

Parameter Estimation of the Poisson Exponentiated Half Logistic Distribution using the Anderson–Darling and Crammer–von Mises Estimation Techniques

Ibrahim Abdullahi^{1*}  and Muhammad Abdullahi Alhassan² 

¹Department of Mathematics, Yusuf Maitama Sule University, Kano, Nigeria.

²Aliko Dangote University of Science and Technology, Wudil, Kano, Nigeria.

*Corresponding Author: ibraabdul@googlemail.com

Abstract

Statistical inference is crucial in statistical studies, each method of parameter estimation possess some certain qualities and benefits depending on different types of data set and applied studies. Therefore, it is necessary to access the performance of various methods on particular statistical models in various applications. In this paper, we discussed the parameter estimation of the Poisson exponentiated half logistic (PEHL) distribution using Anderson–Darling estimation (AD) estimation and Crammer–von Mises (CV) estimation. Simulation studies and real data are considered to observe the suitability of the two methods. Their performance in both simulation and real data set suggested that both CV estimation and AD estimation methods are good alternatives in estimating the three parameters of PEHL depending on the practical application.

Keywords: Cramer–von-Mises, Anderson-Darling, estimation, Poisson exponentiated half logistic, Simulation

1. Introduction

The examination of statistical inferences regarding a statistical distribution holds significant importance in statistical research. Within the literature on statistical inferences, numerous authors have explored various estimation techniques for determining the unknown parameters of a statistical model, adopting diverse perspectives in both classical and frequentist approaches. Among the most commonly examined estimation methods are maximum likelihood and the Bayes method. Each method presents its own set of advantages and disadvantages, contingent on factors such as the nature of the model's basic functions, the optimization procedure employed, and the specific applications in question.

Assessing the performance of various estimation techniques based on a particular statistical model is crucial, even if theoretical comparisons may seem challenging. Simulation studies offer a practical means to compare performance across different parameter values and sample sizes, thereby providing a selection of estimation methods for real-world applications.

2. Related Works

Over the past few decades, numerous authors in the literature have discussed the performance of different estimation techniques for probability models. For instance, studies have addressed various estimations of the generalized exponential (GE) by Gupta and Kundu (2001), power-Pareto (EP) (Caeiro and Norouzrad, 2024), Gumbel distribution (GD) (Mahdi and Cenac, 2005; Aydin and Şenoğlu, 2015), XLindley (Alomair et al., 2024),

Xgamma (Sen et al., 2023), power XLindley (Meriem et al., 2022), two parameter XLindley (Muhammad and Liu, 2019), among others.

In this work, we aim to estimate the parameter of the Poisson exponentiated half logistic distribution (PEHL) (or Poisson generalized half logistic) proposed by (Meriem et al., 2022) using the Anderson–Darling (AD) and Crammer–von Mises (CV) estimation techniques. The PEHL model was demonstrated to have the flexibility in modeling various failure rates, and having closed form quantile function. We want to study how ADE and CVE estimators of the unknown parameters of the PEHL behave for different parameters and different sample sizes, by analyzing the bias and mean square error (MSE) of the estimators; also, to study the performance in practice using real data. This work will provide a guide for applied statisticians when dealing with PEHL in practice to have more choice apart from MLE, depending on the areas of the practical application.

The rest of the paper is organized as follows: In section 3, the PEHL parameter estimation is discussed using Anderson–Darling estimation (ADE) and Crammer–von Mises estimation (CVE). Section 4, provides the Monte Carlo simulation studies and real data examples. Section 5, conclusions.

3. Parameter Estimation by ADE and CVE for PEHL

The three-parameter PEHL distribution was proposed by (Meriem et al., 2022). The density function (pdf) and the cumulative distribution function (cdf) of the PGHL are given by

$$f(x) = \frac{2a\alpha\lambda e^{-\alpha x}(1 - e^{-\alpha x})^{a-1}}{(1 - e^{-\lambda})(1 + e^{-\alpha x})^{a+1}} e^{-\lambda\left(\frac{1-e^{-\alpha x}}{1+e^{-\alpha x}}\right)^a}, \tag{1}$$

and

$$F(x) = \frac{1 - e^{-\lambda\left(\frac{1-e^{-\alpha x}}{1+e^{-\alpha x}}\right)^a}}{1 - e^{-\lambda}}, \tag{2}$$

respectively, where $x, \alpha, \lambda, a > 0$. The quantile function (P(p)) function of the PGHL given as

$$P(p) = a^{-1} \left[\ln \left(1 + K_{\lambda}^{\frac{1}{a}}(p) \right) - \ln \left(1 - K_{\lambda}^{\frac{1}{a}}(p) \right) \right], \tag{3}$$

where $K_{\lambda}(p) = \left(\frac{\log(1-p(1-e^{-\lambda}))}{-\lambda} \right)$ and $0 < p < 1$.

