



Estimation of Parameters of the Two-Parameter Rayleigh Distribution Based on Progressive Type-II Censoring Using Maximum Likelihood Method via the NR and the EM Algorithms

Murithi Daniel Fundi*, Edward Gachangi Njenga, Kemboi George Keitany

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

Email address:

fund.danie@gmail.com (M. D. Fundi), enjenga2@yahoo.co.uk (E. G. Njenga), gekemboi@gmail.com (K. G. Keitany)

*Corresponding author

To cite this article:

Murithi Daniel Fundi, Edward Gachangi Njenga, Kemboi George Keitany. Estimation of Parameters of the Two-Parameter Rayleigh Distribution Based on Progressive Type-II Censoring Using Maximum Likelihood Method via the NR and the EM Algorithms. *American Journal of Theoretical and Applied Statistics*. Vol. 6, No. 1, 2017, pp. 1-9. doi: 10.11648/j.ajtas.20170601.11

Received: November 15, 2016; **Accepted:** November 30, 2016; **Published:** December 20, 2016

Abstract: In this article, Maximum likelihood estimates for the shape and scale parameters of two-parameter Rayleigh distribution are obtained based on progressive type-II censored samples using the Newton-Raphson (NR) method and the Expectation-Maximization (EM) algorithm. A simple algorithm discussed in [2-3] is used for generating progressive type-II censored samples. Based on this censoring scheme, approximate asymptotic variances are derived and used to construct approximate confidence intervals of the parameters. The performance of these two maximum likelihood estimation algorithms is compared in terms of simulation results of root mean squared error (RMSE) and the coverage rates. Simulation results showed that in nearly all the combination of simulation conditions the estimators based on the EM algorithm have less root mean squared error (RMSE) and narrower widths of confidence intervals compared to those obtained using the NR algorithm. Finally, an illustrative example with real-life data sets is provided to illustrate how maximum likelihood estimation using the two algorithms works in practice.

Keywords: Two-Parameter Rayleigh Distribution, Maximum Likelihood Estimation, EM Algorithm, NR Method, Progressive Type-II Censoring

1. Introduction

The two-parameter Rayleigh distribution is a particular case of a Weibull distribution widely used in reliability theory and life testing. Rayleigh [25] introduced this distribution in connection with a problem in acoustics. Rayleigh distribution has a nice relation to other distributions including Chi-Square and most extreme value distributions. In addition, the hazard function of this distribution increases with an increase in time. As a result, the distribution has attracted several researchers as it occurs in different forms including one-parameter Rayleigh distribution, and two-parameter Burr type X distribution also known as the Generalized Rayleigh distribution. According to Surles and Padgett, the two-parameter Rayleigh distribution is an

extreme value distribution that is effective in modeling general life data [26].

In literature, several distinguished authors have extensively studied estimation, inferential, and predictions issues for one-parameter Rayleigh distribution although not much has been done on two-parameter Rayleigh distribution. Interested readers are referred to [9, 10, 15, and 16] for exposure to the Rayleigh distribution.

Recently, Khan, Provost, and Singh [17] considered the predictive inference based on doubly censored samples for the two-parameter Rayleigh distribution. Very recently, Dey, Dey, and Kundu [12] derived interval and point estimates of the scale and location parameters of a two-parameter

Rayleigh distribution using progressive Type-II censored samples.

A continuous random variable X is said to have a two-parameter Rayleigh distribution with a scale parameter λ and location parameter μ , if its density function is given by:

$$f(x; \lambda, \mu) = 2\lambda(x - \mu)e^{-\lambda(x - \mu)^2}; \lambda > 0, x > \mu, 0 < \mu < \infty \quad (1)$$

The corresponding distribution function for $x > \mu$ is given by:

$$F(x; \lambda, \mu) = 1 - e^{-\lambda(x - \mu)^2}; \lambda > 0, x > \mu \quad (2)$$

The presence of the location parameter makes the two-parameter more effective in analyzing real life data sets compared to one-parameter Rayleigh distribution.

In reliability testing, an experimenter may cease testing before all the experimental units fail due to time constraint or lack of funds. Samples that results from such situations are known as censored samples. There are numerous censoring methods available to an experimenter with type-II and type-I censoring schemes being the commonly used schemes in life testing. A mixture of these two schemes results to a hybrid censoring scheme. However, type-II, type I, and hybrid censoring schemes do not give room for removal of experimental units before the terminal point of the experiment. Progressive type-II censoring scheme allows such removal hence it gained popularity in life-testing and reliability experiments. In this paper, we consider progressive Type-II censoring scheme.

In the recent statistical literature, progressive censoring scheme has attracted many reliability practitioners and theoreticians. Interested readers are referred to [2-4]. For more recent references, refer to [24, 27], as well as references, cited therein.

Recently, Lio, Chen, & Tsai [19] investigated inference of the estimated parameters of the generalized Rayleigh distribution based on progressive type-I interval censoring scheme. The study reviewed that use of progressive type I interval censored samples to estimates the MLEs using Expectation Maximization algorithm yields more accurate and precise parametric estimates. Very recent, Dey et al. [12] derived interval and point estimates of the scale and location parameters of a two-parameter Rayleigh distribution using progressive Type-II censored samples.

