TWMS J. App. and Eng. Math. V.15, N.8, 2025, pp. 2019-2030

STOCHASTIC EXPECTATION MAXIMIZATION ALGORITHM FOR EXPONENTIAL-POISSON DISTRIBUTION UNDER TYPE-I PROGRESSIVE INTERVAL CENSORING

A. MOHAMMADI[1,*], A. SHADROKH[2], M. YARMOHAMMADI[3], §

Abstract. The Exponential-Poisson (EP) distribution is generated by combining the Exponential distribution with a zero truncated Poisson distribution as a model for lifetime data with decreasing failure rate. This paper deals with the problem of estimating unknown parameters of the Exponential-Poisson distribution as a lifetime model when samples are observed under progressive type-I interval censoring. We employ the Newton-Raphson (NR), classical expectation maximization (EM) and stochastic expectation maximization (SEM) algorithms to find the maximum likelihood estimates for the unknown parameters. The performance of the proposed SEM estimators are illustrated by a Monte Carlo simulation study and used for a real data set. Our simulation showed that the performance of SEM algorithm is quite satisfactory on the basis of mean square error and by increasing the sample size, the efficiency is also increases.

Keywords: Maximum likelihood estimation, Progressive Interval Censoring, EM and SEM algorithm, Exponential Poisson.

AMS Subject Classification: 90B25, 62B10, 65D30, 65C05.

1. Introduction

Life testing and reliability experiments are effective ways for assessing live/failure life under specified operating circumstances. Such tests are carried out on identical units before (and throughout) the product's release, and the failure times observed are documented. After that, the failure time is further examined in order to evaluate and forecast product attributes. Due to time limits, financial expenses, or inadvertent breakdown, such trials are sometimes discontinued before all life-tested components fail. As a result, the data seen in such circumstances may be incomplete. Type-I and type-II censoring systems are two regularly used classical censorship techniques in which the experiment stops after a certain period and number of failures, respectively. However, neither of these filtering systems

[1] Department of Statistics, Payame Noor University (PNU), P.O.Box 19395-4697, Tehran, Iran.
e-mail: ahmohammadi@pnu.ac.ir; ORCID: https://orcid.org/0000-0001-6100-9402.
[2] Department of Statistics, Payame Noor University (PNU), P.O.Box 19395-4697, Tehran, Iran.
e-mail: A.shadrokh@pnu.ac.ir; ORCID: https://orcid.org/0000-0003-0069-8165.
[3] Department of Statistics, Payame Noor University (PNU), P.O.Box 19395-4697, Tehran, Iran.
e-mail: Masyar@pnu.ac.ir; ORCID: https://orcid.org/0000-0001-7038-7106.
* Corresponding author.
§ Manuscript received: June 18, 2024; accepted: November 30, 2024.

allows the researcher to remove the live unit from the experiment before the end time. The concept of progressive censoring, which involves removing units between tests, has lately gained popularity. Furthermore, it has been shown that in many real-world scenarios, experimenters are unable to continually monitor endurance tests in order to detect accurate failure lives. Medical and clinical research, for example, may not offer correct survival times for people diagnosed with specific medicines. Failure lifetimes are frequently seen at intervals called interval censoring in such scenarios. This censor, however, does not permit the removal of units between experiments. [2] pioneered progressive type-I interval censoring by combining the concepts of type-I, progressive, and interval censoring. Lifetime statistics under this censorship may be shown as follows. Assume that at time $t_0 = 0$, a random sample of n units is subjected to a life test experiment, and that at inspection time t_i, the total number of observed failures in the interval $(t_{i-1}, t_i]$ is D_i. Further assume that R_i be the total number of units removed from the experiment at time t_i, $i = 1, 2, ..., m$. Here the number of surviving units at time t_i, say Y_i is a random variable, and therefore $R_i \leq Y_i$. The general practice to determine R_i is by a prescribed percentage q_i of the remaining Y_i surviving units or by a pre-specified non negative integers in which the actually observed Robs $i = \min\left(R_i, \text{number of surviving units at inspection time } t_i\right), i = 1, 2, \ldots, m-1$ and Robs m = number of surviving units at inspection time t_m. Data observed under this censoring can be represented as $(D_i, R_i, t_i)_{i=1}^{m}$. Among many, lifetime data set based on 112 patients with plasma cell myeloma (see [5] and references cited there in), data set of HMO-HIV study, see [18], and survival times from surgery of a group of 374 patients, see [13] have been statistically analyzed using progressive type-I interval censoring. Further statistical inference under this censoring for Exponentiated Weibull family has been discussed by [3], for Weibull distribution by [12], for generalized Rayleigh distribution by [11], for Gamma distribution by [21], for generalized Exponential distribution by [6] and [14], for Weibull and generalized Exponential distributions by Lin and Lio [15], for lognormal distribution by [16], and for Burr XII distribution by [1].

In this paper we discuss the Exponential Poisson distribution under progressive type-I interval censoring. The Exponential Poisson (EP) distribution is generated by combining the Exponential distribution with a zero truncated Poisson distribution as a model for lifetime data with decreasing failure rate [10]. Let $T_1, T_2, ..., T_W$ be a random sample from distribution with density $f(t; \beta) = \beta e^{-t\beta}, t, \beta \in \mathbb{R}_+$ and W is a Poisson distributed variable truncated at zero with probability mass function

$P(w, \lambda) = e^{-\lambda}\lambda^{w}\Gamma^{-1}(w+1)\left(1-e^{-\lambda}\right)^{-\lambda}, w \in \mathbb{N}, \lambda \in \mathbb{R}_+$ where $\Gamma(.)$ is the Gamma function and W and T_is are independent. Let us define $X = \min(T_1, T_2, ..., T_W)$. Then, $f(x|w; \beta) = \beta w e^{-\beta w x}$, and marginal probability density function of X is

$$f\left(x; \boldsymbol{\theta}\right) = \frac{\lambda\beta e^{-\lambda-\beta x+\lambda \exp\left(-\beta x\right)}}{\left(1-e^{-\lambda}\right)}, \quad x, \beta, \lambda \in \mathbb{R}_+ \tag{1}$$

where $\boldsymbol{\theta} = (\lambda, \beta)$. (It should be noted that [4] derived another lifetime distribution as Poisson-Exponential (PE) distribution by taking $X = \max(T_1, T_2, ..., T_W)$)

