Int. J. Nonlinear Anal. Appl. In Press, 1–10
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2026.39931.5588

A comprehensive assessment of the performance and stability of maximum likelihood estimators for the exponential–poisson distribution: An analysis of sample size and censoring schemes on the NR, EM, and SEM algorithms under type I progressive interval censoring

Ahmad Mohammadi, Ali Shadrokh*, Masoud Yarmohammadi

Department of Statistics, Payame Noor University (PNU), P.O.Box 19395-4697, Tehran, Iran

(Communicated by Sirous Moradi)

Abstract

This study compares Newton-Raphson (NR), EM, and stochastic EM (SEM) for estimating Exponential-Poisson parameters under Type I progressive interval censoring. Simulations across sample sizes (20–200) and censoring schemes show SEM achieves the smallest MSE and highest stability, especially for small samples or complex censoring. EM performs reliably for large samples but is more variable for small samples. NR is highly sensitive to initial values and least stable. A real melanoma dataset confirms EP distribution fit (KS=0.0961) and replicates simulation patterns, emphasizing the need to choose estimation methods carefully.

Keywords: exponential-poisson distribution, type I progressive interval censoring, maximum likelihood estimator, SEM algorithm, EM algorithm, NR algorithm, mean squared error (MSE)
2020 MSC: Primary 62N02; Secondary 62F10

1 Introduction

Lifetime and reliability data are extensively utilized in medical, engineering, economic, and biological research, where incomplete observations frequently complicate statistical analysis. Among various forms of incomplete data, censored observations are particularly important, as the exact event times are not observable for all experimental units. Within the broad spectrum of censoring mechanisms, Type I progressive interval censoring has received considerable attention due to its flexibility and practical applicability in life-testing experiments and clinical studies [6, 9, 10].

A wide range of lifetime distributions has been studied under different censoring schemes, and selecting an appropriate model plays a critical role in achieving accurate inference. One commonly employed model is the Exponential–Poisson (EP) distribution, which is constructed by compounding the exponential distribution with a zero-truncated

*Corresponding author
Email addresses: ahmohammadi@pnu.ac.ir (Ahmad Mohammadi), a.shadrokh@pnu.ac.ir (Ali Shadrokh), masyar@pnu.ac.ir (Masoud Yarmohammadi)

Received: December 2025 *Accepted:* February 2026

Poisson distribution and is characterized as a decreasing failure-rate model. The EP distribution has been successfully applied in diverse fields such as medicine, business sciences, and reliability engineering [3]. While maintaining computational simplicity, it provides greater flexibility than the classical exponential distribution and is particularly suitable for modeling phenomena with decreasing hazard rates.

Parameter estimation for the EP distribution under Type I progressive interval censoring generally requires numerical and iterative techniques, owing to the analytical complexity of the resulting likelihood function. Previous studies have adopted the Newton–Raphson (NR) method, the Expectation–Maximization (EM) algorithm, and its stochastic extension, the Stochastic EM (SEM) algorithm, as practical approaches for obtaining maximum likelihood estimates [6, 7]. Each method has its own advantages and limitations: the NR algorithm is highly sensitive to initial values, the EM algorithm may exhibit slow convergence in certain scenarios, and the SEM algorithm often provides improved stability by incorporating randomization when dealing with complex data structures or substantial missing information.

Given the importance of the EP distribution and the widespread use of progressive interval censoring in applied research, a systematic evaluation and comparison of numerical estimation algorithms is of considerable interest. Building on the existing theoretical framework and the research foundation established in the thesis, this study investigates the properties of the EP distribution under Type I progressive interval censoring, presents the NR, EM, and SEM estimation procedures, and evaluates their empirical performance using real data. The primary objective is to provide a rigorous and reliable assessment of these algorithms and to highlight their respective practical strengths and limitations.

2 Exponential-Poisson lifetime distribution

The Exponential-Poisson distribution is constructed by compounding the exponential distribution with the zero-truncated Poisson distribution. Consider a random sample of size W, denoted by $T_1, T_2, ..., T_W$, drawn from an exponential distribution with the following probability density function:

$$f(t;\beta) = \beta e^{-\beta t} \qquad t \geq 0, \quad \beta > 0 \tag{2.1}$$

where the number of observations in this random sample, denoted by W, follows a zero-truncated Poisson distribution with the following probability mass function:

$$P(w,\lambda) = \frac{e^{-\lambda}\lambda^{w}}{(1-e^{-\lambda})\Gamma(w+1)}, \qquad w \in \mathbb{N}, \quad \lambda > 0 \tag{2.2}$$

where $\Gamma(.)$ denotes the gamma function, and W and the variables T are assumed to be independent. Now, let us define:

$$X = \min(T_1, T_2, ..., T_w). \tag{2.3}$$

It is known that the distribution of the first-order statistic from a random sample of size w drawn from an exponential distribution with parameter beta follows an exponential distribution with parameter $w\beta$.