3.1 Parameter estimation by ADE

The ADE is also based on the cdf of a probability model, it was developed based on the Anderson and Darling statistical tests for detecting how a sample distribution is closed to normality (Anderson and Darling, 1952). Let $\hat{\theta} = (\hat{a}, \hat{\lambda}, \hat{\alpha})^T$ be the Anderson Darling estimators of the parameters of the PGHL, then $\hat{\theta}$ can be computed by the minimization of $A_e(\theta)$ below

$$A_e(\theta) = -\frac{1}{n} - \frac{1}{n} \sum_{i=1}^n (2i - 1) \left[\log \left(1 - e^{-\lambda\left(\frac{1-e^{-\alpha x_i}}{1+e^{-\alpha x_i}}\right)^a} \right) + \log \left(e^{-\lambda\left(\frac{1-e^{-\alpha x_{n+1-i}}}{1+e^{-\alpha x_{n+1-i}}}\right)^a} - e^{-\lambda} \right) - 2 \log(1 - e^{-\lambda}) \right],$$

this can also be achieved by the solutions of the following nonlinear equations

$$\frac{\partial A_e}{\partial a} = -\frac{\lambda}{n} \sum_{i=1}^n (2i-1) \left[\frac{\psi^a(x_i; \alpha) e^{-\lambda \psi^a(x_i; \alpha)} \log \psi(x_i; \alpha)}{1 - e^{-\lambda \psi^a(x_i; \alpha)}} - \frac{\psi^a(x_{n+1-i}; \alpha) e^{-\lambda \psi^a(x_{n+1-i}; \alpha)} \log \psi(x_{n+1-i}; \alpha)}{e^{-\lambda \psi^a(x_{n+1-i}; \alpha)} - e^{-\lambda}} \right],$$

$$\frac{\partial A_e}{\partial \lambda} = -\frac{1}{n} \sum_{i=1}^n (2i-1) \left[\frac{\psi^a(x_i; \alpha) e^{-\lambda \psi^a(x_i; \alpha)}}{1 - e^{-\lambda \psi^a(x_i; \alpha)}} - \frac{\psi^a(x_{n+1-i}; \alpha) e^{-\lambda \psi^a(x_{n+1-i}; \alpha)} + e^{-\lambda}}{e^{-\lambda \psi^a(x_{n+1-i}; \alpha)} - e^{-\lambda}} - \frac{2e^{-\lambda}}{1 - e^{-\lambda}} \right],$$

$$\frac{\partial A_e}{\partial \alpha} = \frac{2a\lambda}{n} \sum_{i=1}^n (2i-1) \left[\frac{x_i \psi^{a-1}(x_i; \alpha) e^{-\lambda \psi^a(x_i; \alpha) - \alpha x_i}}{(1 - e^{-\lambda \psi^a(x_i; \alpha)})(1 + e^{-\alpha x_i})^2} - \frac{x_{n+1-i} \psi^{a-1}(x_{n+1-i}; \alpha) e^{-\lambda \psi^a(x_{n+1-i}; \alpha) - \alpha x_{n+1-i}}}{(e^{-\lambda \psi^a(x_{n+1-i}; \alpha)} - e^{-\lambda})(1 + e^{-\alpha x_{n+1-i}})^2} \right],$$

where

$$\psi(x_i; \alpha) = \frac{1 - e^{-\alpha x_i}}{1 + e^{-\alpha x_i}} \tag{4}$$

3.2 Parameter estimation by CVE

The Cramer–von-Mises estimators of the parameters of the PEHL distribution say $\hat{\theta} = (\hat{a}, \hat{\lambda}, \hat{\alpha})^T$ can be determined by the minimization of the $C(\theta)$ given below as

$$C(\theta) = \frac{1}{12n} + \sum_{i=1}^n \left(\frac{1 - e^{-\lambda \left(\frac{1 - e^{-\alpha x_i}}{1 + e^{-\alpha x_i}} \right)^a}}{1 - 1 - e^{-\lambda}} - \frac{2i-1}{2n} \right)^2.$$

