International Journal of Advanced Studies in Economics and Public Sector Management | IJASEPSM
p-ISSN: 2354-421X | e-ISSN: 2354-4228
Volume 12, Number 2 October, 2024

Comparative Study on Two Method for Estimating Three Parameter Exponentiated Odd Generalized Exponential Exponential Distribution

[1]**Turman Henry,** [2]**S. E Chaku,** [3]**Abubakar M. A,** [4]**Mary U. Adehi, &** [5]**Sunday Charity**
[1&5]*Department of Statistics, Federal School of Statistics, Manchok*
[2,3&4]*Department of Statistics, Nasarawa State University Keffi*

Article DOI: 10.48028/iiprds/ijasepsm.v12.i2.16

Abstract

Maximum likelihood is a relatively simple method of constructing an estimator for an un-known parameter μ. Maximum likelihood estimator (MLE) is obtained by taking the partial differentiation of likelihood function of a distribution. While the penalized MLE is obtained by adding a penalty to the regular MLE. The aim of this work is to make comparison between the regular MLE and penalized MLE of a three (3) parameter (α, λ and θ) Exponentiated Odd Generalized Exponential Exponential Distribution (EOGEED), this distribution has two shape and one scale parameters. Sample size 50, 100, 250 of simulated data were generated to check the performance of regular MLE and PMLE. The results show that the PMLE performs better than regular MLE especial for small sample size.

Keywords: *Maximum likelihood estimator, Penalized Maximum likelihood estimation, Exponentiated Odd Generalized Exponential Exponential Distribution*

Corresponding Author: **Turman Henry**

Background to the Study
Sometimes we have random data that we know comes from a parametric model, but we don't know the parameters. For instance, in an election between two candidates, polling data is drawn from a Bernoulli (p) distribution with an unknown parameter p. In this scenario, we would like to use the data to estimate the value of the parameter p, since it predicts the election outcome (Jeremy & Jonathan, 2014). Maximum likelihood estimation (MLE) is one method among others used to estimate parameter(s).

As earlier mentioned, there are numerous approaches to estimating unknown parameters, and these estimates are based on various data. Since the MLE provides a single value for the unknown parameter (our estimates will eventually incorporate intervals and probabilities), it is an example of a point estimate. The MLE has two benefits: it is frequently simple to calculate and, in basic examples, it matches our intuition. We must first determine the log likelihood of the likelihood function in order to easily obtain MLE. We take a partial derivative since the likelihood and log likelihood maxima coincide because ln (x) is an increasing function. This log-likelihood function will produce a system of non-linear likelihood equations that are too complex to solve analytically, but can be optimised numerically using global optimisation techniques found in programs like SAS, R, and Python.

The question of which parameter value has the highest probability in the observed data is addressed by MLE. In addition to parameter estimation, MLE can be applied to complex models like temporal dependence, covariate effect, and non-stationary models. According to Razira Aniza et al. (2020), the MLE method performs better with a larger sample size than with a smaller one, particularly if the sample size is less than 50. According to Hosking et al. (1985). Lazoglou & Anagnostopoulou (2017), sample sizes do not demonstrate good performance.

A flexible likelihood extension of MLE can be achieved through penalisation, which also helps to improve it on small properties and works well with both large and small data sets. Coles and Dixon (1999) proposed PMLE and demonstrated through an investigation how the PMLE enhances MLE. Compared to a direct estimate without penalties, PMLE is given a smoother estimation that is also more accurate. Instead of maximising the likelihood, penalised likelihood is usually achievable in terms of minimising the negative log-likelihood (also referred to as the loss function). Instead, we will minimise $q(\theta | X) = -L(\theta | X) + p(\theta)$, where the penalty function p penalises what we would consider to be unrealistic or unreasonable values of θ. This is the basic idea behind penalised likelihood (keep in mind that the penalty function is independent of the data). "Extreme" values, like infinite regression parameters, odds ratios near zero or infinity, or probabilities near zero or one, are examples of this unrealistic value. (Fang Wang et al., 2019; Patrick, 2021).

Simply put, Bayesian penalised likelihood refers to the study of the posterior mode's asymptotic and frequentist characteristics, or the maximum a posteriori (MAP) estimator Fraga & Wang, (2005); Alves, Neves, & Ros_ario, (2017). Simulation method studies allow researchers to answer specific question about data analysis statistical power, and best practices for obtaining accurate results in empirical research." (Hallgren, 2013), Simulation is a technique that can be used in any field of study but most people are not aware of it robustness of this tool like statistical procedure under ideal and non-ideal conditions.