$$f(x|w;\beta) = \beta w e^{(-\beta w x)}, \tag{2.4}$$

and the marginal probability density function and cumulative distribution function of X are given by:

$$f(x;\lambda,\beta) = \frac{\lambda\beta}{1-e^{-\lambda}} e^{-\lambda-\beta x+\lambda\exp(-\beta x)}, \qquad x \geq 0, \quad \beta,\lambda > 0, \tag{2.5}$$

$$f(x;\lambda,\beta) = (e^{\lambda\exp(-\beta x)} - e^{\lambda})(1-e^{\lambda})^{-1}, \tag{2.6}$$

The validity of these relationships is presented in Appendix A. The distribution of the random variable X is referred to as the Exponential-Poisson distribution and is denoted by $EP(\lambda,\beta)$. The parameters λ and β represent the distribution's shape and scale parameters, respectively. As λ approaches zero, the Exponential-Poisson distribution converges to the exponential distribution with parameter beta. The probability density function of the Exponential-Poisson distribution is a strictly decreasing function, taking the value $\lambda\beta(1-e^{-\lambda})^{-1}$ at $x = 0$.

An applied use of the Exponential–Poisson (EP) distribution was presented by Adamidis and Loukas [1]. In environmental studies, observations such as rainfall amounts often contain an excessive number of zero measurements, particularly during drought periods. In such datasets, zero-valued observations may occur at different sampling stages, implying that the value zero has a non-negligible probability and should be explicitly incorporated into the model-fitting process.

A similar phenomenon arises in reliability studies, where instantaneous failures may occur due to poor material quality or manufacturing defects. In these situations, standard lifetime distributions such as the gamma, Weibull, or lognormal distributions are inadequate, as they do not allow for a positive probability mass at zero. Although lifetime models are widely applied in economics, medicine, business, and actuarial sciences [8], they are typically formulated for strictly positive lifetimes.

The EP distribution, by allowing a nonzero probability density at zero, provides a suitable and flexible alternative for modeling scenarios in which zero-valued lifetimes are plausible.

3 Progressive censoring

A limitation of conventional Type I and Type II censoring schemes is that they do not permit the removal of test units at times other than the predetermined termination point of the experiment. Progressive censoring eliminates this restriction. This censoring scheme is defined as follows:

A total of n units are placed on a life-testing experiment at time zero. Upon the occurrence of the first failure, R_1 surviving units are randomly removed from the experiment. Upon the second failure, R_2 of the remaining surviving units are withdrawn, and this procedure continues until, at the time of the m-th failure, all remaining surviving units, that is, $R_m = n - m - R_1 - ... - R_{m-1}$, are removed from the life test. In this type of censoring, the values m and the sequence $R_1, R_2, ..., R_m$ are predetermined. See Figure 1 for an illustration.

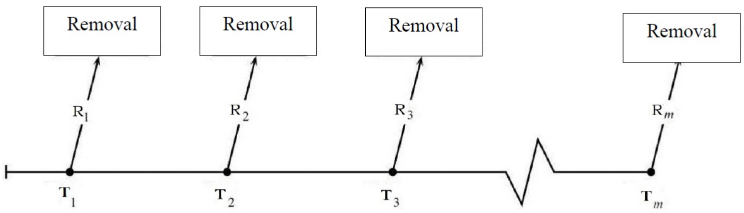

Figure 1: Schematic representation of the progressive censoring structure

The likelihood function corresponding to this censoring scheme is given by:

$$L(\theta) = c\prod_{i=1}^{m} f(x_i, \theta)[1 - F(x_i, \theta)]^{R_i}, \quad x_1 < x_2 < ... < x_m, \tag{3.1}$$

in which

$$c = n(n - R_1 - 1)(n - R_1 - R_2 - 2)...(n - R_1 - R_{m-1} - m + 1). \tag{3.2}$$

Note that

$$R_1 = ... = R_{m-1} = 0, \qquad R_m = n - m \tag{3.3}$$

then the conventional Type II censoring scheme is obtained. Furthermore, if we have:

$$R_1 = ... = R_m = 0, \tag{3.4}$$

then a non-censored scheme (a complete sample) is obtained. For further discussion on progressive censoring, one may refer to the book by Balakrishnan and Aggarwala [4]. The studies by Noori Asl et al. [11, 12, 13] and Arabi Belaghi et al. [3] are among the recent publications addressing progressive censoring schemes.