On the other hand, $\hat{\theta}$ can be obtained by the solution of the nonlinear equations given as

$$\frac{\partial C}{\partial a} = \frac{2\lambda}{1 - e^{-\lambda}} \sum_{i=1}^n \left(\frac{1 - e^{-\lambda \psi^a(x_i; \alpha)}}{1 - e^{-\lambda}} - \frac{2i-1}{2n} \right) \psi^a(x_i; \alpha) e^{-\lambda \psi^a(x_i; \alpha)} \log \psi(x_i; \alpha),$$

$$\frac{\partial C}{\partial \lambda} = \frac{2}{1 - e^{-\lambda}} \sum_{i=1}^n \left(\frac{1 - e^{-\lambda \psi^a(x_i; \alpha)}}{1 - e^{-\lambda}} - \frac{2i-1}{2n} \right) \left[\psi^a(x_i; \alpha) e^{-\lambda \psi^a(x_i; \alpha)} - \frac{(1 - e^{-\lambda \psi^a(x_i; \alpha)}) e^{-\lambda}}{1 - e^{-\lambda}} \right],$$

$$\frac{\partial C}{\partial \alpha} = \frac{4a\lambda}{1 - e^{-\lambda}} \sum_{i=1}^n \left(\frac{1 - e^{-\lambda \psi^a(x_i; \alpha)}}{1 - e^{-\lambda}} - \frac{2i-1}{2n} \right) \frac{x_i \psi^{a-1}(x_i; \alpha)}{(1 + e^{-\alpha x_i})^2} e^{-\lambda \psi^a(x_i; \alpha) - \alpha x_i},$$

where $\psi(x_i; \alpha)$ is given by (4).

4. Simulation Studies and Data Illustration

This section discussed the Monte Carlo simulation results and real data performance of the two estimation methods.

4.1 Simulation studies

Monte Carlo simulation results are used to assess the performance of the estimation techniques proposed for the parameter estimation of the PEHL distribution, namely the ADE and CVE. For the comparison to be more

meaningful, both the estimation methods need to be considered under the same conditions, therefore, the same samples are used based on some selected parameters. We generated 10,000 samples of size $n = (30, 60, 90, \dots, 300)$ from the PEHL distribution for some parameter values from (3). The MSE will determine how close the estimators of the parameters are to actual values of the parameters. In view of these techniques in the estimation of the parameters of the PEHL, the most effective estimation method will yield both the bias and MSE closer to zero. The numerical computations were performed using R3.5.3 (Nash, 2014 and R Team, 2019), the numerical optimization method considered is the nlminb package, which allows bounds constraints and uses a quasi-Newton method (Ibrahim et al., 2023). We provide some important steps for the numerical simulation below:

The resulting simulation are given in the Figures 1 to 5. We observed that both bias and MSE of all estimators decreases as the sample size increases, i.e., the estimates are consistent and asymptotically unbiased for the parameters. The bias is negative in some cases, especially for parameter α . In addition, based on the simulation conducted we observed that λ has the biggest MSE in all cases. The simulation results suggested that both methods are suitable for estimation of PEHL, even though the ADE has the smallest MSE in most cases, showing better performance.