It has found widespread use in economics, medicine, business, and actuarial sciences. Zero values are not permitted in the distributions often used in lifetime experiments such as Gamma, Weibull, and Log-normal, to name a few. A random variable X with probability density function 1 and cumulative distribution function

$$F(x; \boldsymbol{\theta}) = (e^{\lambda \exp(-\beta x)} - e^{\lambda})(1-e^{\lambda})^{-1}. \tag{2}$$

has two parameter EP distribution and is denoted by $EP(\lambda, \beta)$. The parameters λ and β are the shape and scale parameters, respectively. The EP distribution (1) can take value on

0. This allows this distribution to become a viable option for usage in issues involving zero values. The occurrence of zero values in environmental studies is typical during the study of precipitation, particularly during dry seasons. Such behavior is also found in reliability studies, where instantaneous failures can occur owing to lower quality or difficulties during component manufacture. For a detail account of this distribution and its properties one may refer to [10]. This paper focus on the problem of estimating the unknown parameters of the EP distribution under classical approach when data are observed using progressive type-I interval censoring. Section 2 deals with MLEs. Section 3 discusses Newton-Raphson method. Section 4 we obtain estimators using EM and SEM algorithms. Section 5 analyses real data set and simulation study and conclusion is presented in the last section.

2. MLEs

The goal of this section is to find maximum likelihood estimators (MLEs) for the unknown parameters of the EP distribution when data is observed with progressive type-I interval censoring. Suppose that n number of units whose lifetimes are identically distributed random variables with pdf and cdf as defined in 1 and 2, used for life test experiment at time $t_0 = 0$. Further assume that $(D_i, R_i, t_i)_{i=1}^m$ are observed data under progressive type-I interval censoring with the prescribed m inspection times t_i. Then the associated likelihood function of (λ, β) can be written as

$$\mathcal{L}(\boldsymbol{\theta}) \propto \Pi_{i=1}^m \left[F(t_i;\lambda,\beta) - F(t_{i-1};\lambda,\beta)\right]^{D_i} \left[1 - F(t_i;\lambda,\beta)\right]^{R_i}. \tag{3}$$

Now using the log-likelihood function

$$\begin{aligned} l(\lambda,\beta) &= \ln\mathcal{L}(\lambda,\beta) \\ &= -\ln(1-e^{\lambda})\sum_{i=1}^m (D_i+R_i) + \sum_{i=1}^m D_i \ln(e^{\lambda\exp(-\beta t_i)} - e^{\lambda\exp(-\beta t_{i-1})}) \\ &\quad + \sum_{i=1}^m R_i \ln(1 - e^{\lambda\exp(-\beta t_{i-1})}). \end{aligned}$$

The maximum likelihood estimates of (λ, β) are the solution to the following equations

$$\begin{aligned} \frac{\partial l(\lambda,\beta)}{\partial\lambda} &= \frac{e^{\lambda}}{1-e^{\lambda}}\sum_{i=1}^m (D_i+R_i) + \sum_{i=1}^m D_i \left(\frac{e^{-\beta t_i}e^{\lambda\exp(-\beta t_i)} - e^{-\beta t_{i-1}}e^{\lambda\exp(-\beta t_{i-1})}}{e^{\lambda\exp(-\beta t_i)} - e^{\lambda\exp(-\beta t_{i-1})}}\right) \\ &\quad + \sum_{i=1}^m R_i \left(\frac{-e^{-\beta t_i}e^{\lambda\exp(-\beta t_i)}}{1-e^{\lambda\exp(-\beta t_i)}}\right) = 0, \end{aligned} \tag{4}$$

$$\begin{aligned} \frac{\partial l(\lambda,\beta)}{\partial\beta} &= \sum_{i=1}^m D_i \left(\frac{-\lambda t_i e^{-\beta t_i}e^{\lambda\exp(-\beta t_i)} + \lambda t_{i-1}e^{-\beta t_{i-1}}e^{\lambda\exp(-\beta t_{i-1})}}{e^{\lambda\exp(-\beta t_i)} - e^{\lambda\exp(-\beta t_{i-1})}}\right) \\ &\quad + \sum_{i=1}^m R_i \left(\frac{\lambda t_i e^{-\beta t_i}e^{\lambda\exp(-\beta t_i)}}{1-e^{\lambda\exp(-\beta t_i)}}\right) = 0. \end{aligned} \tag{5}$$

However, since the preceding equations have no closed-form solution, numerical methods must be applied. Newton-Raphson is the most often utilized approach in the current literature (NR). The primary disadvantage of this technique is that it needs the second derivative of the logarithmic probability function for each iteration, which can be computationally laborious owing to the complicated structure of the probability function. Alternatively, the following equations may be solved using built-in numerical packages in various computer languages, however the results may change significantly. [8] presented an Expectation-Maximization (EM) technique, which various authors have utilized to achieve maximum likelihood estimates. This approach is far superior than the NR method, especially when the data is incomplete and collected under some censoring scheme. The EM method is then discussed in order to get maximum likelihood estimators of (λ, β).

3. Newton–Raphson method

Approximation to the MLE of unknown parameters λ and β can be obtained using the Newton–Raphson method. Let $\hat{\boldsymbol{\theta}}^0 = (\hat{\lambda}^0, \hat{\beta}^0)^\top$ as an initial estimate in the Newton–Raphson procedure, the next estimate is

$$\hat{\boldsymbol{\theta}}_1 = \hat{\boldsymbol{\theta}}_0 - \left[\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mid_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_0}\right]^\top \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \mid_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_0}\right]^{-1}$$

where

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial l(\lambda,\beta)}{\partial \lambda} \\ \frac{\partial l(\lambda,\beta)}{\partial \beta} \end{pmatrix} \quad \text{and} \quad \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \begin{pmatrix} \frac{\partial^2 l(\lambda,\beta)}{\partial \lambda^2} & \frac{\partial^2 l(\lambda,\beta)}{\partial \lambda \beta} \\ \frac{\partial^2 l(\lambda,\beta)}{\partial \lambda \beta} & \frac{\partial^2 l(\lambda,\beta)}{\partial \beta^2} \end{pmatrix}$$