The purpose of this article is to develop an estimation procedure for the scale and shape parameters of the two-parameter Rayleigh distribution based on progressive type-II censoring scheme. We first derive the maximum likelihood estimators of the unknown parameters. Since the MLEs of the shape and scale parameters of the two-parameter Rayleigh distribution cannot be obtained in the explicit form, we propose the use of the NR and the EM algorithms to compute the MLEs. Progressive type-II right censored samples are considered as incomplete data hence both the EM and the NR algorithms are suitable numerical iterative procedures for finding the MLEs. For more information

regarding the EM algorithm including its application and advantages compared to those of NR method readers are referred to [1, 18, and 29]. For derivation and application of the Newton-type method refer to [20, 21, and 23].

The rest of the article is organized as follows. In section 2, progressive type-II censoring scheme is briefly discussed, the MLEs of the scale and location parameters are derived based on progressive type-II censoring using the EM and NR algorithms. Based on this censoring scheme, approximate asymptotic variances are derived and used to construct approximate confidence intervals of the parameters. In Section 3, simulation results and discussions are provided. In section 4, an illustrative example is provided using real-life data sets. In the final section, a conclusion is provided.

2. Parameter Estimation

2.1. Progressive Type-II Censoring Scheme

Let n identical items be put on a life-testing experiment at time 0 with the corresponding lifetimes $X_1, X_2, X_3, \dots, X_n$ being independent and identically distributed with the density function given in equation (1). Further, suppose that integer $m < n$ is fixed at the beginning of the experiment (where $m < n$ is the number of units to be observed completely until

failure) with $R_j \geq 0$ and $\sum_{j=1}^m R_j + m = n$ specified.

This implies that progressive censoring will occur in m failure stages as follows. At the time of the first failure, a random sample of size R_1 , ($X_1: m: n$) surviving items are randomly drawn from $n-1$ remaining surviving units in the experiment leaving $n-1-R_1$ survival units. At the time of the second failure, a random sample of size R_2 , ($X_2: m: n$) is randomly drawn from $n-2$ surviving units leaving $n-2-R_1$ surviving items in the experiment. The process is continued until the m^{th} failure time $X_{m:n:m}$ is evident (the m^{th} stage) when all $R_m = n - m - R_1 - R_2 - \dots - R_{m-1}$ surviving items are removed from the life-testing experiment. The set of an observed lifetime $X = X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}$ is a progressively type II censored sample. According to Balakrishnan and Aggarwala, a progressively type -II censoring scheme consist of m failure stages and $R_1, R_2, \dots,$

R_m random samples, such that $n - m = \sum_{j=1}^m R_j$ with R_j 's

fixed before the study. Where R_j denotes the j^{th} censored random sample [2].

It is imperative to note that, if $m=0$ there is no censoring, if $R_1 = R_2 = \dots = R_m = 0$, then $n=m$ (complete sample situation), and if $R_1 = R_2 = \dots = R_{m-1} = 0$, then $m = n - R_m$, which is the conventional type-II right censoring scheme. In this article, we will use x_j in place of

$X_{j:m:n}$ for $j=1, 2, 3, \dots, m$ to make the notation simple.

2.2. Maximum Likelihood Estimation Based on Progressive Type-II Censoring

MLE is one of the standard techniques for estimating unknown parameters of distribution or a model. The principle concept behind this method is to select the value of the parameter under which the underlying data is most likely to be observed.

Suppose n identical units are placed at the same time on a

$$L(\lambda, \mu; x_{1:n}, x_{2:n}, \dots, x_{m:n}) = K \prod_{j=1}^m f(x_j; \lambda, \mu) [1 - F(x_j; \mu, \lambda)]^{R_j} \tag{3}$$

Substituting the value of $f(\cdot)$ and $F(\cdot)$ in equation (3), the log likelihood function of μ and λ constructed on progressive type-II censored sample ignoring the constant term can be written as follows:

$$L(\lambda, \mu; x_{1:n}, x_{2:n}, \dots, x_{m:n}) \propto \prod_{j=1}^m \lambda (x_j - \mu) e^{-\lambda(x_j - \mu)^2} \left[1 - \left[1 - e^{-\lambda(x_j - \mu)^2} \right] \right]^{R_j} \tag{4}$$

The log-likelihood function of (4) is written as:

$$l(\lambda, \mu; x) \propto m \ln \lambda + \sum_{j=1}^m \ln(x_j - \mu) - \lambda \sum_{j=1}^m (R_j + 1)(x_j - \mu)^2 \tag{5}$$

2.3. Expectation-Maximization (EM) Algorithm

Let $X = X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}$ with $X_{1:m:n} < X_{2:m:n} < \dots < X_{m:m:n}$ denotes the progressive type-II right-censored data from a population with density function and distribution given in equations (1) and (2), respectively.

We propose the use of EM algorithm discussed in [7] as follows.