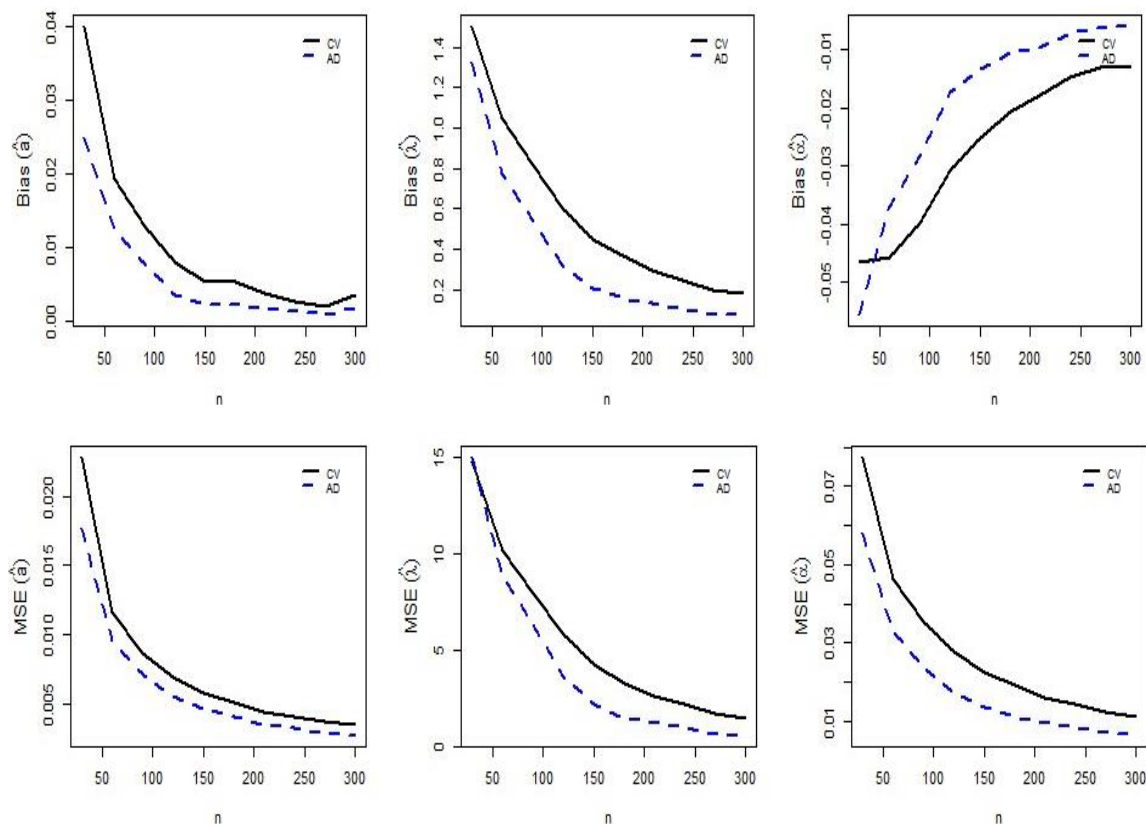


Figure 1: Plots of the bias and MSE of the CVE and ADE for $a = 0.5, \lambda = 0.9, \alpha = 0.5$.

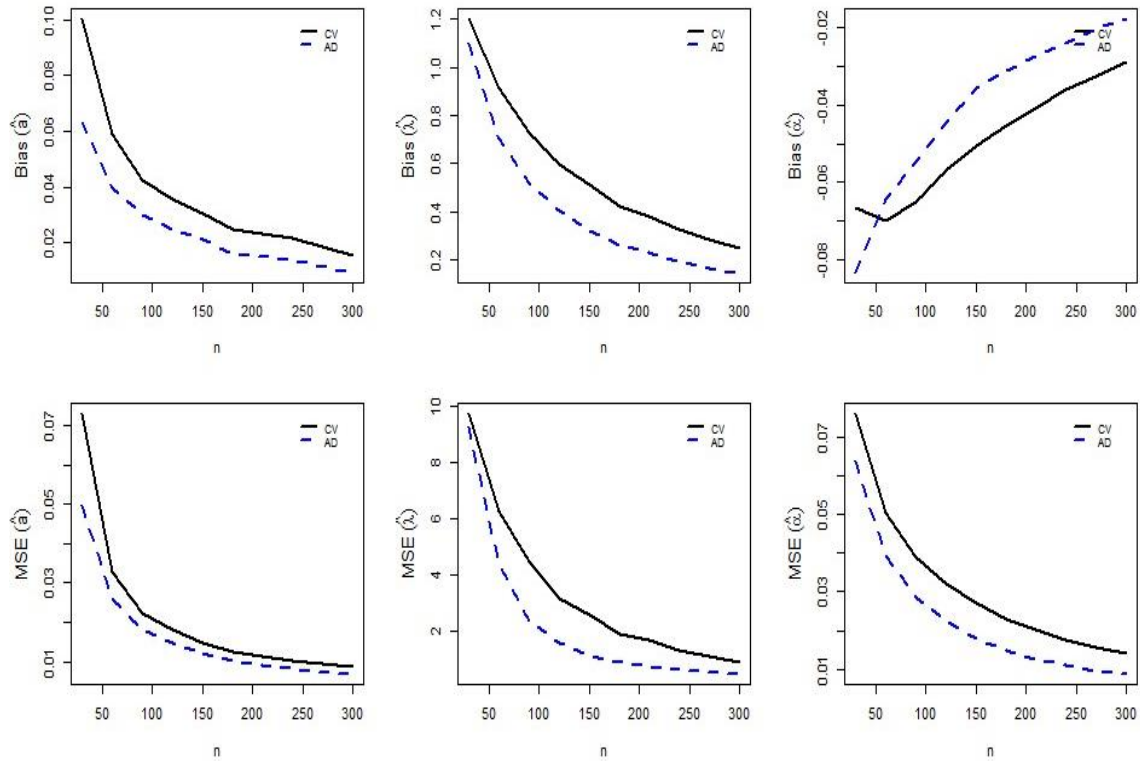


Figure 2: Plots of the bias and MSE of the CVE and ADE for $a = 0.8, \lambda = 0.5, \alpha = 0.7$.