with $(\partial l(\lambda, \beta)/\partial \lambda)$ and $(\partial l(\lambda, \beta)/\partial \beta)$ given in equations (4) and (5) and

$$\begin{aligned} \frac{\partial^2 l(\lambda, \beta)}{\partial \lambda^2} = \sum_{i=1}^{m} D_i & \frac{\left\{ \begin{array}{c} \left[\exp(-2\beta t_i) e^{\lambda \exp(-\beta t_i)} - \exp(-2\beta t_{i-1}) e^{\lambda \exp(-\beta t_{i-1})}\right] \\ \times \left[e^{\lambda \exp(-\beta t_i)} - e^{\lambda \exp(-\beta t_{i-1})}\right] \\ - \left[\exp(-\beta t_i) e^{\lambda \exp(-\beta t_i)} - \exp(-\beta t_{i-1}) e^{\lambda \exp(-\beta t_{i-1})}\right] \\ \times \left[e^{-\beta t_i} e^{\lambda \exp(-\beta t_i)} - e^{-\beta t_{i-1}} e^{\lambda \exp(-\beta t_{i-1})}\right] \end{array} \right\}}{\left[e^{\lambda \exp(-\beta t_i)} - e^{\lambda \exp(-\beta t_{i-1})}\right]^2} \\ + \sum_{i=1}^{m} R_i & \frac{\begin{array}{c} \left[- \exp(-2\beta t_{i-1}) e^{\lambda \exp(-\beta t_{i-1})} \left(1 - e^{\lambda \exp(-\beta t_{i-1})}\right) \right. \\ \left. - \exp(-2\beta t_{i-1}) e^{2\lambda \exp(-\beta t_{i-1})}\right] \end{array}}{\left(1 - e^{\lambda \exp(-\beta t_{i-1})}\right)^2} - \frac{e^\lambda}{(1 - e^\lambda)^2} \sum_{i=1}^{m} (D_i + R_i), \end{aligned} \tag{6}$$

$$\begin{aligned} \frac{\partial^2 l(\lambda, \beta)}{\partial \beta^2} = \sum_{i=1}^{m} D_i & \frac{\left\{ \begin{array}{c} (1 + \lambda) \left(t_{i-1} \exp(-\beta t_{i-1}) e^{\lambda \exp(-\beta t_{i-1})} - t_i \exp(-\beta t_i) e^{\lambda \exp(-\beta t_i)}\right) \\ \times \left[e^{\lambda \exp(-\beta t_i)} - e^{\lambda \exp(-\beta t_{i-1})}\right] \\ + \left[t_{i-1} \lambda \exp(-\beta t_{i-1}) e^{\lambda \exp(-\beta t_i)} + \lambda t_i \exp(-\beta t_{i-1}) e^{\lambda \exp(-\beta t_{i-1})}\right] \\ \left[\exp -\beta t_i e^{\lambda \exp(-\beta t_i)} - \exp -\beta t_{i-1} e^{\lambda \exp(-\beta t_{i-1})}\right] \end{array} \right\}}{\left[e^{\lambda \exp(-\beta t_i)} - e^{\lambda \exp(-\beta t_{i-1})}\right]^2} \\ + \sum_{i=1}^{m} R_i & \frac{\left\{ \begin{array}{c} (1 + \lambda)(t_i \exp(-\beta t_i) e^{\lambda \exp(-\beta t_i)}) \left(1 - e^{\lambda \exp(-\beta t_{i-1})}\right) \\ + \lambda t_i \exp(-2\beta t_i) e^{2\lambda \exp(-\beta t_i)} \end{array} \right\}}{\left(1 - e^{\lambda \exp(-\beta t_{i-1})}\right)^2}, \end{aligned} \tag{7}$$

$$\begin{aligned} \frac{\partial^2 l(\lambda, \beta)}{\partial \lambda \beta} = \sum_{i=1}^{m} D_i & \frac{\left\{ \begin{array}{c} \left[-t_i \exp(-\beta t_i) e^{\lambda \exp(-\beta t_i)} (1 + \lambda \exp(-\beta t_i)) \right. \\ \left. + t_{i-1} \exp(-\beta t_{i-1}) e^{\lambda \exp(-\beta t_{i-1})} (1 + \lambda \exp(-\beta t_{i-1}))\right] \\ \left[e^{\lambda \exp(-\beta t_i)} - e^{\lambda \exp(-\beta t_{i-1})}\right] \\ + \left[t_i \lambda \exp(-\beta t_i) e^{\lambda \exp(-\beta t_i)} (1 - \lambda \exp(-\beta t_i))\right] \\ \left[e^{-\beta t_i} e^{\lambda \exp(-\beta t_i)} - e^{-\beta t_{i-1}} e^{\lambda \exp(-\beta t_{i-1})}\right] \end{array} \right\}}{\left[e^{\lambda \exp(-\beta t_i)} - e^{\lambda \exp(-\beta t_{i-1})}\right]^2} \\ + \sum_{i=1}^{m} R_i & \left[\frac{\left\{ \begin{array}{c} t_i \exp(-\beta t_i) e^{\lambda \exp(-\beta t_i)} (1 - \lambda \exp(-\beta t_i)) \left(1 - e^{\lambda \exp(-\beta t_i)}\right) \\ - \left(t_i \exp(-\beta t_i) e^{\lambda \exp(-\beta t_i)}\right) \end{array} \right\}}{(1 - \lambda \exp(-\beta t_i))^2} \right]. \end{aligned} \tag{8}$$

Using the above algorithm, we can then obtain the one-step approximation estimators. The simulation analysis revealed that, as predicted, one step approximation estimators beat their counterparts in most circumstances; so, we report just the findings in the last section.