Let some of the complete data vectors W be observed such

$$l_c(W; \mu, \lambda) \propto n \ln \lambda + \sum_{j=1}^m \ln(y_j - \mu) - \lambda \sum_{j=1}^m (y_j - \mu)^2 + \sum_{j=1}^m \sum_{k=1}^{R_j} \ln(z_{jk} - \mu) - \lambda \sum_{j=1}^m \sum_{k=1}^{R_j} (z_{jk} - \mu)^2 \tag{6}$$

The MLEs of the parameters λ and μ based on W are obtained as:

$$\frac{\partial l_c(W; \mu, \lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{j=1}^m (y_j - \mu)^2 - \sum_{j=1}^m \sum_{k=1}^{R_j} (z_{jk} - \mu)^2 = 0 \tag{7}$$

$$\frac{\partial l_c(W; \mu, \lambda)}{\partial \mu} = -\sum_{j=1}^m (y_j - \mu)^{-1} + 2\lambda \sum_{j=1}^m (y_j - \mu) - \sum_{j=1}^m \sum_{k=1}^{R_j} (z_{jk} - \mu)^{-1} + 2\lambda \sum_{j=1}^m \sum_{k=1}^{R_j} (z_{jk} - \mu) = 0 \tag{8}$$

The E-Step of the EM algorithm requires substituting any function of Z_{jk} (say $h(Z_{jk})$) by $E(h(Z_{jk}) / Z_{jk} > y_j)$. Hence, equations (7) and (8) becomes

$$\frac{\partial l_c(W; \mu, \lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{j=1}^m (y_j - \mu)^2 - \sum_{j=1}^m \sum_{k=1}^{R_j} E\left\{ (z_{jk} - \mu)^2 / Z_{jk} > y_j \right\} = 0 \tag{9}$$

$$\frac{\partial l_c(W; \mu, \lambda)}{\partial \mu} = -\sum_{j=1}^m (y_j - \mu)^{-1} + 2\lambda \sum_{j=1}^m (y_j - \mu) - \sum_{j=1}^m \sum_{k=1}^{R_j} E\left\{ (z_{jk} - \mu)^{-1} / Z_{jk} > y_j \right\}$$

life-testing experiment. Let $x_{1:n}, x_{2:n}, \dots, x_{m:n}$ be a progressive type-II censored random sample from density function in equation (1). According to Balakrishnan and Aggarwala, m ordered failures out of the sample of size n are observed under this scheme and random samples R_1, R_2, \dots, R_m of survival units drawn and removed from the experiment at each of m^{th} failure stage [2]. The likelihood function based on progressive type-II censored random samples as in [2] is given by:

that $W = (Y; Z)$, where $Z = (z_1, z_2, \dots, z_m)$ and $Z_j = (Z_{j1}, Z_{j2}, \dots, Z_{jR_j})$, for $j=1, 2, \dots, m$ denotes the censored data (missing data) and $Y = (y_1, y_2, \dots, y_m)$ denotes the observed data.

The log-likelihood function of the complete data set can be written as:

$$+2\lambda \sum_{j=1}^m \sum_{k=1}^{R_j} E\left\{(z_{jk} - \mu) / Z_{jk} > y_j\right\} = 0 \quad (10)$$

We make use of theorem by Ng et al. [22] that states that given $Y_j = y_j$; the conditional distribution of Z_{jk} follows a truncated two-parameter Rayleigh distribution with left truncation at y_j . Hence,

$$f_{Z/Y}(z_j / \mathbf{Y}) = \frac{f_W(z_j)}{1 - F_W(y_j)}, z_j > y_j$$

The conditional expectations in equations (9) and (10) are obtained as:

$$E\left\{(z_{jk} - \mu)^2 / Z_{jk} > y_j\right\} = \frac{2\lambda(y_j - \mu)^2 + 3}{2\lambda} \quad (11)$$

$$E\left\{(z_{jk} - \mu)^{-1} / Z_{jk} > y_j\right\} = (y_j - \mu)^{-1} \quad (12)$$

$$E\left\{(z_{jk} - \mu) / Z_{jk} > y_j\right\} = \frac{\lambda(y_j - \mu)^2 + 1}{\lambda(y_j - \mu)} \quad (13)$$

$$\begin{aligned} \frac{\partial l_c(W; \mu, \lambda)}{\partial \mu} &= -\sum_{j=1}^m (y_j - \mu^{(h)})^{-1} + 2\lambda^{(h+1)} \sum_{j=1}^m (y_j - \mu^{(h+1)}) - \sum_{j=1}^m R_j (y_j - \mu^h)^{-1} + \\ &+ 2\lambda^{(h+1)} \sum_{j=1}^m R_j \frac{\lambda^{(h+1)}(y_j - \mu^h)^2 + 1}{\lambda^{(h+1)}(y_j - \mu^h)} = 0 \quad (15) \\ \mu^{(h+1)} &= \frac{\sum_{j=1}^m y_j}{m} - \frac{\left[\sum_{j=1}^m (y_j - \mu^h)^{-1} + \sum_{j=1}^m R_j (y_j - \mu^h)^{-1} - 2\lambda^{(h+1)} \sum_{j=1}^m R_j \frac{\lambda^{(h+1)}(y_j - \mu^h)^2 + 1}{\lambda^{(h+1)}(y_j - \mu^h)} \right]}{2m\lambda^{(h+1)}} \end{aligned}$$

The value $(\lambda^{(h+1)}, \mu^{(h+1)})$ is then used as a new value of (λ, μ) in the succeeding iteration. The MLEs of (λ, μ) can be obtained by repeating the E-step and M-step until convergence.