Simulation can be carried out using different statistical soft ware's like R-software, SPSS, Excel etc and there are different methods of conducting simulation like Monte-Carlos method, inversion method etc. It is use to generate parameters wish can be used to test or descript probabilities models. Monte-Carlo method is handy for transforming problems of probabilistic nature into deterministic computations using the law of large numbers: eg you want to access the future value of your investment and to see what is the worst-case scenario for a given level of probability for such and many more real-life tasks you can use the Monte Carlo method. Ibrahim I. S., Doguwa S. I. *et al* (2023) said that the vast class of computational algorithms known as "Monte Carlo simulations" uses replicated random sampling to produce numerical results. The main idea is to introduce randomness to address problems that could be theoretically deterministic.

Methodology

Derivation of the EOGEED distribution

Maiti & Pramanik (2015) developed a distribution called Odd Generalized Exponential-Exponential Distribution (OGEED) and proposed it for modelling life data. One of the sole aims of developing a new distribution is to make the existing distribution more flexible for use especially for statistician to make use of on real life data to solve existing problems at hand. In statistics, distributions are used to describe real world phenomena.

In order to model or simulate life time data, we plan to suggest a new distribution termed the Extended Odd Generalised Exponentiated Exponential Distribution (EOGEED).

We will use the generalize (G-classes) distribution by Gupta *et al.,* (1998)

$$F(x)=G^{\alpha} \quad (1)$$

Therefore, the CDF of the proposed distribution is

$$F(x)=G^{\alpha} = (1 - e^{-\lambda(e^{\theta x}-1)})^{\alpha} \quad (2)$$

And the PDF is obtained as follows by differentiating F(x) which is our CDF in equation (2) with the respect to x

$$f(x) = \frac{d}{dx}F(x)$$

Using function of a function

$$\text{let } U = 1 - e^{-\lambda(e^{\theta x-1)}}$$

$$F(x) = (U)^{\alpha}$$

$$= \alpha(1 - e^{-\lambda(e^{\theta x-1)}})^{\alpha-1} x - e^{-\lambda(e^{\theta x}-1)} x - \lambda\theta e^{\theta x} \quad (3)$$

$$= \alpha\,\lambda\theta e^{\theta x} e^{-\lambda(e^{\theta x}-1)}(1 - e^{-\lambda(e^{\theta x}-1)})^{\alpha-1} \quad (4)$$

$$f(x;\alpha,\lambda,\theta) = \alpha\lambda\theta e^{\theta x} e^{-\lambda(e^{\theta x}-1)}(1 - e^{-\lambda(e^{\theta x}-1)})^{\alpha-1} \quad (5)$$

Hence equation (5) is the PDF of our propose distribution with $x > 0$ as random variable and $\alpha > 0$, $\lambda > 0$ and $\theta > 0$ where α, λ, and θ are shape and scale parameters respectively.

Maximum likelihood Estimation

The parameters of EOGEED (α, λ, and θ) can be estimated using the method of likelihood estimation (MLE). Let X = (x_1; x_2; ....; x_n) be a sample size from the EOGEED with parameter vector Θ.

Then log-likelihood function for $\theta$is given by

$$L(\theta|x_1 \ldots . x_n) = \prod_{i=1}^{n} f(x,\theta) \tag{6}$$

Recall that f(x) is given in equation (5)

$$\mathrm{L}(\theta) = \prod_{i=1}^{n} \alpha\lambda\theta e^{\theta x_i} e^{-\lambda\left(e^{\theta x_i}-1\right)} \left(1 - e^{-\lambda\left(e^{\theta x_i}-1\right)}\right)^{\alpha-1} \tag{7}$$

$$= \alpha^n \lambda^n \theta^n e^{\theta \sum x_i} e^{-\lambda \sum\left(e^{\theta x}-1\right)} \left(1 - e^{-\lambda\left(e^{\theta x}-1\right)}\right)^{\alpha-1} \tag{8}$$

Then we take natural log of likelihood function

$$l = logL(\theta) = nlog\alpha + nlog\theta + nlog\lambda + \theta \sum_{i=1}^{n} x_i - \lambda \sum_{i=1}^{n}\left(e^{\theta x_i} - 1\right) + (\alpha - 1) \sum_{i=1}^{n} log\left(1 - e^{-\lambda\left(e^{\theta x_i}-1\right)}\right) \tag{9}$$