4. Methodology

4.1. **EM algorithm.** Suppose that $(D_i, R_i, t_i)_{i=1}^m$ are the observed data under progressive type-I interval censoring. Recall that under progressive type-I interval censoring total D_i number of failures are observed instead of exact failure lifetimes. Let us assume that x_{ij} are the exact lifetimes of the total number of observed failures,$i = 1, 2, ..., m$ and $j = 1, 2, ..., D_i$. Further suppose that z_{ij} are the lifetimes of the units those are censored at the inspection times $t_i, i = 1, 2, ..., m$ and $j = 1, 2, ..., R_i$. Then the complete sample of n number of units can be seen as a combination of observed data x_{ij} and censored data z_{ij}. Subsequently the log-likelihood function of (λ, β) given the complete sample can be written as:

$$ln\mathcal{L}_c(\boldsymbol{\theta}) \propto n\text{ln}\lambda + n\text{ln}\beta - n\lambda - n\text{ln}(1 - e^{-\lambda}) - \beta \sum_{i=1}^{m} \sum_{j=1}^{D_i} x_{ij} + \lambda \sum_{i=1}^{m} \sum_{j=1}^{D_i} \exp(-\beta x_{ij})$$
$$- \beta \sum_{i=1}^{m} \sum_{j=1}^{R_i} z_{ij} + \lambda \sum_{i=1}^{m} \sum_{j=1}^{R_i} \exp(-\beta z_{ij}).$$

The partial derivatives of the above function with respective to λ and β, and equating them to zero, gives:

$$\frac{n}{\lambda} = \frac{n(e^{\lambda} + 1)}{e^{\lambda} - 1} - \sum_{i=1}^{m} \left[\sum_{j=1}^{D_i} \exp(-\beta x_{ij}) + \sum_{j=1}^{R_i} \exp(-\beta z_{ij}) \right], \tag{9}$$

$$\frac{n}{\beta} = \sum_{i=1}^{m} \left(\sum_{j=1}^{D_i} x_{ij} + \sum_{j=1}^{R_i} z_{ij} \right) + \lambda \sum_{i=1}^{m} \left[\sum_{j=1}^{D_i} x_{ij} \exp(-\beta x_{ij}) + \sum_{j=1}^{R_i} z_{ij} \exp(-\beta z_{ij}) \right]. \tag{10}$$

The EM algorithm consist expectation step (E-step) and maximization step (M-step). The E-step replaces the expressions of observed and censored lifetimes by their expectations, whereas M-step maximizes the E-step at each iteration. Now let us consider that $(\lambda^{(k)}, \beta^{(k)})$ as the estimates of (λ, β) at the kth stage, then using (9) and (10)) the estimators of (λ, β) at the $(k + 1)$th stage are given by

$$\lambda^{(k+1)} = n \left[\frac{n(e^{\lambda^{(k)}} + 1)}{e^{\lambda^{(k)}} - 1} - \sum_{i=1}^{m} \Big(D_i \text{E}_{1i}(\lambda^{(k)}, \beta^{(k)}) + R_i \text{E}_{2i}(\lambda^{(k)}, \beta^{(k)}) \Big) \right]^{-1},$$
$$\beta^{(k+1)} = n \Bigg[\sum_{i=1}^{m} \Big(D_i \text{E}_{3i}(\lambda^{(k)}, \beta^{(k)}) + R_i \text{E}_{4i}(\lambda^{(k)}, \beta^{(k)}) \Big)$$
$$+ \lambda^{(k)} \sum_{i=1}^{m} \Big(D_i \text{E}_{5i}(\lambda^{(k)}, \beta^{(k)}) + R_i \text{E}_{6i}(\lambda^{(k)}, \beta^{(k)}) \Big) \Bigg]^{-1},$$

where

$$\begin{aligned}
\mathrm{E}_{1i}(\lambda^{(k)},\beta^{(k)}) &= \mathrm{E}\left(\exp(-\beta x_{ij})|x_{ij}\in(t_{i-1},t_i]\right)\\
&= \frac{1-e^{\lambda}}{e^{\lambda\exp(-\beta t_i)}-e^{\lambda\exp(-\beta t_{i-1})}}\int_{t_{i-1}}^{t_i} e^{-\beta x}f(x)dx,\\
\mathrm{E}_{2i}(\lambda^{(k)},\beta^{(k)}) &= \mathrm{E}\left(\exp(-\beta z_{ij})|z_{ij}\in(t_i,\infty)\right) = \frac{1-e^{\lambda}}{1-e^{\lambda\exp(-\beta t_i)}}\int_{t_i}^{\infty} e^{-\beta x}f(x)dx,\\
\mathrm{E}_{3i}(\lambda^{(k)},\beta^{(k)}) &= \mathrm{E}\left(x_{ij}|x_{ij}\in(t_{i-1},t_i]\right) = \frac{1-e^{\lambda}}{e^{\lambda\exp(-\beta t_i)}-e^{\lambda\exp(-\beta t_{i-1})}}\int_{t_{i-1}}^{t_i} xf(x)dx,\\
\mathrm{E}_{4i}(\lambda^{(k)},\beta^{(k)}) &= \mathrm{E}\left(z_{ij}|z_{ij}\in(t_i,\infty)\right) = \frac{1-e^{\lambda}}{1-e^{\lambda\exp(-\beta t_i)}}\int_{t_i}^{\infty} xf(x)dx,\\
\mathrm{E}_{5i}(\lambda^{(k)},\beta^{(k)}) &= \mathrm{E}\left(x_{ij}\exp(-\beta x_{ij})|x_{ij}\in(t_{i-1},t_i]\right)\\
&= \frac{1-e^{\lambda}}{e^{\lambda\exp(-\beta t_i)}-e^{\lambda\exp(-\beta t_{i-1})}}\int_{t_{i-1}}^{t_i} xe^{-\beta x}f(x)dx,\\
\mathrm{E}_{6i}(\lambda^{(k)},\beta^{(k)}) &= \mathrm{E}\left(z_{ij}\exp(-\beta z_{ij})|z_{ij}\in(t_i,\infty)\right) = \frac{1-e^{\lambda}}{1-e^{\lambda\exp(-\beta t_i)}}\int_{t_i}^{\infty} xe^{-\beta x}f(x)dx.
\end{aligned}$$

Further using an iterative procedure, the desired maximum likelihood estimates of (λ,β) can be obtained. Here the procedure can be terminated when a desired convergence is achieved for a small value of $\epsilon>0$ satisfying $|\lambda^{(k+1)}-\lambda^{(k)}|+|\beta^{(k+1)}-\beta^{(k)}|<\epsilon$. Observe that in the EM algorithm the expressions like $\mathrm{E}_{1_i},\mathrm{E}_{2_i}\ldots$ do not admit a closed form, subsequently these expressions need to be computed numerically. To avoid this situation we next propose Stochastic expectation maximization algorithm (SEM).