4 Type I progressive interval censoring

Type I progressive interval censoring is a combination of Type I interval censoring and progressive censoring. Suppose n components are placed on a life test at time t_0, and they are inspected at m predetermined inspection times $t_1, t_2, ..., t_m$, where t_m denotes the terminal inspection time. At the i-th inspection time, t_i, a total of D_i failures are recorded within the interval $(t_{i-1}, t_i]$, and R_i surviving components are randomly selected and removed from the experiment. Finally, at the time t_m, a total of D_m failures are observed within the interval $(t_{m-1}, t_m]$, and

$$R_m = n - \sum_{j=1}^{m} D_j - \sum_{j=1}^{m-1} R_j, \tag{4.1}$$

is removed from the experiment. The values R_i may be specified in advance as proportions $q_1, q_2, ..., q_{m-1}$, with $q_m = 1$, defined relative to the number of surviving components at time t_i for $i = (1, 2, ..., m)$. If y_i denotes the number of surviving components within the interval $(t_{i-1}, t_i]$, then, the number of randomly removed components at the i-th inspection time can be expressed as $R_i = [q_i \times y_i]$, for $i = 1, 2, ..., m$. Accordingly, a Type I progressive interval-censored sample can be represented as $(D_i, R_i, t_i)_{i=1}^m$ to m, where the total sample size is $n = \sum_{i=1}^m (D_i + R_i)$ from $i = 1$ to m. If $R_i = 0$ for $i = 1, 2, ..., m-1$, then the Type I progressive interval-censored sample reduces to the ordinary Type I interval-censored sample. The likelihood function for a Type I progressive interval-censored sample with distribution function $F(t, \theta)$ is given by:

$$L(\theta) \propto \prod_{i=1}^{m} [F(t_i, \theta) - F(t_{i-1}, \theta)]^{D_i} [1 - F(t_i, \theta)]^{R_i}. \tag{4.2}$$

Type I progressive interval censoring was introduced by Aggarwala [2], who developed statistical inference for the exponential distribution based on data subject to this censoring scheme. Subsequently, Chen and Lio [6] estimated the parameters of the generalized exponential distribution under the same censoring structure.

5 Model and methods

In this study, the Exponential–Poisson (EP) distribution is employed as the primary model for analyzing survival data subject to Type I progressive interval censoring. The data consist of patient follow-up times observed over a set of predetermined inspection intervals. At the end of each interval, the numbers of observed failures, removed units, and individuals remaining at risk are recorded. This data structure conforms to the formal definition of Type I progressive interval censoring in the survival analysis literature and enables the likelihood function to be constructed based on the available observed information.

After specifying the censoring scheme, the likelihood function under the EP model was formulated, and the unknown parameters were estimated using three numerical approaches: the Newton–Raphson (NR) method, the Expectation–Maximization (EM) algorithm, and the Stochastic Expectation–Maximization (SEM) algorithm. The NR method updates parameter estimates using the first and second derivatives of the log-likelihood function and iterates until convergence. The EM algorithm avoids direct solution of the likelihood equations by maximizing the expected complete-data log-likelihood conditional on the observed data. In contrast, the SEM algorithm replaces the expectation step with random sampling from the conditional distribution of the missing data, followed by a maximization step similar to that of the EM algorithm. This stochastic component often improves numerical stability when analyzing censored datasets.

To assess model adequacy, the empirical distribution function of the observed data was compared with the corresponding theoretical distribution function. In addition, the Kolmogorov–Smirnov test was applied to evaluate the goodness of fit. The resulting test statistic (KS = 0.0961) indicates that the EP model provides an adequate fit and captures the underlying behavior of the data. Overall, these results suggest that the proposed model and estimation procedures constitute an effective and reliable framework for analyzing survival data under Type I progressive interval censoring.

6 Findings

This section reports the performance of three estimation methods—the Newton–Raphson (NR) algorithm, the Expectation–Maximization (EM) algorithm, and its stochastic variant, the Stochastic EM (SEM) algorithm—for the Exponential–Poisson (EP) model under the Type I progressive interval censoring scheme. The primary objective is to compare the accuracy and stability of these algorithms in estimating the EP model parameters across different sample sizes and censoring patterns.

The results are summarized in Tables 2—9, which present the mean parameter estimates and their corresponding mean squared errors (MSEs) for each estimation method. All results reflect the behavior of the algorithms under various sample sizes (n=20,30,50,100,150,200) and three distinct censoring schemes (p1, p2, p3).

The numerical results indicate that the SEM algorithm generally produces more stable estimates with fewer fluctuations across most sample sizes. The EM estimator performs satisfactorily when the sample size is relatively large, but exhibits increased variability in smaller samples. In contrast, the NR method shows sensitivity to the choice of initial values and, in several scenarios, yields estimates with comparatively higher dispersion.

A detailed comparison of the results reported in the tables reveals that:

- the mean estimates obtained using the SEM algorithm are typically closer to the true parameter values;
- the EM algorithm demonstrates intermediate performance and is more reliable for larger sample sizes;
- the NR method achieves acceptable accuracy in certain cases, but its instability becomes more pronounced as the sample size decreases.

