \sum_{k=1}^{R_j} E\left[Z_{jk}^{-2}/_{Z_{jk}>x_j}\right]$$

$$+\sum_{j=1}^m \sum_{k=1}^{R_j} E\left[lne^{\frac{-\lambda}{z_{jk}}/_{Z_{jk}}} > x_j\right] + (\theta-1)\sum_{j=1}^m \sum_{k=1}^{R_j} E\left[ln\left(1-e^{\frac{-\lambda}{z_{jk}}/_{Z_{jk}}} > x_j\right)\right] \tag{7}$$

Thus, given $X_j = x_j$ the conditional distribution of Z_{jk} is a left truncated GIE distribution at x_j.
That is

$$f\left(Z_j/_{X_j=x_j}\right) = \frac{f_w(z_j)}{1-F_w(z_j)}, Z_j > x_j \tag{8}$$

Equation (7) becomes:

$$E\left(Z_{jk}^{-2}/_{Z_{jk}>x_j}\right) = \frac{\lambda\theta}{1-e^{\frac{-\lambda}{x_j}}} \int_{x_j}^1 e^{\frac{-\lambda}{z_j}} Z_j^{-2}\left(1-e^{\frac{-\lambda}{z_j}}\right)^{\theta-1} z_{jk}^{-2} dz_j \tag{9}$$

Similarly

$$E\left(e^{\frac{-\lambda}{z_{jk}}/z_{jk}} > x_j\right) = \frac{\lambda\theta}{\left(1-e^{\frac{-\lambda}{x_j}}\right)^\theta} \int_{x_j}^1 e^{\frac{-\lambda}{x_j}} z_j^{-2}\left(1-e^{\frac{-\lambda}{z_j}}\right)^{\theta-1} e^{\frac{-\lambda}{z_{jk}}} dz_j \tag{10}$$

and

$$E\left(1-e^{\frac{-\lambda}{z_{jk}}/z_{jk}} > x_j\right) = \frac{\lambda\theta}{\left(1-e^{\frac{-\lambda}{x_j}}\right)^\theta} \int_{x_j}^1 e^{\frac{-\lambda}{x_j}} z_j^{-2}\left(1-e^{\frac{-\lambda}{z_j}}\right)^{\theta-1}\left(1-e^{\frac{-\lambda}{z_{jk}}}\right) dz_j \tag{11}$$

The M-step maximizes the pseudo-likelihood function by substituting the conditional expectations in equation (7) with $E_1(v;\theta,\lambda)$ and $E_2(v;\theta,\lambda)$ gotten in equations (9, 10 and 11) respectively. Assuming that at the kth stage, the estimates for (θ,λ) are $\theta^{(k)}$, and $\lambda^{(k)}$ then differentiating equation (12) below with respect to θ and λ gives the values of $\theta^{(k+1)}$; $\lambda^{(k+1)}$

$$H^*(w,\theta,\lambda)\alpha\, mln\lambda + mln\theta + \sum_{j=1}^m lnx_j^{-2} + \sum_{j=1}^m lne^{\frac{-\lambda}{x_j}} + (\theta-1)\sum_{j=1}^m ln\left(1-e^{\frac{-\lambda}{x_j}}\right) + \sum_{j=1}^m R_j E_1\left(\theta_j;\theta^{(k)},\lambda^{(k)}\right)$$

$$+\sum_{j=1}^m R_j E_2\left(\theta_j;\theta^{(k)},\lambda^{(k)}\right) + (\theta-1)\sum_{j=1}^m R_j E_3\left(\theta_j;\theta^{(k)},\lambda^{(k)}\right) \tag{12}$$

Once θ^{k+1} is obtained λ^{k+1} is then obtained by solving the equation

$$\frac{\partial H^*(w,\theta,\lambda)}{\partial\lambda} = \frac{m}{\lambda} - \sum_{j=1}^m \frac{1}{x_j} + (\hat{\theta}-1)\sum_{j=1}^m \frac{e^{\frac{-\lambda}{x_j}}}{x_j\left[1-e^{\frac{-\lambda}{x_j}}\right]} = 0 \tag{13}$$