2.4. The Newton-Raphson Algorithm

We will directly extend the argument for deriving the Newton-Raphson algorithm for optimization in one dimension to two-dimensional problems as discussed by Devore and Berk [8] giving the two-parameter Newton-Raphson method as:

$$\boldsymbol{\theta}^{(h+1)} = \boldsymbol{\theta}^h + J^{-1}(\boldsymbol{\theta}^h) S(\boldsymbol{\theta}^h) \quad (16)$$

Where $J(\boldsymbol{\theta})$ is the Hessian matrix (a matrix with (i, j) entry equal to the second derivatives with respect to θ^j and θ^i)

The M-step of the (h+1)th iteration of the EM algorithm is completed by substituting the above conditional expectations on to equations (9) and (10) as follows:

$$\frac{\partial l_c(W; \mu, \lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{j=1}^m (y_j - \mu)^2 - \sum_{j=1}^m R_j \frac{2\lambda(y_j - \mu)^2 + 3}{2\lambda} = 0$$

Hence,

$$\lambda^{(h+1)} = \frac{n}{\sum_{j=1}^m (y_j - \mu^h)^2 + \sum_{j=1}^m R_j \frac{2\lambda^h (y_j - \mu^h)^2 + 3}{2\lambda^h}} \quad (14)$$

$\lambda^{(h+1)}$ is the estimate of λ at the (h+1)th iteration of the EM algorithm.

Once $\lambda^{(h+1)}$ is obtained, $\mu^{(h+1)}$ is obtained as follows;

and $S(\boldsymbol{\theta})$ is the score function (a vector of derivatives).

From equation (5), $J(\boldsymbol{\theta})$ and $S(\boldsymbol{\theta})$ are obtained as:

$$S(\boldsymbol{\theta}) = \begin{pmatrix} \frac{m}{\lambda} - \sum_{j=1}^m (R_j + 1)(x_j - \mu)^2 \\ -\sum_{j=1}^m (x_j - \mu)^{-1} + 2\lambda \sum_{j=1}^m (R_j + 1)(x_j - \mu) \end{pmatrix} \quad (17)$$

$$J(\boldsymbol{\theta}^h) = \begin{pmatrix} -\frac{m}{\lambda^{2h}} & 2\sum_{j=1}^m (R_j + 1)(x_j - \mu^h) \\ 2\sum_{j=1}^m (R_j + 1)(x_j - \mu^h) & -\sum_{j=1}^m (x_j - \mu^h)^{-2} - 2\lambda^h n \end{pmatrix} \quad (18)$$

Hence, equation (16) becomes

$$\begin{pmatrix} \lambda \\ \mu \end{pmatrix}^{h+1} = \begin{pmatrix} \lambda^h \\ \mu^h \end{pmatrix} + \begin{bmatrix} -\frac{m}{\lambda^{2h}} & 2 \sum_{j=1}^m (R_j + 1)(x_j - \mu^h) \\ 2 \sum_{j=1}^m (R_j + 1)(x_j - \mu^h) & -\sum_{j=1}^m (x_j - \mu^h)^{-2} - 2\lambda^h n \end{bmatrix}^{-1} \begin{pmatrix} \frac{m}{\lambda^h} - \sum_{j=1}^m (R_j + 1)(x_j - \mu^h)^2 \\ -\sum_{j=1}^m (x_j - \mu^h)^{-1} + 2\lambda^h \sum_{j=1}^m (R_j + 1)(x_j - \mu^h) \end{pmatrix} \quad (19)$$

The procedure is reiterated until there is no significant difference between $\begin{pmatrix} \lambda \\ \mu \end{pmatrix}^{h+1}$ and $\begin{pmatrix} \lambda^h \\ \mu^h \end{pmatrix}$

2.5. Approximate Interval Estimation

The approximate asymptotic variances of the shape and the scale parameters and the confidence intervals are obtained as follows:

Let $X = X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}$ with $X_{1:m:n} < X_{2:m:n} < \dots < X_{m:m:n}$ denoting a progressive type-II right-censored sample from a population with density and distribution functions given in equations (1) and (2), respectively.

The Fisher information matrix $I(\lambda, \mu)$ is then obtained by taking the expectation of minus the second derivatives of equation (3.6) with respect to θ^j and θ^i . Cox and Hinkley [6], established that if (μ, λ) belongs to an open interval of a real line, then some of Cramer-Rao regularity conditions are satisfied, and as the sample size increases, the distribution of the MLE tends to be approximately bivariate normally distributed with mean (λ, μ) and covariance matrix $I^{-1}(\lambda, \mu)$. In practice, $I^{-1}(\lambda, \mu)$ is estimated by $I^{-1}(\hat{\lambda}, \hat{\mu})$.

The distribution of the MLEs is denoted by:

$(\hat{\lambda}, \hat{\mu}) \rightarrow N_2((\lambda, \mu), I_0^{-1}(\hat{\lambda}, \hat{\mu}))$, where $I_0(\hat{\lambda}, \hat{\mu})$ is the observed information matrix given by

$$I_0(\hat{\lambda}, \hat{\mu}) = \begin{bmatrix} \text{var}(\hat{\lambda}) & \text{cov}(\hat{\lambda}, \hat{\mu}) \\ \text{cov}(\hat{\lambda}, \hat{\mu}) & \text{var}(\hat{\mu}) \end{bmatrix} = \begin{bmatrix} -E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \lambda^2} \right\} & -E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \lambda \partial \mu} \right\} \\ -E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \lambda \partial \mu} \right\} & -E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \mu^2} \right\} \end{bmatrix}_{(\hat{\lambda}, \hat{\mu})}^{-1} \quad (20)$$