Overall, these findings suggest that the SEM algorithm constitutes a particularly suitable approach for estimating the parameters of the EP model under Type I progressive interval censoring, especially in situations involving small sample sizes or complex censoring structures.

Table 1: Mean maximum likelihood estimates of the EP(0.5, 5) distribution using the SEM, EM, and NR algorithms for different sample sizes

	N	**20**	**30**	**50**	**100**	**150**	**200**
SEM	$\hat{\lambda}$	0.5116	0.4857	0.5098	0.5039	0.5070	0.5025
	$MSE(\hat{\lambda})$	0.3344	0.1017	0.0609	0.0099	0.0096	0.0066
	$\hat{\beta}$	5.1076	5.0473	5.1855	5.1942	5.1511	5.1078
	$MSE\hat{\beta}$	2.4991	5.8869	2.9315	1.4508	0.8266	0.6636
EM	$\hat{\lambda}$	0.7334	0.5687	0.5693	0.5580	0.5068	0.4994
	$MSE(\hat{\lambda})$	1.6882	1.5623	1.0762	0.7421	0.4431	0.2064
	$\hat{\beta}$	4.9856	5.4223	4.8007	4.8395	5.0374	4.9418
	$MSE\hat{\beta}$	2.0629	2.0135	1.7635	1.1905	0.7439	0.5541
NR	$\hat{\lambda}$	0.6872	0.6676	0.6614	0.6168	0.5438	0.5319
	$MSE(\hat{\lambda})$	2.9337	1.5215	0.5752	1.2810	1.7595	0.3454
	$\hat{\beta}$	5.8170	5.8475	5.7429	5.6276	5.5599	5.5020
	$MSE\hat{\beta}$	2.4114	3.5077	2.5315	3.7614	2.7021	2.4010

This table compares the performance of the SEM, EM, and NR algorithms for estimating the parameters λ and β across different sample sizes. Overall, the SEM algorithm consistently provides estimates closest to the true parameter values and yields the smallest mean squared errors (MSEs) for both parameters, particularly for small and moderate sample sizes. This indicates a higher level of stability and robustness of SEM under the considered censoring scheme.

The EM algorithm also demonstrates satisfactory performance, especially as the sample size increases, with its estimates gradually converging toward the true values and MSEs decreasing accordingly. In contrast, the NR method exhibits noticeably larger variability and higher MSEs, particularly for smaller sample sizes, reflecting its sensitivity to sample size and initial values. As the sample size increases, the performance of all three algorithms improves; however, SEM remains the most reliable method in terms of accuracy and stability across all scenarios considered.

Table 2: Mean parameter estimates of the EP (0.5, 5) distribution obtained from the three algorithms

Algorithm	**SEM**	**EM**	**NR**
Mean of $\hat{\lambda}$s	0.5040	0.5307	0.5754
Mean of $MSE\hat{\lambda}$s	0.0301	0.5733	1.0803
Mean of $\hat{\beta}$s	5.1390	4.9642	5.5928
Mean of $MSE\hat{\lambda}$s	1.4090	0.9660	2.8031

This table summarizes the overall performance of the NR, EM, and SEM algorithms by reporting the mean estimates and mean squared errors (MSEs) of the parameters λ and β across all considered sample sizes. The results clearly indicate that the SEM algorithm achieves the highest level of accuracy, as evidenced by the smallest average MSEs for both parameters. This confirms the superior stability of SEM observed in the individual simulation scenarios.

The EM algorithm ranks second in terms of performance, producing relatively accurate estimates with moderate variability, particularly benefiting from larger sample sizes. In contrast, the NR method records the largest average estimation errors and higher dispersion, reflecting its sensitivity to the data structure and initial values. Overall, this aggregated comparison reinforces the behavioral patterns reported in the preceding tables and highlights SEM as the most reliable estimation approach for the EP model under Type I progressive interval censoring.

Table 3: Mean maximum likelihood estimates of the EP(2, 1) distribution using the SEM, EM, and NR algorithms for different sample sizes

	N	**20**	**30**	**50**	**100**	**150**	**200**
SEM	$\hat{\lambda}$	1.8850	1.9495	1.9973	2.0076	2.0165	2.0165
	$MSE(\hat{\lambda})$	0.2754	0.1772	0.0510	0.0193	0.0115	0.0088
	$\hat{\beta}$	1.0934	1.0658	1.0138	0.9878	0.9948	0.9889
	$MSE\hat{\beta}$	0.1472	0.1088	0.304	0.0113	0.0062	0.0050
EM	$\hat{\lambda}$	1.7064	1.8066	1.7587	1.8213	1.8463	1.8980
	$MSE(\hat{\lambda})$	1.3109	0.9708	0.4284	0.2902	0.2403	0.1629
	$\hat{\beta}$	1.2218	1.1667	1.1334	1.0790	1.0760	1.0418
	$MSE\hat{\beta}$	0.3334	0.2498	0.1441	0.0929	0.09630	0.0382
NR	$\hat{\lambda}$	2.3834	2.6492	2.6798	2.6426	2.5458	2.4793
	$MSE(\hat{\lambda})$	1.6567	0.4603	1.2417	2.5712	0.6926	1.1247
	$\hat{\beta}$	1.5093	1.4709	1.4416	1.4151	1.3968	1.3758
	$MSE\hat{\beta}$	2.2276	3.0725	1.1681	0.5431	1.7382	0.4224