2.4. Newton Raphson Method

In this context, the MLEs of θ and λ can be obtained by differentiating equation (5) with respect to θ and λ then equate the

normal equations to 0.
Then;

$$\hat{\theta} = \theta(\lambda) = \frac{-m}{\sum_{i=1}^{m} ln\left(1-e^{\frac{-\lambda}{x_i}}\right)+\sum_{i=1}^{m} R_i ln\left(1-e^{\frac{-\lambda}{x_i}}\right)} \tag{14}$$

$$\frac{\partial l(x;\theta,\lambda)}{\partial \lambda} = \frac{m}{\lambda} - \sum_{i=1}^{m}\frac{1}{x_i} + (\theta-1)\sum_{i=1}^{m}\frac{e^{\frac{-\lambda}{x_i}}}{x_i\left(1-e^{\frac{-\lambda}{x_i}}\right)} + \theta\sum_{i=1}^{m} R_i \frac{e^{\frac{-\lambda}{x_i}}}{\left(1-e^{\frac{-\lambda}{x_i}}\right)} \tag{15}$$

combining (14) and (15) gives a non-linear normal equation in

$$\frac{\partial l(x;\hat{\theta},\lambda)}{\partial \lambda} = \frac{m}{\lambda} - \sum_{i=1}^{m}\frac{1}{x_i} + \left(\hat{\theta}-1\right)\sum_{i=1}^{m}\frac{e^{\frac{-\lambda}{x_i}}}{x_i\left(1-e^{\frac{-\lambda}{x_i}}\right)} + \hat{\theta}\sum_{i=1}^{m} R_i \frac{e^{\frac{-\lambda}{x_i}}}{\left(1-e^{\frac{-\lambda}{x_i}}\right)} = 0 \tag{16}$$

The NR algorithm is then executed using the following procedure:
let

$$U(\lambda,x) = \frac{m}{\lambda} - \sum_{i=1}^{m}\frac{1}{x_i} + \left(\hat{\theta}-1\right)\sum_{i=1}^{m}\frac{e^{\frac{-\lambda}{x_i}}}{x_i\left(1-e^{\frac{-\lambda}{x_i}}\right)} + \hat{\theta}\sum_{i=1}^{m} R_i \frac{e^{\frac{-\lambda}{x_i}}}{\left(1-e^{\frac{-\lambda}{x_i}}\right)}$$

$$U'(\lambda,x) = \frac{m}{\lambda^2} + \left(\hat{\theta}-1\right)\sum_{i=1}^{m}\frac{e^{\frac{-\lambda}{x_i}}}{x_i^2\left(1-e^{\frac{-\lambda}{x_i}}\right)}\left(\frac{e^{\frac{-\lambda}{x_i}}}{1-e^{\frac{-\lambda}{x_i}}} - 1\right) \tag{17}$$

Giving the information matrix

$$I(\lambda,x) = -\frac{m}{\lambda^2} - \left(\hat{\theta}-1\right)\sum_{i=1}^{m}\frac{e^{\frac{-\lambda}{x_i}}}{x_i^2\left(1-e^{\frac{-\lambda}{x_i}}\right)}\left(\frac{e^{\frac{-\lambda}{x_i}}}{1-e^{\frac{-\lambda}{x_i}}} - 1\right) \tag{18}$$

The MLE of λ is obtained by evaluating:

$$\lambda^{(i)} = \lambda^{(i-1)}\frac{U(\lambda^{(I-1)},x)}{I(\lambda^{(i)},x)} \tag{19}$$

Until convergence. The obtained value is then embedded in equation (14) to obtain the MLE of θ

$$\theta^{(k+1)=\hat{\theta}=} \frac{-m}{\sum_{j=1}^{m}\left(1-e^{\frac{-\lambda}{x_j}}\right)\sum_{j=1}^{m} R_j\, E_3\left(\theta_j;\theta^{(k)},\lambda^{(k)}\right)}$$

2.5. Asymptotic Variance Covariance Matrix of MLEs

The variance–covariance matrix quantifies the accuracy of the parameter estimators using the log-likelihood function. The variance– covariance matrix of parameters θ and λ from the GIE distribution complete data set is first computed.

$$lnf = ln(\lambda) + ln(\theta) - \frac{\lambda}{z_j} + 2lnz_j + (\theta-1)\ln\left(1-e^{\frac{-\lambda}{z_j}}\right) - \theta ln\left(1-e^{\frac{-\lambda}{x_j}}\right) \tag{20}$$

To obtain the observed information matrix, a procedure by Louis and Tanner [11] is adopted. The Fisher information matrix for complete data set is given as;

$$I_c(\theta,\lambda) = -E\begin{bmatrix} \frac{d^2 lnl_c}{d^2\theta} & \frac{d^2 lnl_c}{d\theta d\lambda} \\ \frac{d^2 lnl_c}{d\theta d\lambda} & \frac{d^2 lnl_c}{d^2\theta} \end{bmatrix} = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix} \tag{21}$$