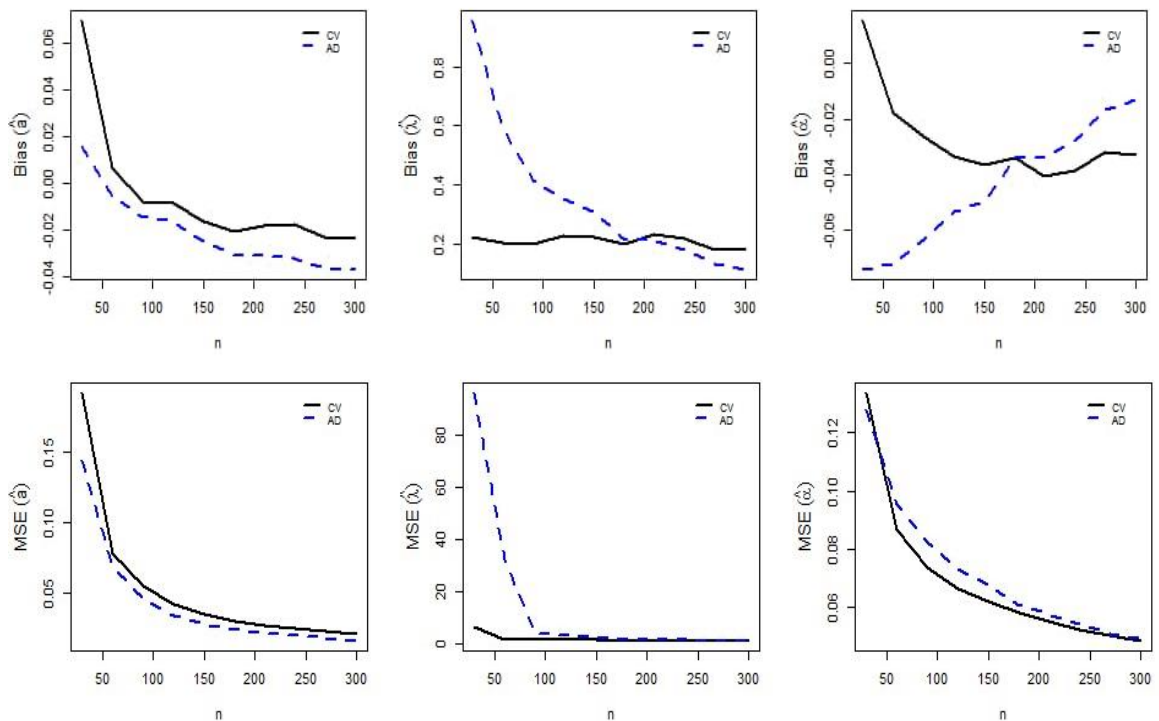


Figure 3: Plots of the bias and MSE of the CVE and ADE for $a = 1.5, \lambda = 1.0, \alpha = 1.2$.