This table reports the estimation performance of the NR, EM, and SEM algorithms for the EP model under an alternative censoring configuration across different sample sizes. The results indicate that both SEM and EM yield estimates of the parameter β that are close to the true value, particularly as the sample size increases. However, SEM clearly outperforms the other methods in estimating λ, as evidenced by substantially smaller MSEs across all sample sizes.

The EM algorithm demonstrates reasonable accuracy with moderate variability, especially for larger samples, while its performance deteriorates for smaller sample sizes. In contrast, the NR method exhibits considerable dispersion and instability, particularly for small sample sizes, leading to higher MSEs for both parameters. As the sample size increases, the estimation accuracy of all algorithms improves; nevertheless, SEM remains the most stable and reliable approach under this censoring scheme.

Table 4: Mean parameter estimates of the EP (2, 1) distribution obtained from the three algorithms

Algorithm	**SEM**	**EM**	**NR**
Mean of $\hat{\lambda}$s	2.0065	1.8453	2.5511
Mean of $MSE\hat{\lambda}$s	0.0341	0.3171	1.2636
Mean of $\hat{\beta}$s	1.0006	1.0796	1.4047
Mean of $MSE\hat{\lambda}$s	0.0196	0.0868	1.0812

The mean values clearly indicate that the SEM algorithm achieves the smallest estimation error for both parameters, substantially outperforming the EM method and exhibiting a pronounced improvement over the NR approach, which records the largest errors.

Table 5: Mean maximum likelihood estimates of the EP (2.5, 1) distribution using the SEM, EM, and NR algorithms for different sample sizes

	N	**20**	**30**	**50**	**100**	**150**	**200**
SEM	$\hat{\lambda}$	3.0003	2.6758	2.5882	2.5867	2.5366	2.5276
	$MSE(\hat{\lambda})$	0.6827	0.5723	0.0914	0.0345	0.0233	0.0181
	$\hat{\beta}$	1.0246	0.9673	0.9822	0.9932	0.9882	0.9858
	$MSE\hat{\beta}$	0.1228	0.0337	0.0167	0.0069	0.0050	0.0034
EM	$\hat{\lambda}$	3.3151	2.7314	2.7516	2.6622	2.5680	2.5582
	$MSE(\hat{\lambda})$	1.7972	1.7404	1.3329	0.5694	0.0961	0.0748
	$\hat{\beta}$	1.0526	0.9860	0.9901	0.9884	0.9976	0.9869
	$MSE\hat{\beta}$	0.2438	0.1370	0.0711	0.0357	0.0193	0.0078
NR	$\hat{\lambda}$	4.3075	3.9374	3.4673	3.0984	3.0271	2.9188
	$MSE(\hat{\lambda})$	1.2731	1.2913	0.9248	2.3353	1.5783	0.8618
	$\hat{\beta}$	0.8801	0.8620	0.8610	0.8648	0.8719	0.8727
	$MSE\hat{\beta}$	3.1395	2.5622	1.5606	2.5928	1.8122	1.9403

This table presents the estimation results of the NR, EM, and SEM algorithms for the EP model under the considered censoring scheme across different sample sizes. Overall, the SEM algorithm demonstrates the highest level of stability and accuracy in estimating both parameters. In particular, the SEM estimates of β remain consistently close to the true value across all sample sizes, accompanied by very small MSEs that decrease further as the sample size increases. Similarly, the SEM estimates of λ exhibit limited deviation from the true value, with substantially smaller MSEs compared to the EM and NR methods, especially for small and moderate sample sizes.

The EM algorithm shows moderate performance, producing reasonably accurate estimates of β and acceptable estimates of λ for larger sample sizes. However, its estimation accuracy deteriorates noticeably in smaller samples, as reflected by the increased MSE values, particularly for λ In contrast, the NR algorithm yields the least reliable performance, with estimates of λ and β that deviate considerably from the true parameter values and exhibit large MSEs across most sample sizes, indicating substantial variability and sensitivity to the data structure. As expected, increasing the sample size improves the estimation accuracy of all three algorithms; nevertheless, SEM consistently outperforms the competing methods in terms of both accuracy and stability under this scenario.