4.2. **SEM algorithm.** This section discusses the SEM algorithm for obtaining MLEs of (λ,β). In the preceding section, it was discovered that E-step had complicated and intractable calculations. Several authors have proposed several ways to eliminate the computational expense in the existing literature. [19] proposed the idea to approximate the expectations in the E-step by the Monte Carlo average, but still the maximization procedure turn out to have complicated and more time-consuming, see [20]. [9] proposed the notion of replacing the E-step with a stochastic step (S-step) and running it via simulation In many cases, the idea of SEM has been shown to be more suited than the EM method in terms of computing burden, see [22] and [23]. Next we use the same idea, and first generate the independent D_i number of samples of $x_{ij}, i=1,2,...,m$ and $j=1,2,...,D_i$ from the following conditional distribution function

$$F_D(x_{ij};\lambda,\beta|x_{ij}\in(t_{i-1},t_i]) = F(t_i;\lambda,\beta)-F(t_{i-1};\lambda,\beta), \quad x_{ij}\in(t_{i-1},t_i].$$

We further generate R_i number of samples of $z_{ij}, i=1,2,...,m$ and $j=1,2,...,R_i$ from the following conditional distribution function

$$F_Z(z_{ij};\lambda,\beta|z_{ij}>t_i) = \frac{F(z_{ij};\lambda,\beta)-F(t_i;\lambda,\beta)}{1-F(t_i;\lambda,\beta)}, \quad z_{ij}>t_j.$$

Now using these generated observations in 9 and 10, the $(k+1)$th stage estimator of (λ,β) can be obtained as

$$\lambda^{(k+1)} = n\left[\frac{n(e^{\lambda(k)}+1)}{e^{\lambda(k)}-1} - \sum_{i=1}^{m}\left(\sum_{j=1}^{D_i}\exp(-\beta x_{ij}) + \sum_{j=1}^{R_i}\exp(-\beta z_{ij})\right)\right]^{-1},$$

$$\beta^{(k+1)} = n\left[\sum_{i=1}^{m}\left(\sum_{j=1}^{D_i}x_{ij} + \sum_{j=1}^{R_i}z_{ij}\right) + \lambda^{(k)}\sum_{i=1}^{m}\left(\sum_{j=1}^{D_i}x_{ij}\exp(-\beta x_{ij}) + \sum_{j=1}^{R_i}z_{ij}\exp(-\beta z_{ij})\right)\right]^{-1},$$

Finally we denote the MLEs of (λ,β) as $(\hat{\lambda},\hat{\beta})$ after the termination of iterative procedure. Now the observed Fisher information matrix of the MLEs of (λ,β) can be obtained as

$$S(\lambda,\beta) = \begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix} = \begin{bmatrix}-\frac{\partial^2 l(\lambda,\beta)}{\partial\lambda^2} & -\frac{\partial^2 l(\lambda,\beta)}{\partial\lambda\partial\beta}\\ -\frac{\partial^2 l(\lambda,\beta)}{\partial\lambda\partial\beta} & -\frac{\partial^2 l(\lambda,\beta)}{\partial\beta^2}\end{bmatrix}^{-1}_{\lambda=\hat{\lambda},\beta=\hat{\beta}}$$

here the involved expressions are reported in 6, 7 and 8. Further the corresponding $100(1-\gamma)\%$ asymptotic confidence interval estimates for λ and β can be obtained as $\hat{\lambda} \pm Z_{\gamma/2}\sqrt{\sigma_{11}}$ and $\hat{\beta} \pm Z_{\gamma/2}\sqrt{\sigma_{22}}$ respectively, where $Z_{\gamma/2}$ is the upper $(\gamma/2)$th percentile of the standard normal distribution.

5. Simulation study and data analysis

5.1. **Simulation study.** The objective of this section is to investigate the performance of the proposed estimators through a simulation study. We first generate the data $(D_i,R_i,t_i)_{i=1}^m$ for given n,m, prefixed inspection times and censoring schemes under progressive type-I interval censoring using the algorithm proposed by [2]. According to the algorithm, first generate $D_1 \sim Bin(n,F(t_1;\lambda,\beta))$ and given the value of D_1, generate $R_1 = [q_1\times(n-D_1)]$, here $Bin(.,.)$ represents the binomial distribution. Further for $i=2,3,...,m$, we have

$$D_i|(D_{i-1},R_{i-1},...,D_1,R_1) \sim Bin\left(n-\sum_{j=1}^{i-1}(D_j+R_j), \frac{F(t_i;\lambda,\beta)-F(t_{i-1};\lambda.\beta)}{1-F(t_{i-1};\lambda,\beta)}\right),$$

with $R_i = [q_i\times(n-\sum_{j=1}^{t-1}(D_j+R_j))]$.

For simulation studies we consider different values of n such as $n=\{20,30,40,50,100,150,200\}$ with inspection times $t_1=0.1, t_2=0.3, t_3=0.5, t_4=0.7$ and $t_5=0.9$, and censoring schemes: $p1=(q_1=0.25,q_2=0.25,q_3=0,q_4=0,q_5=1), p2=(q_1=0,q_2=0,q_3=0.25,q_4=0.25,q_5=1)$ and $p3=(q_1=0,q_2=0,q_3=0,q_4=0,q_5=1)$. Notice that the censoring scheme p3 corresponds to the traditional type-I interval censoring.

For these values of n and p we generate simulated data using $EP(0.5,2.5)$ and $EP(2,1)$ distributions. The average of maximum likelihood estimates of the parameters (λ,β) and their mean square error (MSE) based on 100 Monte Carlo replications using SEM, EM and NR algorithms for mentioned distributions are shown in Tables 1 and 2 respectively. From these tables the estimates obtained using EM and SEM algorithms are almost close to each other for all the considered schemes. However the NR method estimates are marginally higher than the other methods. It is clear that the estimators using the SEM method have lower MSE values compared to other methods. On the other hand, the NR method has the least efficiency. The results show that by increasing the sample size,the estimated values of the parameters became closer to the true values and the efficiency is also increases.

Table 1. Maximum likelihood estimates of $EP(0.5, 2.5)$ distribution using SEM, EM and NR algorithms.