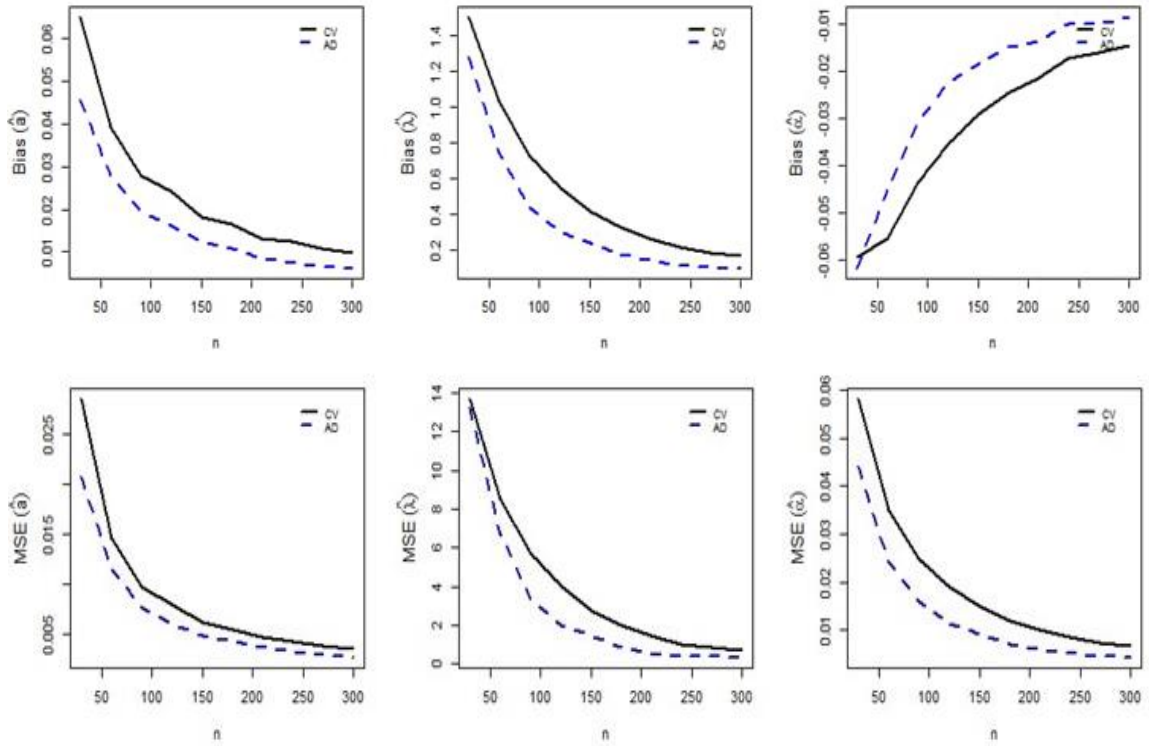


Figure 4: Plots of the bias and MSE of the CVE and ADE for $a = 0.5, \lambda = 0.5, \alpha = 0.5$.

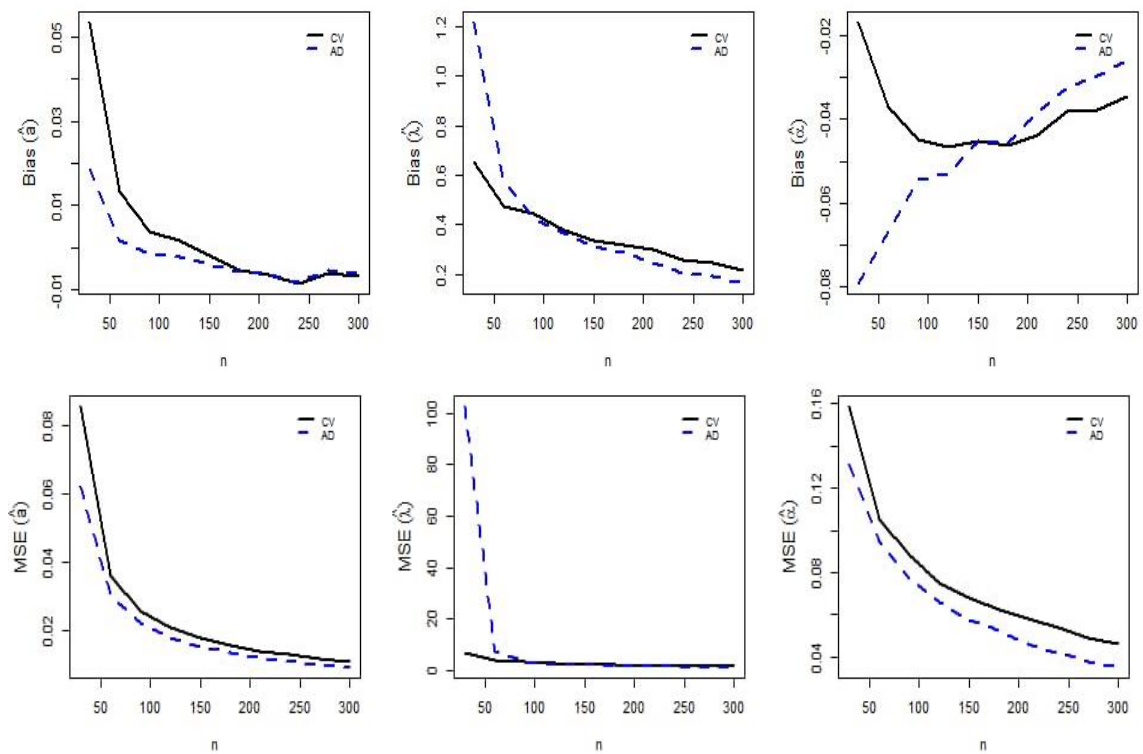


Figure 5: Plots of the bias and MSE of the CVE and ADE for $a = 1.0, \lambda = 1.1, \alpha = 1.0$.