The variance-covariance matrix of parameters of λ and θ

is then obtained by

$$\begin{bmatrix} var(\hat{\theta}) & var(\hat{\theta}, \hat{\lambda}) \\ var(\hat{\theta}, \hat{\lambda}) & var(\hat{\lambda}) \end{bmatrix} = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}^{-1} \quad (22)$$

The missing information matrix is obtained using:

$$I_{miss} = \sum_{j=1}^{m} R_j I_{miss}^{(j)} (\eta)$$

Therefore, the variance-covariance matrix of parameter $\boldsymbol{\theta}$ can be obtained by

$$I_{obs}^{-1} = \left\{ \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix} - \begin{bmatrix} \sum_{j=1}^{m} R_j I_{11}^{(J)} & \sum_{j=1}^{m} R_j I_{12}^{(J)} \\ \sum_{j=1}^{m} R_j I_{21}^{(J)} & \sum_{j=1}^{m} R_j I_{22}^{(J)} \end{bmatrix} \right\}^{-1} \quad (23)$$

From equation (23) an approximate $100(1 - \alpha)\%$ confidence intervals for θ and λ are given as;

$$\hat{\theta} \pm z_{\frac{\alpha}{2}} \sqrt{var(\hat{\theta})} \text{ and } \hat{\lambda} \pm z_{\frac{\alpha}{2}} \sqrt{var(\hat{\lambda})}$$

3. Results and Discussion

3.1. Illustration with Simulated Data

Using R statistical software, different progressive type II right censored samples from GIE distribution with parametric space; ($\theta = 1.4$; $\lambda = 1.65$) and ($\theta = 0.95$; $\lambda = 1.2$) are generated. Balakrishnan and Sandhu [5] and Aggarwala and Balakrishnan [3] algorithms were considered in generating these samples. The algorithms have the following steps;

(i) Generate m independent U (0, 1) random variables W1; W2; ……Wm:

(ii) For given values of the progressive censoring scheme R1; R2; ……. Rm:

(iii) Set

$$V_i = W_i^{\frac{1}{(i+R_m+R_{m-1}+\cdots+R_{m-i+1})}}, i = 1, \dots, m$$

iv) Set

$$U_i = 1 - V_M, V_{m-1}, \dots, V_{m-i+1} \ i = 1, \dots, m$$

is a progressive Type II censored sample of size m from U(0,1).

(v) Set $Y_i = F^{-1}(U_i) \ i = 1, \dots, m$ as the required progressive Type II censored sample of size m from GIE distribution.

These steps were repeated several times for different sample sizes and different censoring schemes. The term different censoring schemes refers to different sets of $R_i s$. In this study, samples of sizes 20 and 45 were used and the censoring schemes considered were as follows:

scheme 1 1 $R_1 = \cdots = R_{m-1} = 0, R_m = n - m$

scheme 2 2 $R_1 = n - m$ and $R_2 = \cdots = R_m = 0$

scheme 3 3 $R_1 = \cdots = R_{m-1}$ and $R_m = n - 2m + 1$

Table 1. *Average Bias of the MLE estimators of θ and λ obtained via NR and EM algorithms.*

n	m	scheme	Bias of θ		Bias of λ	
			NR	EM	NR	EM
when $\theta = 1.40, \lambda = 1.65$						
20	8	1	0.390	0.381	0.489	0.435
20	8	2	0.046	0.037	0.145	0.438
20	8	3	0.099	0.089	0.198	0.146
45	20	1	0.123	0.114	0.222	0.171
45	20	2	0.016	0.007	0.115	0.064
45	20	3	0.110	0.101	0.209	0.158
When $\theta = 0.95 \ \lambda = 1.2$						
20	8	1	0.193	0.237	0.292	0.294
20	8	2	0.071	0.0231	0.170	0.081
20	8	3	0.179	0.103	0.278	0.160
45	20	1	0.059	0.048	0.158	0.104
45	20	2	0.021	0.018	0.120	0.075
45	20	3	0.073	0.039	0.172	0.096