Table 6: Mean parameter estimates of the EP (2.5,1) distribution obtained from the three algorithms

Algorithm	**SEM**	**EM**	**NR**
Mean of $\hat{\lambda}$s	2.5716	3.0271	3.1369
Mean of $MSE\hat{\lambda}$s	0.0836	0.4384	1.3692
Mean of $\hat{\beta}$s	0.9879	0.9927	0.8697
Mean of $MSE\hat{\lambda}$s	0.0117	0.0374	2.0670

The mean values once again confirm the superiority of the SEM algorithm across both parameters. In particular, SEM yields the smallest mean squared errors for $\hat{\lambda}$ and $\hat{\beta}$, indicating higher accuracy and stability compared to the EM and NR methods. While the EM algorithm demonstrates acceptable performance, especially in estimating β, its MSEs remain noticeably larger than those of SEM. In contrast, the NR approach exhibits the largest estimation errors and the greatest deviation from the true parameter values, highlighting its sensitivity and reduced reliability under this scenario.

Table 7: Mean maximum likelihood estimates of the EP(5, 1) distribution using the SEM, EM, and NR algorithms for different sample sizes

	N	**20**	**30**	**50**	**100**	**150**	**200**
SEM	$\hat{\lambda}$	5.7309	5.5333	5.4144	5.2222	5.1738	5.0459
	$MSE(\hat{\lambda})$	0.4549	0.4167	0.2556	0.2170	0.1561	0.1186
	$\hat{\beta}$	0.7010	0.7688	0.8151	0.8782	0.8947	0.9323
	$MSE\hat{\beta}$	0.1813	0.1212	0.0964	0.0732	0.0659	0.0590
EM	$\hat{\lambda}$	6.3687	5.8071	5.6695	5.5371	5.3238	5.1067
	$MSE(\hat{\lambda})$	1.5117	1.2945	1.1918	1.2462	1.1772	1.0212
	$\hat{\beta}$	0.6998	0.7960	0.7991	0.8486	0.8742	0.8986
	$MSE\hat{\beta}$	0.2386	0.1728	0.1331	0.0984	0.0951	0.0768
NR	$\hat{\lambda}$	6.9715	6.6906	6.0598	5.8396	5.6309	5.2246
	$MSE(\hat{\lambda})$	1.6656	1.2497	2.2841	1.6277	1.2783	0.4189
	$\hat{\beta}$	0.6465	0.6897	0.7216	0.7366	0.7814	0.8071
	$MSE\hat{\beta}$	2.3906	1.7098	1.8510	1.8043	1.3043	1.6178

As the true value of λ increases, the differences among the estimation algorithms in terms of the mean squared error of $\hat{\lambda}$ become more pronounced across all sample sizes. The SEM algorithm consistently yields the smallest MSEs for both $\hat{\lambda}$ and $\hat{\beta}$, demonstrating superior accuracy and robustness even when the parameter magnitude is relatively large. The EM method exhibits moderate and relatively stable performance, with estimation errors that are systematically higher than those of SEM but substantially lower than those produced by the NR approach. In contrast, the NR algorithm shows considerable instability, particularly for smaller sample sizes, as reflected by its markedly larger MSE values. These results indicate that the advantage of SEM becomes increasingly evident as the underlying parameter value grows, highlighting its reliability compared to EM and, especially, NR under more challenging estimation scenarios.

The mean values for this scenario clearly indicate that the SEM algorithm delivers the best overall performance by a considerable margin, as evidenced by its substantially smaller mean squared errors for both $\hat{\lambda}$ and $\hat{\beta}$. The EM method ranks second, exhibiting moderate estimation accuracy with MSE values notably lower than those of the NR approach but higher than those achieved by SEM. In contrast, the NR algorithm performs the weakest, primarily due to its large MSE values, reflecting greater instability and reduced reliability under this setting.

Table 8: Mean parameter estimates of the EP(5, 1) distribution obtained from the three algorithms

Algorithm	**SEM**	**EM**	**NR**
Mean of $\hat{\lambda}$s	5.1978	5.3794	5.6667
Mean of $MSE\hat{\lambda}$s	0.1877	1.1529	1.1333
Mean of $\hat{\beta}$s	0.8842	0.8610	0.7672
Mean of $MSE\hat{\lambda}$s	0.0747	0.1019	1.6205

6.1 Real data analysis

To evaluate the performance of the Exponential–Poisson (EP) model and the associated estimation algorithms on real data, a survival dataset consisting of 112 patients diagnosed with malignant melanoma, originally reported by Carbon et al. [5], was analyzed. The data comprise a sequence of predetermined inspection times along with the corresponding numbers of individuals remaining at risk within each interval, which is consistent with a Type I progressive interval censoring framework.