		SEM						EM						NR					
	n	20	30	50	100	150	200	20	30	50	100	150	200	20	30	50	100	150	200
$p1$	$\hat{\lambda}$	0.5179	0.5127	0.5167	0.4888	0.4931	0.4954	0.8395	0.6997	0.6974	0.6555	0.6509	0.6309	2.0787	1.2514	1.0459	0.9235	0.9064	0.8832
	$MSE(\hat{\lambda})$	0.0629	0.0555	0.0573	0.0006	0.0014	0.0003	0.9813	0.5128	0.4833	0.4573	0.3743	0.2571	0.8513	0.0611	0.0324	0.3935	0.7058	0.0678
	$\hat{\beta}$	2.6906	2.5658	2.4847	2.5530	2.4802	2.5075	2.5379	2.4718	2.3423	2.5187	2.3621	2.4478	3.0129	2.8877	2.7774	2.7297	2.6565	2.6191
	$MSE(\hat{\beta})$	0.8340	0.2694	0.1510	0.0897	0.0495	0.0374	1.1434	0.4359	0.3967	0.1607	0.1983	0.1451	10.080	5.2579	1.4416	4.3872	7.7222	7.7135
$p2$	$\hat{\lambda}$	0.5637	0.4906	0.4945	0.4848	0.4928	0.4924	0.8632	0.8252	0.7415	0.7312	0.5913	0.6074	0.8262	0.8100	0.8223	0.8056	0.8055	0.8096
	$MSE(\hat{\lambda})$	0.1816	0.0046	0.0020	0.0009	0.0005	0.0003	0.9317	0.7732	0.6126	0.5150	0.2012	0.1978	7.3298	2.4317	0.0086	0.6628	1.7693	12.689
	$\hat{\beta}$	2.5552	2.5440	2.5830	2.5066	2.5060	2.5012	2.4328	2.3778	2.4640	2.3757	2.5051	2.4952	2.6785	2.6830	2.6695	2.6621	2.6437	2.6275
	$MSE(\hat{\beta})$	0.6085	0.2647	0.2140	0.0683	0.0617	0.0513	0.9754	0.5834	0.4663	0.3018	0.1768	0.1513	0.2125	13.773	12.546	0.5896	1.5333	0.3366
$p3$	$\hat{\lambda}$	0.6001	0.5989	0.4877	0.4871	0.4918	0.4991	0.8490	1.0981	0.6703	0.7397	0.6319	0.5678	0.7641	0.7854	0.7728	0.7636	0.7639	0.7544
	$MSE(\hat{\lambda})$	0.2803	0.2499	0.0012	0.0005	0.0003	0.0002	1.0920	1.3513	0.4407	0.5285	0.2597	0.1210	12.916	2.5614	0.0468	0.8437	0.3090	0.3150
	$\hat{\beta}$	2.5304	2.5371	2.4851	2.4982	2.5296	2.5074	2.4338	2.2784	2.3960	2.3613	2.4001	2.4904	2.6629	2.6471	2.6405	2.6332	2.6210	2.6149
	$MSE(\hat{\beta})$	0.4267	0.3517	0.1666	0.0752	0.0427	0.0380	0.7422	0.7607	0.3565	0.2821	0.1550	0.0956	12.325	0.0103	6.3219	8.6560	1.8806	5.5610

Table 2. Maximum likelihood estimates of $EP(2, 1)$ distribution using SEM, EM and NR algorithms.

		SEM						EM						NR					
	n	20	30	50	100	150	200	20	30	50	100	150	200	20	30	50	100	150	200
$p1$	$\hat{\lambda}$	1.8106	1.9215	1.9835	2.0132	2.0169	2.0141	1.8257	1.8784	1.8615	1.9315	2.1071	2.0633	1.5677	2.1743	2.2905	2.2168	2.0856	2.0142
	$MSE(\hat{\lambda})$	0.3551	0.1690	0.0834	0.0251	0.0095	0.0079	1.4338	1.2677	0.4472	0.3446	0.2545	0.1736	2.4551	0.4235	2.6584	7.7856	1.4356	0.6324
	$\hat{\beta}$	1.1255	1.0688	1.0317	1.0011	0.9907	0.9865	1.2884	1.1853	1.1265	1.1014	1.0756	1.0408	1.7261	1.5776	1.4859	1.4324	1.3954	1.3608
	$MSE(\hat{\beta})$	0.1738	0.1245	0.0430	0.0143	0.0053	0.0044	0.4093	0.2649	0.1169	0.1105	0.0618	0.0303	0.9809	2.1193	11.139	0.2507	1.5681	0.1182
$p2$	$\hat{\lambda}$	1.9307	1.9586	2.0308	1.9936	2.0281	2.0020	1.7699	1.7516	1.7413	1.7357	1.8973	1.9784	2.6177	2.7228	2.7294	2.7525	2.6377	2.5575
	$MSE(\hat{\lambda})$	0.1989	0.1960	0.0291	0.0189	0.0129	0.0093	1.2597	1.0446	0.3735	0.2448	0.1806	0.1750	0.1490	0.0001	0.6559	1.3266	0.0065	0.1357
	$\hat{\beta}$	1.0569	1.0531	1.0036	0.9731	0.9998	0.9772	1.1901	1.1931	1.1235	1.0527	1.0610	1.0349	1.3993	1.4340	1.4380	1.4197	1.4105	1.3938
	$MSE(\hat{\beta})$	0.1359	0.1218	0.0161	0.0116	0.0069	0.0057	0.3492	0.3214	0.1277	0.0886	0.0502	0.0477	12.575	8.4211	1.1836	1.3706	0.4117	0.6881
$p3$	$\hat{\lambda}$	1.9138	1.9686	1.9776	2.0162	2.0250	2.0334	1.7930	1.8248	1.7016	1.8172	1.8166	1.9303	2.9650	3.0505	3.0197	2.9586	2.9141	2.8663
	$MSE(\hat{\lambda})$	0.2723	0.1668	0.0406	0.0139	0.0121	0.0093	1.2394	0.6003	0.4645	0.2814	0.2858	0.1403	15.366	0.9574	0.4108	2.6016	0.6358	16.606
	$\hat{\beta}$	1.0980	1.0756	1.0061	0.9894	0.9939	1.0032	1.1871	1.1217	1.1503	1.0831	1.0916	1.0498	1.4026	1.4013	1.4011	1.3934	1.3846	1.3729
	$MSE(\hat{\beta})$	0.1321	0.0802	0.0323	0.0081	0.0066	0.0050	0.2417	0.1831	0.1678	0.0798	0.0772	0.0366	2.1271	3.6771	1.1819	0.0082	3.2348	0.4609

Table 3. Average estimates based on 100 replications using EP(0.5,2.5) distribution.