Table 2. *Average RMSEs of the MLE estimators of θ and λ obtained via NR and EM algorithms.*

n	m	scheme	Bias of θ		Bias of λ	
			NR	EM	NR	EM
When $\theta = 1.40, \lambda = 1.65$						
20	8	1	0.479	0.0.471	0.577	0.523
20	8	2	0.308	0.303	0.406	0.356
20	8	3	0197	0.192	0.295	0.248
45	20	1	0.310	0.273	0.408	0.329
45	20	2	0.240	0.231	0.338	0.2875
45	20	3	0.281	0.302	0.379	0.358
When $\theta = 0.95 \ \lambda = 1.2$						

n	m	scheme	Bias of θ		Bias of λ	
			NR	EM	NR	EM
20	8	1	0.211	0.178	0.309	0.234
20	8	2	0.069	0.052	0.167	0.103
20	8	3	0.117	0.086	0.214	0.142
45	20	1	0.129	0.118	0.227	0.174
45	20	2	0.081	0.077	0.179	0.1333
45	20	3	0.120	0.117	0.231	0.173

***Table 3.** Approximate 95% confidence interval widths of the MLE estimators of obtained via NR and EM algorithms.*

n	m	scheme	Bias of θ		Bias of λ	
			NR	EM	NR	EM
When $\theta = 1.40, \lambda = 1.65$						
20	8	1	1.8474	1.2416	2.2892	1.4307
20	8	2	1.8634	1.2423	2.3039	1.4399
20	8	3	1.9187	1.2792	2.3549	1.4719
45	20	1	1.8390	1.2260	2.2815	1.4259
45	20	2	1.8410	1.2273	2.2833	1.4271
45	20	3	1.8445	1.2297	2.2865	1.4291
When $\theta = 0.95 \; \lambda = 1.2$						
20	8	1	1.7037	1.1358	2.1586	1.3491
20	8	2	1.7729	1.1819	2.2211	1.3882
20	8	3	1.8076	1.2052	2.2529	1.4080
45	20	1	1.6759	1.1172	2.1337	1.3335
45	20	2	1.6999	1.1333	2.1552	1.3470
45	20	3	1.5406	1.0271	2.0144	1.2589

The results from tables 1, 2 and 3 indicates that:

An increase in the sample size leads to a decrease in the bias estimates irrespective of the censoring schemes.

The bias estimates, RMSE and the 95% confidence interval widths obtained by the EM algorithm are relatively small compared to those obtained via NR algorithm.

A decrease in the true parameter values of θ and λ leads to a decrease in bias estimates.

In the context of censoring schemes, the bias estimates, RMSE associated with censoring scheme 2 are relatively smaller compared to those associated with censoring schemes 1 and 3.

The RMSEs estimates decrease with an increase in the sample size for a fixed number of failures.

For a fixed sample size, the RMSEs and 95% confidence interval widths decrease as the number of failures increases for both EM and NR methods.

A decrease in the true parameter values of θ and λ causes a decrease in RMSEs estimates.

When the number of failures (m) is fixed, the 95% confidence interval widths decreases as the sample sizes (n) increases for both EM and NR algorithms.

3.2. Illustration with Real Data

3.2.1. Communication Transceiver Maintenance Data

The maintenance data reported on active repair times (hours) for an airborne communication transceiver by Von Alven [15] is considered.

The repair times are as follows:

0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1.0, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 7.0, 7.5, 3.0, 3.0, 3.3, 3.3, 4.0, 4.4, 4.5, 4.7, 5.0, 5.4, 5.4, 8.8, 9.0, 10.3, 22.0, 24.5.

The data set is confirmed to provide a good fit to GIE distribution (Kolmogorov Smirnoff statistic of 0.152 (p-value=0.2393)). These results are consistent with previous work by Sanku Dey and Tanujit Dey [12]

***Table 4.** Progressively type II censored samples from maintenance data.*

n	m	$\sum_{j=1}^{m} R_j$	Censoring scheme	Samples generated from original data
46	12	34	$1(6,5,3,0^{*6}, 10,5,5)$	0.2,0.6,1.1,1.3,1.5,1.5,1.5,1.5,2.0,4.0, 5.4
464 46	20	26	$2. (3,2,1,0^{*14}, 5,5,10)$	0.2,0.5,0.6,0.7,0.7,0.8,0.8,1.0,1.0, 1.0,1.0,1.1,1.3,1.5,1.5,1.5,1.5,2.0,7.5,4.4
46	30	16	$3. (0^{*26}, 3,3,5,5)$	0.2,0.3,0.5,0.5,0.5,0.5,0.5,0.6,0.6,0.7,0.7,0.8,0.8,1.0,1.0,1.0,0.8,0.8,1.0, 1.0,1.0,1.0,1.1,1.3,1.5,1.5,1.5,1.5,2.0,0.7,2.0,2.2,2.5,3.0,4.0,5.4