4.2 Real data illustration

Here, we fitted the PEHL using ADE and CVE to two real data set for performance illustration. The first data can be found in Muhammad and Liu (2019). It represents the failure times of 50 components and the observations are: 0.036, 0.058, 0.061, 0.074, 0.078, 0.086, 0.102, 0.103, 0.114, 0.116, 0.148, 0.183, 0.192, 0.254, 0.262, 0.379, 0.381, 0.538, 0.570, 0.574, 0.590, 0.618, 0.645, 0.961, 1.228, 1.600, 2.006, 2.054, 2.804, 3.058, 3.076, 3.147, 3.625, 3.704, 3.931, 4.073, 4.393, 4.534, 4.893, 6.274, 6.816, 7.896, 7.904, 8.022, 9.337, 10.940, 11.020, 13.880, 14.730, 15.080.

The second data set can be found in Muhammad and Liu (2019). It is the failure times of 50 components (Unit: 1000 h), and the observations are: 0.061, 0.073, 0.075, 0.084, 0.086, 0.087, 0.088, 0.089, 0.089, 0.089, 0.099, 0.102, 0.117, 0.118, 0.119, 0.120, 0.123, 0.135, 0.143, 0.168, 0.183, 0.185, 0.191, 0.192, 0.199, 0.203, 0.213, 0.215, 0.257, 0.258, 0.275, 0.297, 0.297, 0.298, 0.299, 0.308, 0.314, 0.315, 0.330, 0.374, 0.388, 0.403, 0.497, 0.714, 0.790, 0.815, 0.817, 0.859, 0.909, 1.286.

Table 1 provided the numerical results obtained for the CVE and ADE methods, along with the goodness of fit tests. It can be noticed that both methods performed well, but CVE performed better than ADE because only the KS of second data the ADE is ranked first but all other measures CVE rank first in both data set. Hence, due to their performance in both simulation and real data, we can suggest both CVE and ADE as a better alternative, depending on the practical application. Figure 6, provided the plots of the density and survival curve for the first data, while Figure 7 for the second data, both showing how CVE and ADE provide good fits.

Table 1: First six moments of the CPGHL for some parameter values.

	Method	α	α	λ	KS	AD	CV
First Data	CVE	0.4959	0.1305	1.7078	0.1045 (0.6032)	0.8666 (0.0243)	0.1375 (0.0335)
	ADE	0.5216	0.1493	1.6334	0.1058 (0.5936)	0.8713 (0.0236)	0.1378 (0.0332)
Second Data	CVE	1.6272	4.0480	2.8014	0.1018 (0.6402)	1.1551 (0.0046)	0.1514 (0.0218)
	ADE	1.5060	3.4401	3.0212	0.0912 (0.7653)	1.1913 (0.0037)	0.1568 (0.0185)