We take the partial derivative w.r.t to the parameters α, λ *and* θ and equate to zero.

$$\frac{\partial logl}{\partial \hat{\alpha}} = \frac{n}{\hat{\alpha}} + \sum_{i=1}^{n} log\left(1 - e^{-\lambda\left(e^{\theta x}-1\right)}\right) = 0 \tag{10}$$

$$\frac{\partial logl}{\partial \hat{\lambda}} = \frac{n}{\hat{\lambda}} - \sum_{i=1}^{n}\left(e^{\theta x} - 1\right) + (\alpha - 1) \times \sum_{i=1}^{n} \frac{-\left(e^{\theta x}-1\right)}{1-e^{-\lambda\left(e^{\theta x}-1\right)}} = 0 \tag{11}$$

$$= \frac{n}{\lambda} - \sum\left(e^{\theta x} - 1\right) - (\alpha - 1)\sum \frac{e^{\theta x} - 1}{\left(1 - e^{-\lambda\left(e^{\theta x}-1\right)}\right)} \tag{12}$$

$$\frac{\partial logl}{\partial \hat{\theta}} = \frac{n}{\theta} + \sum_{i=1}^{n} x_i - \lambda \sum_{i=1}^{n} e^{\theta x} * x_i + (\alpha - 1) \sum_{i=1}^{n} \frac{-\lambda x e^{\theta x}}{1-e^{-\lambda\left(e^{\theta x}-1\right)}} = 0 \tag{13}$$

$$= \frac{n}{\hat{\theta}} + \sum x_i - \lambda \sum x_i e^{\theta x} - \lambda\left(\alpha - 1 \sum \frac{x e^{\theta x}}{1 - e^{-\lambda\left(e^{\theta x}-1\right)}}\right) \tag{14}$$

Penalized MLE

Then we take natural log of likelihood function with penalty term

$$l = logL(\theta) = \log(x - \theta) + nlog\alpha + nlog\theta + nlog\lambda + \theta \sum_{i=1}^{n} x_i - \lambda \sum_{i=1}^{n}\left(e^{\theta x_i} - 1\right) + (\alpha - 1) \sum_{i=1}^{n} log\left(1 - e^{-\lambda\left(e^{\theta x_i}-1\right)}\right) \tag{15}$$

We take the partial derivative w.r.t to the parameters α, λ, and θ and equate to zero.

$$\frac{\partial \log l}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} log\left(1 - e^{-\lambda(e^{\theta x_i} - 1)}\right) = 0$$

$$\frac{\partial \log l}{\partial \theta} = \frac{-1}{x-\theta} + \frac{n}{\theta} + \sum_{i=1}^{n} x_i - \lambda \sum_{i=1}^{n} x_i e^{\theta x_i} - (\alpha - 1)\lambda \sum_{i=1}^{n} x_i e^{\theta x_i} e^{-\lambda(e^{\theta x_i} - 1)} = 0 \tag{16}$$

$$\frac{\partial \log l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}(e^{\theta x_i} - 1) - (\alpha - 1)\sum_{i=1}^{n}\left(\frac{e^{-\lambda(e^{\theta x_i} - 1)}}{1 - e^{-\lambda(e^{\theta x_i} - 1)}}\left(e^{\theta x_i} - 1\right)\right) = 0. \tag{17}$$

Results

Here we used the Monte-Carlo method to generate random data from the Extended Odds Generalized Exponential-Exponential Distribution.

A Monte-Carlo simulation study was carried out considering N=1000 times for selected values of α, λ, and θ. Samples of size 250, 100 and 50 were considered and the required numerical evaluations are carried out.

Table 1: The descriptive Summary of the simulated dataset base on different sample size

N	Mean	Median	Mode	Variance	Skewness	Kurtosis	Min	Max
250	0.57546	0.47927	0.1	0.17018	0.61615	-0.48074	0.00066	1.70018
100	0.50317	0.3844	0.1,0.3	0.15065	0.93421	0.09022	0.01609	1.596
50	0.64529	0.57784	0.1	0.19743	0.40497	-0.71457	0.00066	1.65605

Table 2: Maximum likelihood estimate

Parameter	n=250	n=100	n=50
α	0.7025016	1.7575490	0.4422880
θ	0.8936218	1.1651659	0.7553843
λ	1.2263013	0.8156637	1.3375071

Table 3: Penalized maximum likelihood estimate

Parameter	n=250	n=100	n=50
α	0.039251	0.859213	0.168762
θ	0.021395	0.864526	0.175367
λ	0.026763	0.975933	0.180994