The inspection times and the associated numbers of patients at risk are summarized in Table 10. Based on this censoring structure, the Newton–Raphson (NR), Expectation–Maximization (EM), and Stochastic Expectation–Maximization (SEM) algorithms were applied to estimate the parameters of the EP model.

The resulting maximum likelihood estimates obtained from the three algorithms are reported in Table 10. As shown in the table, the parameter estimates produced by the NR, EM, and SEM methods are generally close for this real dataset. Nevertheless, the SEM algorithm yields smoother and more stable estimates, whereas the NR method exhibits relatively greater deviation in some cases, a pattern that is consistent with the findings of the simulation study.

To assess the adequacy of the fitted model, the Kolmogorov–Smirnov (KS) test was employed. The resulting test statistic, 0.0961, indicates that the EP distribution provides an adequate fit to the observed data. In addition, Figure 2 presents a graphical comparison between the empirical distribution function and the fitted theoretical EP distribution, offering further visual evidence of the model's suitability.

Table 9: Survival times for patients with plasma cell myeloma

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, ∞)	0	0	0

Table 10: Parameter estimates based on different estimation methods for the dataset of Carbon et al. [5]

	Algorithms		
Parameter	**NR**	**EM**	**SEM**
Shape (lambda)	0.89872	0.83664	0.83657
Scale (beta)	0.034314	0.03511	0.03512

The results of this study demonstrate that, for the Exponential–Poisson (EP) model under Type I progressive interval censoring, the performance and stability of the three maximum likelihood estimation algorithms are strongly influenced by both the sample size and the censoring pattern. Across all simulation scenarios, the SEM algorithm consistently produced the lowest mean squared errors and estimates closest to the true parameter values, indicating superior accuracy and stability, particularly for small sample sizes or more complex censoring schemes. The EM algorithm exhibited reliable performance for larger samples, while the NR method, although adequate in certain cases, required careful initialization and tuning due to its sensitivity to starting values and relatively higher variability.

The analysis of the real melanoma patient dataset further confirmed these findings. The EP distribution provided an adequate fit to the data, as evidenced by the Kolmogorov–Smirnov statistic (0.0961), and the SEM algorithm maintained its superior performance in practical application. Overall, these results highlight the importance of selecting

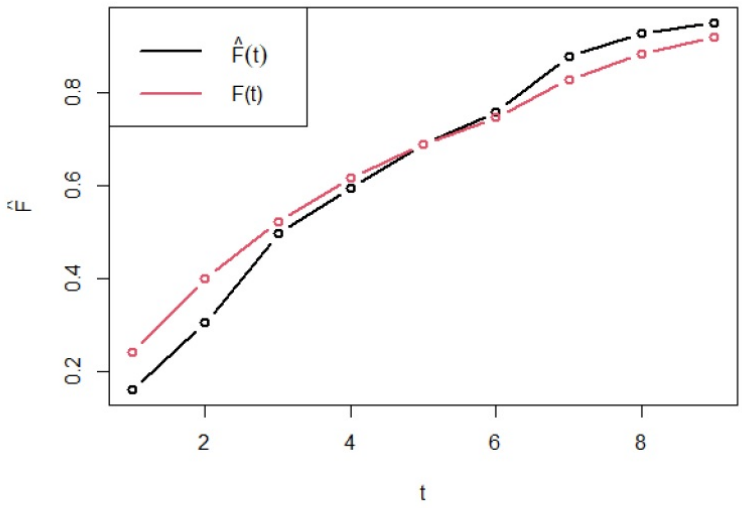

Figure 2: Difference between empirical distribution and theoretical distribution EP

an appropriate estimation algorithm to ensure accurate and stable inference when analyzing survival data subject to Type I progressive interval censoring.

7 Discussion and conclusion

In this study, the performance of three parameter estimation methods for the Exponential–Poisson (EP) distribution—namely the Newton–Raphson (NR), Expectation–Maximization (EM), and Stochastic Expectation–Maximization (SEM) algorithms—was investigated under a Type I progressive interval censoring scheme. The primary objective was to compare the stability, accuracy, and convergence behavior of these algorithms across different sample sizes and censoring patterns, and to evaluate their applicability to real-world data.

The simulation results clearly indicate that the SEM algorithm outperforms the other methods in terms of estimation accuracy and numerical stability. In nearly all scenarios, SEM produced estimates closer to the true parameter values and achieved the smallest mean squared errors (MSEs). This superior performance can be attributed to the stochastic component introduced in the E-step, which helps mitigate numerical instability and reduces sensitivity to complex censoring structures, particularly in small samples.

The EM algorithm also exhibited satisfactory and relatively stable behavior, especially as the sample size increased. Its deterministic E-step leads to regular convergence patterns; however, compared with SEM, EM showed slightly higher variability in settings involving heavier censoring or smaller sample sizes. These findings suggest that EM remains a reliable alternative when computational simplicity is preferred and sufficient sample information is available.