***Table 5.** Average Bias of θ and λ and affiliated RMSE (in parenthesis) of the MLE estimators calculated using NR and EM algorithms for maintenance data.*

n	m	scheme	Bias &RMSE of θ		Bias &RMSE of λ	
			NR	EM	NR	EM
46	12	1	0.0526 (0.0415)	.0281 (0.0296)	0.1357 (0.1539)	0.0739 (0.0537)
46	20	2	0.0245 (0.0617)	0.0112 (0.0343)	0.2110 (0.0797)	0.1292 (0.0.0256)
46	30	3	0.0064 (0.0312)	0.00371 (0.0119)	0.0689 (0.0212)	0.06459 (0.0199)

Table 6. *Approximate 95% Confidence Interval widths of the MLE estimators of θ and λ for Communication Transceiver Maintenance data.*

n	m	scheme	95%CIwidthof θ		95%CI of λ	
			NR	EM	NR	EM
46	12	$(6,5,3,0^{*6},10,5,5)$	1.1809	1.0501	1.1392	1.0625
46	20	$(3,2,1,0^{*14},5,5,10)$	1.0191	0.9649	1.0399	0.9798
46	30	$(0^{*36},3,3,5,5)$	1.03	0.9877	1.0808	1.0191

The results in tables 5 and 6 indicate that:

The bias and RMSE of the maximum likelihood estimates obtained via EM method are smaller compared to those obtained via NR algorithm.

For a fixed sample size, an increase in the number of failures leads to reduced bias across most of the censoring schemes.

The confidence interval widths using EM method are smaller compared to those obtained via NR algorithm.

For a fixed sample size, the higher the number of failures the smaller the confidence interval widths within which the true parameter values lie.

Generally, censoring scheme 2 yields smaller confidence interval widths compared to other censoring schemes. This implies that schemes whose most units are censored at the Centre of the failure process tends to give better and more accurate confidence interval widths.

3.2.2. Ball Bearing Failure Data

The data originated from tests on survival of deep groove ball bearings. The data shows the number of million cycles before failure for each of the ball bearings in the lifetime as shown below:

17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64,

68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, and 173.40.

A Kolmogorov-Smirnoff test shows that the data fits GIE distribution (K-S statistic, D=0.090524 (p-value=0.9917)). These results are in consistence with previous work by Abouammoh and Alshingiti [2] and Krishna and Kumar [7].

Table 7. *Table of progressively type II censored samples from ball bearing data.*

n	m	$\sum_{j=1}^{m} R_j I_{miss}^{(j)}$	Censoring scheme	Samples generated from original data
23	7	16	$1 \begin{pmatrix} 0,4,0,2,0,3,0,2,0,1 \\ ,3,0,1 \end{pmatrix}$	17.88,45.60,51.96,68.64,84.12,98.64,128.04.
23	12	11	$2. \begin{pmatrix} 0^*,2,6,0^*,3,1,0,10^* \\ 4,3,0^*,2 \end{pmatrix}$	17.88,28.92,51.90,54.12,55.56,68.64,68.88,84.12,93.12,98.64,128.04,173.40
23	15	8	$3. \begin{pmatrix} 1,0^*,4,0^*,4,1,0^*,3,3, \\ 0^*1,0^*,3 \end{pmatrix}$	28.92,33.00,45.60,48.80,51.84,51.96,55.50,67.80,68.64,93.12,98.64,105.12,127. 92,128,04,173.40

Table 8. *Average Bias θ and λ and affiliated RMSE (in parenthesis) of the MLE estimators calculated using NR and EM algorithms for the ball bearing data.*

n	m	scheme	Bias &RMSE of θ		Bias &RMSE of λ	
			NR	EM	NR	EM
23	7	1	0.07337 (0.597)	0.07133 (0.489)	0.3366 (0.536)	0.2611 (0.129)
23	12	2	0.0456 (0.318)	0.0312 (0.311)	0.09077 (0.109)	0.0821 (0.117)
23	15	3	0.0541 (0.297)	0.0399 (0.276)	0.1833 (0.332)	0.1041 (0.079)

Table 9. *Approximate 95%confidence interval widths of the MLE estimators of θ and λ for ball bearing failure data.*

n	m	Scheme	95%CIwidth of θ		95%CI width of λ	
46	12	(0,4,0,2,0,3,0,2,0,1,3,0,1)	2.5972	2.8995	3.7577	3.6291
46	20	$(0^*,2,6,0^*,3,1,0,1,0^*4,3,0^*,2,)$	2.4232	2.7097	3.5904	3.4851
46	30	$(1,0^*,2,4,0^*,4,1,0^*,3,3,0^*,3,1,0^*,3)$	2.4595	2.8013	3.6755	3.5543

Results in both tables 8 and 9 indicates that:

The bias and RMSE of the maximum likelihood estimates obtained via EM method are smaller compared to those obtained via NR algorithm. Similarly, an increase in the number of failures leads to smaller values of bias.