Table 4: Sample performance comparisons for the simulated data (MLE)

Performance measure	n=250	n=100	n=50
AIC	190.0613	59.42628	49.70368
CAIC	190.1589	59.67628	50.22541
BIC	200.6257	67.24179	55.43970
HQIC	194.3132	62.58936	51.888
(log(likelihood))	92.0360	26.71314	21.85184

Table 5: Some estimated test statistics

Sample size (n)	Anderson darling	Kolomogorov Smirov	
		D	p-value
250	0.2067454	0.32672	0.9523
100	0.4133306	0.090121	0.3911
50	--0.2162163	0.075214	0.9196

Table 6: Bias of EOGEED parameter estimation of MLE, and PMLE

Methods	Bias		
	α	θ	λ
MLE	-0.000205	0.000542	-0.000092
PMLE	-0.000241	0.000596	-0.000122

Table 7: RMSE of EOGEED parameter estimation of MLE, and P MLE

Methods	RMSE		
	α	θ	λ
MLE	0.016520	0.026041	0.018067
PMLE	0.016461	0.022900	0.017712

Result and discussion of finding for the simulated data set

Using the simulated data set shows that our distribution is positively skewed which means that it will fit right skewed data set as show in table 1. Three sample size was considered n=250, 100 and 50. Table 4 show that our distribution fits better with smaller sample size because the values of AIC, CAIC, BIC, HQIC have their least value when n=50, (49.70368, 50.22541, 55.43970, 51.888 and 21.85184 respectively). The Anderson – Darling values of all samples show that they follow or come from same probability distribution. Also, the Kolomogorov Smirov result (p-value) shows that the samples are from same population and come from same probability distribution. The values of the biases of both of MLE and PMLE are tending to zero as seen in Table 6. Also, the result of the RMSE show that the PMLE is better base on the value on Table 7.

Conclusion

The simulation study shows that the PMLE gives a better estimate compared to MLE and our distribution perform better with small sample size than bigger sample as seen in our result above.

References

Coles, S., & Dixon, M. (1999). Likelihood based inference for extreme value models. *Extremes 2*(1), 5-23.

Fang, W., Liang, P., Yongcheng, Q. & Meiping, X. (2019). Maximum penalized likelihood estimation for the endpoint and exponent of a distribution, *Statistica Sinica* 29 (2019), 203-224 doi.org/10.5705/ss.202016.0336

Fraga, A. I., Neves, C. & Ros_ario, P. (2017). A general estimator for the right endpoint with an application to supercentenarian womens records, *Extremes 20, 199*{237.

Hall, P. & Wang, J. Z. (2005). Bayesian likelihood methods for estimating the end point of a distribution. *J. R. Stat. Soc.* Ser B Stat. Methodol. 67, 717{729.

Hosking, J. R. M., Wallis J. R. &wood, E. F. (1985). Estimation of the Generalizied Extreme-value Distribution by the method of probability-weighed moment. *Technometrics* 27(3)25-261

Ibrahim A. S. & Doguwa S. I. S., *et al* (2023). New generalized Odd Fréchet-G (NGOF-G) Family of Distribution with Statistical Properties and Applications, *A periodical of the Faculty of Natural and Applied Sciences,* UMYU, Katsina. 2(3), 100-107 https://doi.org/10.56919/usci.2323.016

Jeremy, O. & Jonathan, B. (2014). Maximum likelihood estimates lecture note Class 10,18.05 Spring.

Lazoglou, G. & Anagnostopoulou, C. (2017). An overview of statistical methods for studying the extreme rainfalls in mediterranean, *Proceedings, 1*(5), 681.

Maiti, S. S., & Pramanik, S. (2015). Odds generalized exponential – exponential distribution, *Journal of Data Science 13*(2015), 733-754

Muhammad, H. T. & Gauss, M. C. (2016). Compounding of distributions: a survey and new generalized classes, *Journal of Statistical Distributions and Applications* (2016) 3:13DOI 10.1186/s40488-016-0052-1

Patrick, B. (2021). Penalize likelihood. University of Iowa. Likelihood theory (BIOS 7110)

Tahir, M. H., Cordeiro, G. M., Alizadeh, M. & Mansoor, M. (2015). The Odd generalized family of distributions with applications. *Journal of Statistical Distributions and Applications.* Available at http:// doi: 10.1186/s40488-014-0024-2.