	Algorithms		
	SEM	EM	NR
Average of $\hat{\lambda}$	0.511561	0.730028	0.920694
Average of $MSE(\hat{\lambda})$	0.050017	0.555044	2.444167
Average of $\hat{\beta}$	2.531422	2.428400	2.692622
Average of $MSE(\hat{\beta})$	0.211133	0.418178	5.574900

Table 4. Average estimates based on 100 replications using EP(2,1) distribution.

	Algorithms		
	SEM	EM	NR
Average of $\hat{\lambda}$	1.979900	1.806233	2.563383
Average of $MSE(\hat{\lambda})$	0.090561	0.567294	3.013444
Average of $\hat{\beta}$	1.024122	1.119828	1.434967
Average of $MSE(\hat{\beta})$	0.051533	0.153594	2.862017

Tables 3 and 4 show the average of the estimated λ and β parameters and the MSE values for three algorithms obtained from table 1 and 2 respectively. Table 3 shows that the SEM algorithm has the closest estimates (i.e. 0.511561,2.531422) to the real ones (i.e. 0.5,2.5). The average of MSE values for this algorithm is minimum. Table 4 shows that the SEM algorithm has the closest estimates (i.e. 1.9799,1.024122) to the real ones (i.e. 2,1). The average of MSE values for this algorithm is also minimum.

5.2. **Data analysis.** We evaluate a research including 112 patients with plasma cell myeloma treated at the National Cancer Institute (N.C.I) [7] to demonstrate the approaches used in this publication. This dataset is presented in table 5. We denoted the number of patients dropped out from the study at the end of the time interval $(t_{i-1}, t_i]$ by D_is. Although the progressive censoring numbers are not determined before to the research, the statistical inference processes investigated in this article are based on $R_i,\ i = 1, ..., m$, as numbers of withdrawals described in section 2, and therefore apply to this dataset. We consider $EP(\lambda, \beta)$ distribution to fit the data based on plot 1 of empirical distribution of survival and population distribution. Our estimates summarized in table 6.

Further we compare the goodness-of-fit of the data set to EP distribution with the other distribution such as Weibull distribution, with the shape and scale parameters as λ and β respectively, using negative log-likelihood criterion (NL) and Kolmogorov–Smirnov (K–S) test. Notice that for the K–S test, we define the maximum distance

$$D_n(F) = \sup_{0 \leq t \leq \infty} |\hat{F}(t; \lambda, \beta) - F(t; \hat{\lambda}, \hat{\beta})|,$$

as the distance between the empirical distribution $\hat{F}(t; \lambda, \beta)$ for the observed data using progressive type-I interval censoring and the population distribution $F(t; \hat{\lambda}, \hat{\beta})$. The empirical CDF at each inspection time t_i can be estimated as

$$\hat{F}(t; \lambda, \beta) = 1 - \Pi_{j=1}^{i}(1 - \hat{p}_j), \quad i = 1, 2, ..., m,$$

Table 5. Gathered survival data by N.C.I .

Interval (Months)	Number at risk	D_i	R_i
$[0, 5.5]$	112	18	1
$[5.5, 10.5]$	93	16	1
$[10.5, 15.5]$	76	18	3
$[15.5, 20.5]$	55	10	0
$[20.5, 25.5]$	45	11	0
$[25.5, 30.5]$	34	8	1
$[30.5, 40.5]$	25	13	2
$[40.5, 50.5]$	10	4	3
$[50.5, 60.5]$	3	1	2
$[60.5, \infty)$	0	0	0

Table 6. Estimates based on NR, EM and SEM algorithms.

	Algorithms		
	NR	EM	SEM
shape (λ)	0.89872	0.83664	0.83657
scale (β)	0.034314	0.03511	0.03512

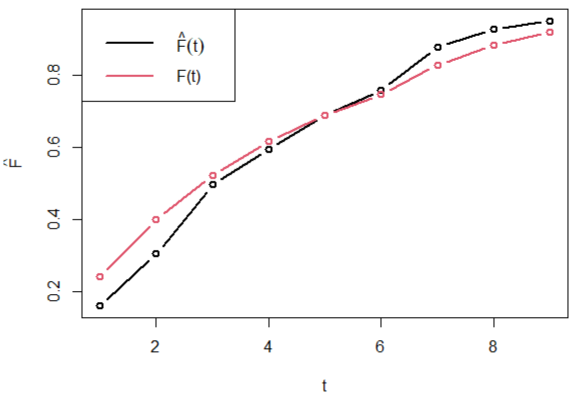

Figure 1. Difference of empirical and population distributions.

where $\hat{p}_1 = \frac{D_1}{n}$ and $\hat{p}_j = \frac{D_j}{n-\sum_{k=1}^{j-1}(D_k+R_k)}, \quad j = 2, 3, ..., m.$

We found that corresponding to EP distribution maximum likelihood estimates are (0.83657, 0.03512), NL is 194.55, and K–S value is 0.0961 whereas correspond to Weibull distribution maximum likelihood estimates are (1.3619, 0.0152), NL is 198.43, and K–S value is 0.0972. From all the calculated values it is seen that the EP distribution has smaller values of NL criterion and K–S test statistic, subsequently the EP distribution fits the data set reasonably well as compared to the Weibull distribution (by [12]). Now for the considered data the asymptotic 95% confidence interval estimates of λ and β are (0.8153, 0.8578) and (0.0334, 0.0368), respectively. Figure 1 shows difference of empirical and population (EP) distributions.

6. Conclusion

In this paper we considered lifetime data following EP distribution under progressive type-I interval censoring. We observed that MLEs of the unknown parameters of the distribution do not admit closed form, and further the implementation of EM algorithm still require optimization technique to solve the involved expressions in E-step. To avoid the numerical technique we considered SEM algorithm to obtain the MLEs. In simulation study we presented a comparison between the estimates obtained by SEM algorithm and estimates using NR and EM algorithms. Our simulation study showed that the performance of SEM algorithm is quite satisfactory on the basis of mean square error and by increasing the sample size, the estimated values of the parameters became closer to the true values and the efficiency is also increases. For illustration purpose we also considered a real data set.