Using equation (3.6), the elements of the observed information matrix can be obtained as follows:

$$-E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \lambda^2} \right\} = \frac{m}{\lambda^2} \quad (21)$$

$$-E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \lambda \partial \mu} \right\} = 2m \frac{e^{-\lambda \mu^2}}{\lambda \mu} [\lambda \mu^2 + 1] \quad (22)$$

$$-E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \mu^2} \right\} = 2\lambda n + \int_0^\infty \frac{2\lambda e^{-\lambda(x_j - \mu)^2}}{(x_j - \mu)} dx_j = \infty$$

Since,

$$\int_0^\infty \frac{e^{-\lambda(x_j - \mu)^2}}{(x_j - \mu)} dx_j = \infty$$

The exact asymptotic variance of $\hat{\mu}$ cannot be obtained in explicit form. We rely on the results of Dey, Dey, and Kundu [11] who applied Corollary of Theorem 3 of Smith [28], to approximate the asymptotic variance of $\hat{\mu}$ by $V(\hat{\mu} - \mu)$ using the inverse of the observed information as:

$$V(\hat{\mu} - \mu) \approx \frac{1}{2\mu + \frac{1}{n} \sum_{j=1}^m (x_j - \mu)^{-2}} \quad (23)$$

Hence,

$$-E \left\{ \frac{\partial^2 l(X; \mu, \lambda)}{\partial \mu^2} \right\} \approx \frac{1}{2\hat{\mu} + \frac{1}{n} \sum_{j=1}^m (x_j - \hat{\mu})^{-2}} \quad (24)$$

Using equation (20), a $(1-\alpha)100\%$ approximate confidence intervals for λ and μ are obtained as

$$\hat{\lambda} \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\lambda})}, \text{ and } \hat{\mu} \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\mu})} \quad (25)$$

Respectively, where, $z_{\frac{\alpha}{2}}$ is the $z_{\frac{\alpha}{2}}$ -th percentile point of the standard normal distribution.

3. Results and Discussions

In this section, a simulation study is performed to compare the performance of MLEs of the two-parameter Rayleigh distribution obtained using the NR method and the EM algorithm based on progressive type-II censored samples.

Progressive type-II is right censored samples from two-parameter Rayleigh distribution were generated using the algorithms discussed in [2-3].

In comparing the performance of the MLEs, four measures considered were the root mean squared error (RMSE) and the 95% approximate confidence width of MLEs. Suppose $\hat{\theta}_k$ is the MLE of θ for the i^{th} replication of the k^{th} algorithm method simulated, then the Bias and RMSE of $\hat{\theta}_k$ are computed as follows:

$$i. Bias(\hat{\theta}_k) = \frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}_k - \theta), \text{ where } \theta = (\lambda, \mu) \text{ and } k = EM \text{ or } NR$$

$$ii. RMSE(\hat{\theta}_k) = \sqrt{\frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}_k - \theta)^2}$$

In this paper, samples of sizes 20, 30, 40, 50, and 70 were used and the censoring schemes considered are given in Table 1 and 2 below.

Table 1. Censoring Schemes $R = (r_1, r_2, \dots, r_m)$ for $\lambda = 0.5, 1$, and $\mu = 0.3, 0.6, 1$.

n	m	$\sum_{j=1}^m R_j$	Censoring Schemes (i, ii, iii, iv)
20	15	5	1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1
	18	2	1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1
30	20	10	1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1
	25	5	1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1

Table 2. Censoring Schemes $R = (r_1, r_2, \dots, r_m)$ for $\theta = (\lambda = 1, \mu = 0.3)$ with fixed number of failures.

n	m	$\sum_{j=1}^m R_j$	Censoring Schemes
20		2	0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1
30		12	0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
40	18	22	1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2
50		32	2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 2, 2, 1, 2
70		52	3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 2

It is imperative to note that, in all the above censoring schemes no restriction has been imposed on the maximum number of iterations and convergence is assumed to occur when the absolute differences between successive estimates are less than 0.0001.

Table 3. The RMSE and the width of 95% approximate confidence intervals of the MLEs for the parameters of two-parameter Rayleigh distribution under progressive type-II censoring for the EM and NR algorithms using scheme 1.

λ	μ	RMSE of ($\hat{\lambda}$)		RMSE of ($\hat{\mu}$)		Width of ($\hat{\lambda}$)		Width of ($\hat{\mu}$)	
		EM	NR	EM	NR	EM	NR	EM	NR
0.5	0.3	0.03373	0.14692	0.04120	0.10349	2.21181	2.86311	3.14514	4.25909
	0.6	0.03820	0.14833	0.09925	0.13286	2.19209	2.52055	3.11384	3.57085
	1	0.03912	0.16444	0.21569	0.13532	2.16135	2.22008	3.05599	2.94556
1	0.3	0.10479	0.33258	0.07826	0.10511	2.39961	2.52933	2.86799	2.94822
	0.6	0.12312	0.34156	0.11592	0.18279	2.39146	2.47646	2.84873	2.82692
	1	0.14261	0.36837	0.12607	0.19510	2.32661	2.46077	2.71757	2.71373

Table 4. The RMSE and the width of 95% approximate confidence intervals of the MLEs for the parameters of two-parameter Rayleigh distribution under progressive type-II censoring by the EM and NR algorithms using scheme 2.