In contrast, the NR algorithm demonstrated the weakest overall performance. The results reveal that NR is highly sensitive to initial values and censoring configurations, leading to unstable estimates and noticeably larger MSEs in several cases. Consequently, the use of NR for estimating EP model parameters under Type I progressive interval censoring should be restricted to situations where good initial values are available and additional safeguards are employed to control numerical instability. An important observation from the simulation study is that increasing the sample size improves estimation accuracy for all three algorithms. This pattern reflects the consistency of the estimators and confirms that larger samples reduce variability and enhance the reliability of parameter estimates under progressive interval censoring.

The real data analysis further supports the conclusions drawn from the simulation study. The EP distribution provided an adequate fit to the melanoma dataset, as indicated by the Kolmogorov–Smirnov statistic of 0.0961. Consistent with the simulation findings, the SEM algorithm again exhibited the most stable and regular estimation behavior, reinforcing its practical advantage in real-world applications involving censored survival data.

8 Limitations and future directions

Despite the favorable performance of the proposed estimation methods, particularly the SEM algorithm, this study is subject to certain limitations. The analysis is restricted to the Exponential–Poisson distribution under a Type I progressive interval censoring scheme, and the simulation scenarios consider a limited range of sample sizes and parameter configurations. As a result, the generalizability of the findings to other censoring mechanisms or lifetime distributions may be limited. Future research may extend the proposed framework to alternative censoring schemes, such as Type II or hybrid censoring, and examine the performance of the estimation algorithms under different lifetime distributions. Additionally, further methodological developments of the SEM algorithm and its application to large-scale or more complex survival datasets represent promising directions for future work.

Overall, the results of this study demonstrate that the SEM algorithm is the most suitable method for estimating the parameters of the Exponential–Poisson distribution under Type I progressive interval censoring, particularly in settings with limited sample sizes or complex censoring patterns. The EM algorithm represents a dependable alternative with reasonable stability, while the NR method requires careful implementation and tuning. The agreement between the simulation and real data analyses confirms the robustness of these conclusions and supports the practical applicability of the proposed modeling framework.

References

[1] K. Adamidis and S. Loukas, *A lifetime distribution with decreasing failure rate*, Stat. Probab. Lett. **39** (1998), 35–42.

[2] R. Aggarwala, *Progressive interval censoring: some mathematical results with applications to inference*, Commun. Stat. Theory Meth. **30** (2001), 1921–1935.

[3] R. Arabi Belaghi, M. Noori Asl, and S. Singh, *On estimating the parameters of the Burr XII model under progressive type-I interval censoring*, J. Stat. Comput. Simul. **87** (2017), no. 16, 3132–3151.

[4] N. Balakrishnan and R. Aggarwala, *Progressive Censoring: Theory, Methods, and Applications*, Springer Science & Business Media, 2000.

[5] S. Carbon, A. Ireland, C.J. Mungall, S. Shu, B. Marshall, S. Lewis, AmiGO Hub, and Web Presence Working Group, *AmiGO: online access to ontology and annotation data*, Bioinformatics, **25** (2009), no. 2, 288–289.

[6] D.G. Chen and Y.L. Lio, *Parameter estimations for generalized exponential distribution under progressive type-I interval censoring*, Comput. Stat. Data Anal. **54** (2010), 1581–1591.

[7] S. Dey, S. Singh, and Y.M. Tripathi, *Estimation and prediction for a progressively censored generalized inverted exponential distribution*, Stat. Methodol. **32** (2016), 185–202.

[8] J. Lawless, *Statistical Models and Methods for Lifetime Data*, New York: John Wiley & Sons, 1982.

[9] C.T. Lin, S.J.S. Wu, and N. Balakrishnan, *Planning life tests with progressively type-I interval censored data from the lognormal distribution*, J. Stat. Plann. Inf. **139** (2009), 54–61.

[10] H.K.T. Ng and Z. Wang, *Statistical estimation for the parameters of the Weibull distribution based on progressively type-I interval censored sample*, J. Stat. Comput. Simul. **79** (2009), 145–159.

[11] M. Noori Asl, R.A. Arabi Belaghi, and H. Bevrani, *On Burr XII distribution analysis under progressive type-II hybrid censored data*, Methodol. Comput. Appl. Probab. **19** (2017), no. 2, 665–683.

[12] M. Noori Asl, R. Arabi Belaghi, and H. Bevrani, *Classical and Bayesian inferential approaches using Lomax model under progressively type-I hybrid censoring*, J. Comput. Appl. Math. **343** (2018), 397–412.

[13] B. Pradhan and E.V. Gijo, *Parameter estimation of lognormal distribution under progressive type-I interval censoring*, Technical report no. sqcor-2013-02, 2013.