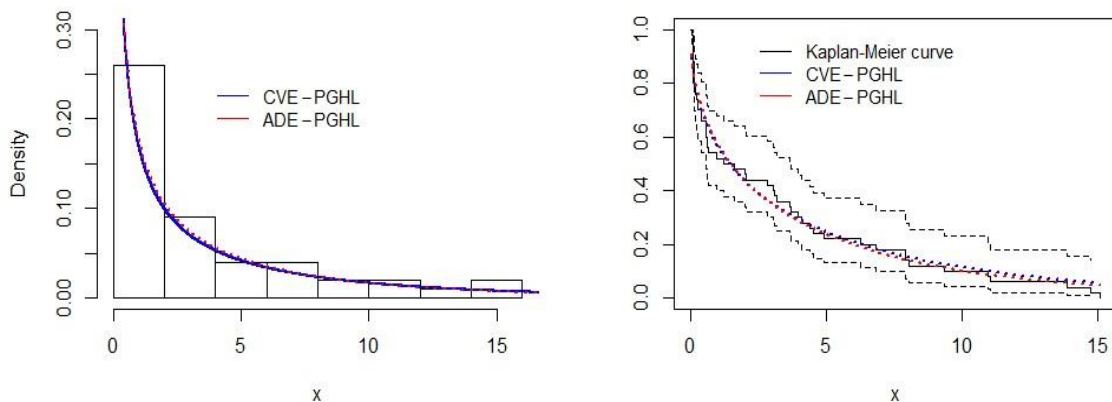


Figure 6: Plots of the density and survival curve for the second data

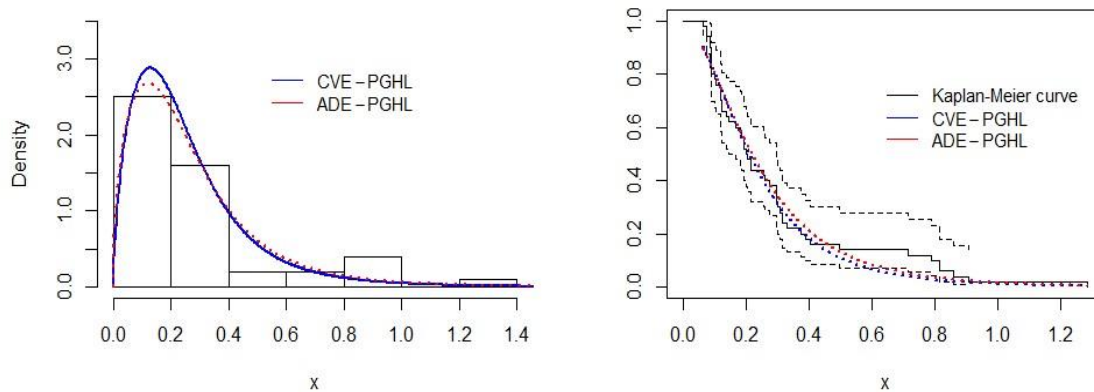


Figure 7: Plots of the density and survival curve for the second data

5. Conclusions

In this work, we provided parameter estimation of the PEHL using ADE and CVE. Simulation and real data are used to comment on the performance of the two estimation methods. The simulation results for some parameter values suggested that both methods are suitable for estimation of PEHL, even though the ADE has the smallest MSE in most cases, showing better performance. Also, both methods performed well on the real data, but CVE performed better than ADE as ranked using KS, AD, and CV goodness of fit in both data set. Hence, due to their performance in both simulation and real data, we can suggest both CVE and ADE as a good alternative, depending on the practical application for estimating the PEHL parameters.

Acknowledgment

I will like to thank the editor, and referees for their useful comments which improve the paper.

Conflicts of Interest

The author declares no conflict of interest.

References

- Alomair, A. M., Ahmed, M., Tariq, S., Ahsan-ul-Haq, M., & Talib, J. (2024). An exponentiated XLindley distribution with properties, inference and applications. *Heliyon*, 10(3).
- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23, 193–212.
- Aydin, D., & Şenoğlu, B. (2015). Monte Carlo comparison of the parameter estimation methods for the two-parameter Gumbel distribution. *Journal of Modern Applied Statistical Methods*, 14(2), 12.
- Caeiro, F., & Norouzirad, M. (2024). Comparing estimation methods for the power-Pareto distribution. *Econometrics*, 12(3), 20.
- Gupta, R. D., & Kundu, D. (2001). Generalized exponential distribution: Different methods of estimation. *Journal of Statistical Computation and Simulation*, 69(4), 315–337.
- Ibrahim, M., Shah, M. K. A., & Ahsan-ul-Haq, M. (2023). New two-parameter XLindley distribution with statistical properties, simulation and applications on lifetime data. *International Journal of Modelling and Simulation*, 1-14.

- Mahdi, S., & Cenac, M. (2005). Estimating parameters of Gumbel distribution using the methods of moments, probability weighted moments and maximum likelihood. *Revista de Matemática: Teoría y Aplicaciones*, 12(1-2), 151–156.
- Meriem, B., Gemeay, A. M., Almetwally, E. M., Halim, Z., Alshawarbeh, E., Abdulrahman, A. T., ... & Hussam, E. (2022). The Power XLindley Distribution: Statistical Inference, Fuzzy Reliability, and COVID-19 Application. *Journal of Function Spaces*, 2022(1), 9094078.
- Muhammad, M., & Liu, L. (2019). A new extension of the generalized half logistic distribution with applications to real data. *Entropy*, 21(4), 339.
- Nash, J. C. (2014). On best practice optimization methods in R. *Journal of Statistical Software*, 60(1), 1–14.
- R Core Team. (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.
- Sen, S., Alizadeh, M., Aboraya, M. A., M. Yousof, H., & Ibrahim, M. (2023). On truncated versions of Xgamma distribution: Various estimation methods and statistical modeling. *Statistics, Optimization & Information Computing*, 12(4), 943–961.