λ	μ	RMSE of ($\hat{\lambda}$)		RMSE of ($\hat{\mu}$)		Width of ($\hat{\lambda}$)		Width of ($\hat{\mu}$)	
		EM	NR	EM	NR	EM	NR	EM	NR
0.5	0.3	0.01329	0.06867	0.03287	0.05661	2.23332	2.31569	3.18513	3.23865
	0.6	0.03779	0.08229	0.12121	0.11653	1.75011	1.84819	2.22171	2.29613
	1	0.02631	0.12001	0.17521	0.21878	2.05572	2.12799	2.82191	2.92994
1	0.3	0.06543	0.27941	0.05217	0.13848	2.29415	2.37595	2.74638	2.77699
	0.6	0.08114	0.30678	0.06869	0.26369	2.25967	2.33232	2.70264	2.73948
	1	0.09298	0.33543	0.07928	0.39274	2.21544	2.36355	2.61828	2.67069

Table 5. The RMSE and the width of 95% approximate confidence intervals of the MLEs for the parameters of two-parameter Rayleigh distribution under progressive type-II censoring by the EM and NR algorithms using scheme 3.

λ	μ	RMSE of ($\hat{\lambda}$)		RMSE of ($\hat{\mu}$)		Width of ($\hat{\lambda}$)		Width of ($\hat{\mu}$)	
		EM	NR	EM	NR	EM	NR	EM	NR
0.5	0.3	0.12187	0.31759	0.15321	0.25552	2.25621	2.72891	3.33628	3.62619
	0.6	0.15032	0.32572	0.20077	0.26295	2.18092	2.40089	3.23173	3.37592
	1	0.21156	0.34757	0.25555	0.30160	2.05645	2.28516	3.06527	3.13232
1	0.3	0.23239	0.34905	0.16239	0.14971	2.32438	2.32551	3.00063	2.94679
	0.6	0.26722	0.39766	0.19507	0.19629	2.29471	2.29563	2.95629	2.90378
	1	0.29689	0.44535	0.23105	0.27945	2.22596	2.71548	2.84875	2.87882

Table 6. The RMSE and the width of 95% approximate confidence intervals of the MLEs for the parameters of two-parameter Rayleigh distribution under progressive type-II censoring by the EM and NR algorithms using scheme 4.

λ	μ	RMSE of ($\hat{\lambda}$)		RMSE of ($\hat{\mu}$)		Width of ($\hat{\lambda}$)		Width of ($\hat{\mu}$)	
		EM	NR	EM	NR	EM	NR	EM	NR
0.5	0.3	0.07763	0.18979	0.12649	0.18551	2.28813	2.58962	3.32442	3.44308
	0.6	0.09508	0.23211	0.16778	0.21962	2.13476	2.30886	3.20137	3.29517
	1	0.14153	0.27896	0.22908	0.23665	2.03380	2.20992	3.04246	3.05544
1	0.3	0.16660	0.32116	0.13359	0.12908	2.30450	2.31269	2.98479	2.93725
	0.6	0.23003	0.33009	0.16648	0.15105	2.26412	2.29418	2.92940	2.86821
	1	0.23649	0.37962	0.20816	0.21444	2.20525	2.28209	2.82867	2.83401

Table 7. The RMSE and the width of 95% approximate confidence intervals of the MLEs for the parameters of two-parameter Rayleigh distribution under progressive type II censoring by the EM and NR algorithms with fixed number of failures when $\theta = (\lambda=1, \mu=0.3)$.

Sample		RMSE of ($\hat{\lambda}$)		RMSE of ($\hat{\mu}$)		Width of ($\hat{\lambda}$)		Width of ($\hat{\mu}$)	
n	m	EM	NR	EM	NR	EM	NR	EM	NR
20		0.27922	0.53333	0.25149	0.35995	2.42022	2.62368	3.04994	3.31897
30		0.26571	0.49180	0.15816	0.27571	2.28307	2.46398	2.92599	3.20908
40	18	0.22567	0.43848	0.14770	0.22269	2.23012	2.38406	2.91659	3.18989
50		0.20262	0.35030	0.13070	0.17341	2.12142	2.21689	2.83493	2.92433
70		0.16258	0.26483	0.11252	0.11146	2.05005	2.08809	2.78633	2.78706

A Summary of Results from Tables 3-7 is Provided Below

- i. The MLEs realized using the EM algorithm have lower levels of RMSE compared to those obtained by the NR method in nearly all combinations of simulation conditions.
- ii. The widths of 95% approximated confidence intervals of parameters λ and μ obtained using the EM algorithm tends to be lesser compared to those obtained by NR method in nearly all combinations of simulation conditions. According to Gulhar et al., a smaller width is better because it captures the true parameter value (CV) within a small span and the results are more accurate and precise [13].
- iii. For a fixed sample size n (e.g. $n=30$), we noted that as the number of failures (m) increases (i.e. from 20 to 25), the RMSE and widths of confidence intervals of MLEs obtained using both the EM and NR algorithms decreases. For RMSE and widths of the confidence intervals (compare Table 3 and 4, and Table 5 and 6). This implies that the performance of MLEs becomes better.
- iv. When the number of failures m is fixed, we observed that as the sample size n increases the RMSE, and the

widths of 95% approximate confidence intervals of MLEs obtained using both the EM and NR algorithms decreases (see Table 7). This indicates that the MLEs are consistent in nature.