COPYRIGHTS

References

[1] Arabi Belaghi, R., Noori Asl, M. and Sukhdev, S., (2017), On estimating the parameters of the Burr XII model under progressive type-I interval censoring, J. Stat. Comput. Simul, 87 (16), pp. 3132-3151.
[2] Aggarwala, R. (2001), Progressive interval censoring: some mathematical results with applications to inference, Commun. Stat. - Theory Methods, 30 (8-9), pp. 1921-1935.
[3] Ashour, S. and Afify, W., (2007), Statistical analysis of exponentiated Weibull family under type I progressive interval censoring with random removals, J. Appl. Sci. Res., 3 (12), pp. 1851-1863.
[4] Cancho, V. G., Louzada-Neto, F. and Barriga, G.D.C., (2011), The Poisson-Exponential lifetime distributionl, Comput. Stat. Data Anal., 55 (1), pp. 677–686.
[5] Chen, D. G. and Lio, Y. L., (2010), Parameter estimations for generalized Exponential distribution under progressive type-I interval censoring, Comput. Stat. Data Anal., 54 (6), pp.1581–1591.
[6] Chen, D. G, Lio, Y. L. and Jiang, N., (2013), Lower confidence limits on the generalized Exponential distribution percentiles under progressive type-I interval censoring, Commun. Stat. Simul. Comput. 42 (9), pp.2106–2117.
[7] Carbone, P. P., Kellerhouse, L. E. and Gehan, E. A., (1967), Plasmacytic myeloma: A study of the relationship of survival to various clinical manifestations and anomalous protein type in 112 patients, Am. J. Med. Sci., 42 (6), pp. 937–948.
[8] Dempster, A. P. and Laird, N. M., and Rubin, D.B., (1977), Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc., series B (methodological), 39 (1), pp. 1-22.
[9] Diebolt, J. and Celeux, G., (1993), Asymptotic properties of a stochastic em algorithm for estimating mixing proportions, Stoch. Models, 9 (4), pp. 599-613.
[10] Kus, C., (2007), A new lifetime distribution, Comput. Stat. Data Anal., 51 (9), pp.4497–4509.
[11] YL, L., Ding-Geng, C. and Tzong-Ru, T. (2011), Parameter estimations for generalized rayleigh distribution under progressively type-I intervalcensored data, Open J. Stat., 1, pp. 46– 57.
[12] Ng, H. K. T. and Wang, Z., (2009), Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample, J. Stat. Comput. Simul., 79 (2), pp. 145–159.
[13] Pradhan, B. and Gijo, E. V., (2013), Parameter estimation of lognormal distribution under progressive type-I Interval censoring, Tech. Rep. No. SQCOR-2013-02
[14] Peng, X. Y. and Yan, Z. Z., (2013), Bayesian estimation for generalized Exponential distribution based on progressive type-I interval censoring, Acta Math. Appl. Sin. Engl. Ser, 29 (2), pp. 391–402.
[15] Lin, Y. J. and Lio, Y. L., (2012), Bayesian inference under progressive type-I interval censoring, J. Appl. Stat., 39 (8), pp.1811–1824.
[16] Lin, C. T., Wu, S. J. and Balakrishnan, N., (2009), Planning life tests with progressively type-I interval censored data from the lognormal distribution, J. Stat. Plan. Infer., 139 (1), pp. 54–61.

[17] Wingo, D. R., (1993), Maximum likelihood estimation of burr xii distribution parameters under type II censoring, Microelectron Reliab. 33 (9), pp. 1251–1257.
[18] Singh, S., Tripathi, Y. M., (2016), Estimating the parameters of an inverse Weibull distribution under progressive type-I interval censoring, Stat. Pap. 59, pp. 21-56.
[19] Wei, G. C. and Tanner, M. A., (1990), A Monte Carlo implementation of the em algorithm and the poor man's data augmentation algorithms, J. Am. Stat. Assoc. 85 (411), pp. 699–704.
[20] Wang, F. K. and Cheng, Y. F., (2010), Em algorithm for estimating the burr XII parameters with multiple censored data, Qual. Reliab. Eng. Int,. 26 (6), pp. 615–630.
[21] Xiuyun, P., Zaizai, Y., (2011), Parameter estimations with gamma distribution based on progressive type-I interval censoring, In 2011 IEEE international conference on computer science and automation engineering, 1, pp. 449-453.
[22] Tregouet, D. A., Escolano, S., Tiret, L., Mallet, A., and Golmard, J. L., (2004), A new algorithm for haplotype-based association analysis: the stochastic-em algorithm, Ann. Hum. Genet. 68 (2), pp. 165–177.
[23] Zhang, M., Ye, Z. and Xie, M., (2014), A stochastic em algorithm for progressively censored data analysis, Qual. Reliab. Eng. Int., 30 (5), pp. 711–722.

Ahmad Mohammadi was born in 1979. He is currently an instructor in the Department of Statistics in Payame Noor University in Iran. He received his BSc in Statistics in 2003 from Tabriz University in Iran, his MSc in statistics in 2005 from Shahid Bahonar university in Iran. His research interests are in Reliability and Statistical Inferences.

Ali Shadrokh was born in 1967. He is currently an associate professor in the department of statistics at the Payame Noor University in Iran. He received his BSc in Statistics in 1991 from Ahvaz University in Iran, his MSc in statistics in 1996 from Tarbiat Modares university in Iran and his Ph.D in applied statistics in 2006 from University of Joseph Fourier in France. His research interests are in linear models, sequential regression, resampling methods and model selection.

Masoud Yarmohammadi was born in 1961. He is currently an associate professor in the department of statistics at the Payame Noor University in Iran. He received his BSc in Statistics in 1985 from Ahvaz University in Iran, his MSc in statistics in 1989 from UMIST in UK and his Ph.D in applied statistics in 1994 from University of Wales in UK. He published about 70 papers in preferred Journals and 9 books in statistics. From August 2022 to September 2024, he was a member of the board of directors of the Iranian Statistical Society. His areas of interests are time series analysis, singular spectrum analysis and multivariate analysis.