The confidence interval widths using EM method are fairly smaller compared to those obtained via NR algorithm.

For a fixed sample size, the higher the number of failures the smaller the confidence interval widths.

Generally, censoring scheme 2 yields smaller confidence interval widths compared to other censoring schemes.

4. Conclusion

The study focused on estimating the parameters of GIE distribution based on progressively type II censored samples. The MLEs were obtained using the EM as well as NR approaches. The results were assessed based on bias, root mean squared errors and confidence interval widths using both simulated and real data. From the simulated data, the results reveal the following:

The average estimates of bias, and RMSEs obtained by EM method are fairly smaller than those obtained by NR

algorithm. This means that the MLEs via EM are more consistent.

The 95% confidence interval widths for MLE via EM are shorter than via NR.

a decrease in the true parameter values of θ and λ leads to a decrease in the average estimates of bias, RMSEs and confidence interval widths.

These results are consistent with previous studies that have used both EM and NR algorithms such as Amal et al [1], Wang, and cheng [16]. Salem and Abo-Kasem [13].

This means that EM algorithm outperforms the NR algorithm especially in analysis of large data.

References

[1] Amal Helu, Hani Samawi and Mohammad Z. Raqab (2013). Estimation on lomax progressive censoring using the EM algorithm. Journal of Statistical Computation and Simulation, 837-861.

[2] Abouammoh A. M and Alshingiti (2009). Reliability estimation of the generalized inverted exponential distribution. Journal of Statistical Computation and Simulation, 79 (11), 1301-1315.

[3] Aggarwala R. and Balakrishnan N. (1998). Some Properties of Progressive Censored Order Statistics from arbitrary and uniform distributions with applications to Inference and Simulations. Journal of Statistics and Planning Inference, 70, 35-49.

[4] Bakoban R. A (2012). Estimation in step stress Partially accelerated life tests for the Generalized inverted exponential distribution using type I censoring. American Journal of scientific Research, 3, 25-35.

[5] Balakrishnan N. and Sandhu A. (1995). A simple simulation algorithm for generating progressive Type II censored sample. American journal of Statistics, 49, 229-230.

[6] Childs A. and Balakrishnan N. (2000). Conditional inference procedures for the Laplace distribution when the observed samples are progressively censored. Journal of Metrika, 52 (3), 253-265.

[7] Dempster A. P, Laird N. M, Rubin D. D. (1977). Maximum likelihood from incomplete data via EM algorithm. Journal of the Royal Statistical Society, Series B, 39, 1-38.

[8] Dey and Pradhan (2014). Generalized exponential distribution under hybrid censoring. Statistical Methodology, 18, 101-114.

[9] Horst R. (2009). The Weibull Distribution Handbook. CRC Press, Taylor and Francis Group.

[10] Krishna H and Kumar K (2013). Reliability estimation in generalized inverted exponential distribution with progressively type II censored sample. Journal of statistical computation and simulation, 83, 1007-1019.

[11] Louis T. A (1982). Finding the observed matrix when using the EM algorithm. *Journal of the Royal Statistical Society Series B.*, 44, 226-233.

[12] Sanku D. and Tanujit D. (2016). Statistical inference for the Generalized inverted exponential distribution based on upper record values. Journal of Mathematics and Computers in simulation, 4200.

[13] Salem A. M and Abo-Kasem O. E (2011). International journal of Contemporary Mathematics and Sciences. (6) 35, 1713-1724.

[14] Singh S. K, Singh U. and Kumar M. (2013). Singh S. K, Singh U. and Kumar M. (2013). Estimation of parameters of generalized inverted exponential distribution for progressive type II censored sample with Binomial removals. Journal of probability and statistics, Article ID 183652 http://dxi.org./10.1155/2013/183652.

[15] Von Alven (1964 pg 156).

[16] Wang and Cheng (2009) EM algorithm for estimating the Burr XII parameters with multiple censored data. Quality and reliability Engineering Journal International. 30 (5), 1622-1638.