- v. When the value of λ is fixed, we noted that as the value of μ increases, the RMSE for all the estimates increases, which indicates the consistency of the estimators.
- vi. Additionally, if $m=0$ there is no censoring, under this condition zero samples are generated hence it is not possible to obtain the corresponding MLEs. On the other hand, if $R_1 = R_2 = \dots = R_m = 0$, then $n=m$ (complete sample situation), under this condition estimates are extremely biased.

4. Example Using Real-Life Data

Now consider a real-life data set to illustrate how maximum likelihood estimation using the NR method and the EM algorithm for the two-parameter Rayleigh distribution works in practice. We have utilized progressive type-II censoring to analyze a real data representing the survival times (in years) of 46 patients given chemotherapy treatment

as discussed in [5]. The discussion indicated that the Rayleigh Distribution is acceptance for this data set (provides a good fit). The data set is given as follows:

0.047, 0.115, 0.121, 0.132, 0.164, 0.197, 0.203, 0.260, 0.282, 0.296, 0.334, 0.395, 0.458, 0.466, 0.501, 0.507, 0.529, 0.534, 0.540, 0.570, 0.641, 0.644, 0.696,

0.841, 0.863, 1.099, 1.219, 1.271, 1.326, 1.447, 1.485, 1.553, 1.581, 1.589, 2.178, 2.343, 2.416, 2.444, 2.825, 2.830, 3.578, 3.658, 3.743, 3.978, 4.003, 4.033

From the above data, progressive type-II censored samples were generated with $m=20, 30,$ and 40 as follows:

Table 8. Censoring scheme $R = (r_1, r_2, \dots, r_m)$ and progressive type-II censored samples for different values of m using real-life data sets.

n	m	$\sum_{j=1}^m R_j$	Censoring Schemes	Progressive type-II Censored samples from the original data
46	20	26	(3,2,1,0 ¹⁴ ,5,5,10)	0.047, 0.164, 0.260, 0.296, 0.334, 0.395, 0.458, 0.466, 0.501, 0.507, 0.529, 0.534, 0.540, 0.570, 0.641, 0.644, 0.696, 0.841, 1.447, 2.343
	30	16	(0 ²⁶ ,3,3,5,5)	0.047, 0.115, 0.121, 0.132, 0.164, 0.197, 0.203, 0.260, 0.282, 0.296, 0.334, 0.395, 0.458, 0.466, 0.501, 0.507, 0.529, 0.534, 0.540, 0.570, 0.641, 0.644, 0.696, 0.841, 0.863, 1.099, 1.219, 1.485, 2.178, 3.578
	40	6	(3,3,0 ³⁸)	0.047, 0.164, 0.282, 0.296, 0.334, 0.395, 0.458, 0.466, 0.501, 0.507, 0.529, 0.534, 0.540, 0.570, 0.641, 0.644, 0.696, 0.841, 0.863, 1.099, 1.219, 1.271, 1.326, 1.447, 1.485, 1.553, 1.581, 1.589, 2.178, 2.343, 2.416, 2.444, 2.825, 2.830, 3.578, 3.658, 3.743, 3.978, 4.003, 4.033

Table 9. The MLEs and the corresponding widths of 95% approximate confidence intervals for the unknown parameters of TR distribution for different values of m .

Sample	$\hat{\lambda}$		$\hat{\mu}$		Width of ($\hat{\lambda}$)		Width of ($\hat{\mu}$)	
	EM	NR	EM	NR	EM	NR	EM	NR
20	0.04869	0.42304	0.46229	0.58781	0.02814	0.24451	0.07393	0.35723
30	0.01932	0.09094	0.39005	0.68766	0.01117	0.05256	0.12923	0.21104
40	0.00646	0.01108	0.19252	0.57345	0.00373	0.00641	0.64586	0.08998

From the table above, it is observed that:

- i. The MLEs obtained using the EM algorithm have narrower widths of confidence intervals compared to those obtained using NR method except for $\hat{\mu}$ when $m=40$. A smaller width is better because it captures the true parameter value (CV) within a small span and the results are more accurate and precise.
- ii. For both methods, the MLEs and the width of 95% approximate confidence intervals decrease as the number of failures increases (i.e., from 20 to 40) for nearly all the values of m , which indicates the consistency of the estimators.

5. Conclusions

In this study, the problem of estimation of the MLEs for the parameters of the two-parameter Rayleigh distribution based on generated progressive type-II censored samples was addressed. In particular, the MLEs were derived using the NR and the EM algorithms. Approximate asymptotic variances of the MLEs were also derived and used to construct approximate confidence intervals of the parameters.

The simulation results clearly show that the MLEs obtained using the EM algorithm have lower levels of RMSE and narrower widths of the corresponding confidence intervals compared to those obtained using the NR algorithm. However, the NR method may yield better estimates

especially when $\left(\frac{n}{m}\right)$ is greater than 70%. This shows that

both the EM and NR algorithms can be used in estimation problem, but we can conclude that the EM algorithm is

highly recommended as it provides better estimates. Al-Zahrani and Gindwan [1] and Helu, Samawi, & Raqab [14] obtained similar simulation results.

References

- [1] Al-Zahrani, Bander M., and Maha S. Gindwan. "Estimating the parameter of the Lindley distribution under progressive type-II censored data." *Electronic Journal of Applied Statistical Analysis* 8.1 (2015): 100-111.
- [2] Balakrishnan, Narayanaswamy, and Rita Aggarwala. *Progressive Censoring: theory, methods, and applications*. Springer Science & Business Media, 2000.
- [3] Balakrishnan, N., and R. A. Sandhu. "A simple simulational algorithm for generating progressive Type-II censored samples." *The American Statistician* 49.2 (1995): 229-230.
- [4] Balakrishnan, N. "Progressive censoring methodology: an appraisal." *Test* 16.2 (2007): 211-259.
- [5] Bekker, A., J. J. J. Roux, and P. J. Mosteit. "A generalization of the compound Rayleigh distribution: using a Bayesian method on cancer survival times." *Communications in Statistics-Theory and Methods* 29.7 (2000): 1419-1433.
- [6] Cox, D. R., and D. V. Hinkley. "Theoretical Statistics Chapman and Hall, London." *See Also* (1974).
- [7] Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)* (1977): 1-38.
- [8] Devore, Jay L., and Kenneth N. Berk. *Modern mathematical statistics with applications*. Cengage Learning, 2007.

- [9] Dey, Sanku. "Comparison of Bayes Estimators of the parameter and reliability function for Rayleigh distribution under different loss functions." *Malaysian Journal of Mathematical Sciences* 3 (2009): 247-264.
- [10] Dey, Sanku, and M. K. Das. "A note on prediction interval for a Rayleigh distribution: Bayesian approach." *American Journal of Mathematical and Management Sciences* 27.1-2 (2007): 43-48.
- [11] Dey, Sanku, Tanujit Dey, and Debasis Kundu. "Two-parameter Rayleigh distribution: different methods of estimation." *American Journal of Mathematical and Management Sciences* 33.1 (2014): 55-74.
- [12] Dey, Tanujit, Sanku Dey, and Debasis Kundu. "On Progressively Type-II Censored Two-parameter Rayleigh Distribution." *Communications in Statistics-Simulation and Computation* 45.2 (2016): 438-455.
- [13] Gulhar, Monika, et al. "A comparison of some confidence intervals for estimating the population coefficient of variation: a simulation study." *Sort: Statistics and Operations Research Transactions* 36.1 (2012): 45-68.
- [14] Helu, Amal, Hani Samawi, and Mohammad Z. Raqab. "Estimation on Lomax progressive censoring using the EM algorithm." *Journal of Statistical Computation and Simulation* 85.5 (2015): 1035-1052.
- [15] Howlader, H. A., and A. Hossain. "On Bayesian estimation and prediction from Rayleigh based on type II censored data." *Communications in Statistics-Theory and Methods* 24.9 (1995): 2251-2259.
- [16] Johnson, Norman L., Samuel Kotz, and N. Balakrishnan. *Continuous multivariate distributions, volume 1, models and applications*. Vol. 59. New York: John Wiley & Sons, 2002.
- [17] Khan, Hafiz MR, Serge B. Provost, and Ashima Singh. "Predictive inference from a two-parameter Rayleigh life model given a doubly censored sample." *Communications in Statistics-Theory, and Methods* 39.7 (2010): 1237-1246.
- [18] Kundu, Debasis, and Arabin Kumar Dey. "Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm." *Computational Statistics & Data Analysis* 53.4 (2009): 956-965.
- [19] Lio, Y. L., Ding-Geng Chen, and Tzong-Ru Tsai. "Parameter Estimations for Generalized Rayleigh Distribution under Progressively Type-I Interval Censored Data." *Open Journal of Statistics* 1.02 (2011): 46.
- [20] Top of Form McLachlan, Geoffrey J, and Thriyambakam Krishnan. *The EM Algorithm and Extensions*. New York: J. Wiley, 1997. Print.
- [21] Bottom of Form Lee, Jason, Yuekai Sun, and Michael Saunders. "Proximal Newton-type methods for convex optimization." *Advances in Neural Information Processing Systems*. 2012.
- [22] Ng, H. K. T., P. S. Chan, and N. Balakrishnan. "Estimation of parameters from progressively censored data using EM algorithm." *Computational Statistics & Data Analysis* 39.4 (2002): 371-386.
- [23] Ortega, James M., and Werner C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*. Vol. 30. Siam, 1970.
- [24] Pradhan, Biswabrata, and Debasis Kundu. "Bayes estimation and prediction of the two-parameter gamma distribution." *Journal of Statistical Computation and Simulation* 81.9 (2011): 1187-1198.
- [25] Rayleigh, Lord. "On the capillary phenomena of jets." *Proc. R. Soc. London*. Vol. 29. No. 196-199. 1879.
- [26] Surles, J. G., and W. J. Padgett. "Inference for P (Y < X) in the Burr type X model." *Communications in Statistics-Theory and Methods* 7 (1998): 225-238.
- [27] Seo, Eun-Hyung, and Suk-Bok Kang. "AMLEs for Rayleigh distribution based on progressive Type-II censored data." *Communications for Statistical Applications and Methods* 14.2 (2007): 329-344.
- [28] Smith, Richard L. "Maximum likelihood estimation in a class of nonregular cases." *Biometrika* 72.1 (1985): 67-90.
- [29] Watanabe, Michiko, and Kazunori Yamaguchi, eds. *The EM algorithm and related statistical models*. CRC Press, 2003.