Proceedings of the 1st International Conference on Statistical Sciences and Data Analytics

A SIMULATION ANALYSIS ON COMPARING THE EFFICIENCY AND CONSISTENCY OF MAXIMUM LIKELIHOOD ESTIMATION AND MAXIMUM PRODUCT OF SPACING METHODS OF ODD GOMPERTZ EXPONENTIAL DISTRIBUTION

Jibril Yahaya Kajuru[1,*], Ibrahim Abubakar Sadiq[1], Reuben Oluwabukunmi David[1], Yakubu Aliyu[1], Abubakar Usman[1] and Yahaya Zakari[1]

Department of Statistics, Ahmadu Bello University, Zaria
[*]Corresponding author Email: jibrilyk4u@gmail.com

ABSTRACT

Simulation studies are essential tools for evaluating the performance of parameter estimation methods in probability distributions, particularly when analytical solutions are complex or intractable. This study investigates the efficiency and consistency of Maximum Likelihood Estimation (MLE) and Maximum Product of Spacing (MPS) methods for the Odd Gompertz Exponential Distribution (OGED). This research estimates the parameters of the OGED via MLE and MPS methods and conducts a Monte-Carlo simulation-based comparison across varying sample sizes and parameter settings, evaluating their performance using bias, root mean squared error (RMSE), and convergence properties. Findings from the study reveal that both estimation techniques were consistent, as the Bias and RMSE decrease at different sample sizes of 20 to 1000. However, for all sample sizes and different actual values of parameters, the MLE proved to be a better estimator than MPS because the estimates converged more to the true parameter values and had a lower RMSE across all sample sizes. Thus, the findings offer practical insights into the relative strengths of these estimation methods for complex distributions like the OGED.

Keywords: Simulation study, OGED, Parameter Estimation, Maximum Likelihood, Maximum Product Spacing, Bias, and RMSE.

1. Introduction

Over the decades, simulation studies have played a vital role in the development and validation of new probability distributions through data generated under controlled conditions to study the behaviour of parameters in probability models. It allows researchers to generate data from known probability distributions for assessing the efficiency, consistency, and robustness of parameter estimators of the model (Burton *et al*., 2006). Unlike analytical methods, simulation studies offer flexibility in modelling complex scenarios where theoretical derivation is difficult, and they also assess the performance of various parameter estimation techniques, such as MLE and MPS (Kozubowski & Panorska, 2005). Simulation studies aid in assessing the practical merits and demerits of each method across a variety of probability distributions, thereby guiding researchers toward the most appropriate estimation method for their specific applications. Despite its adaptability, simulation requires careful design to avoid confusing conclusions. Factors such as model misspecification, inappropriate metrics, or inadequate replications can compromise the validity of findings (Carpenter *et al*., 2007).

The Gompertz distribution is a generalization of the exponential distribution and is widely applied in various fields of study, particularly to lifetime and survival analysis (Kajuru, 2025). It has also received considerable attention from different researchers in the fields of demography, biology, economics, and so on. The exponential distribution is widely used in reliability engineering and survival analysis of data due to its simplicity, memoryless property, and analytical tractability (Maurya *et al*., 2016). It also exhibits the same features of a constant failure rate (Oguntunde, 2017). However, the failure rate property of the exponential distribution renders it inappropriate for modelling real-world scenarios with bathtub and inverted bathtub shapes (Lemonte, 2013).

Parameter estimation serves as a fundamental aspect of statistical modelling and inference, aiming to determine unknown parameters of a probability distribution based on observed data. Parameter estimation for compound distributions like OGED often relies on likelihood-based methods, and various estimation methods are widely discussed in the literature. Researchers often use traditional estimation approaches such as the method of moments, method of least squares, method of weighted least squares, and method of maximum likelihood estimation (MLE), with each having its merit and limitations, but among these methods, the most popular method of estimation is the maximum likelihood estimation method (Singh, 2014). The MLE, pioneered by Fisher (1922), is widely favoured for its consistency, sufficiency, and asymptotic efficiency under regular conditions (Casella & Berger, 2002). However, its performance can be reduced in small samples. The Method of Moments (MoM), introduced by Pearson, is a simpler alternative that equates sample moments to hypothetical moments, offering natural estimators, though often less efficient than MLE (Kendall & Stuart, 1979). Least Squares (LS) and Weighted Least Squares (WLS) estimators are especially popular in regression analysis (Rasheed *et al*., 2014). The MPS introduced by Cheng and Amin (1983) serves as a competitor of MLE in cases where the estimates from MLE break down, and it maximizes the geometric mean of spacing in ordered data and performs well for skewed and heavy-tailed distributions (Ranneby, 1984; Cheng & Amin, 1983).

Habu *et al*. (2024) introduced the Topp-Leone Distribution using Maximum Product Spacing (MPS) and Maximum Likelihood Estimate (MLE). The parameters of Log-Topp-Leone (L-T-L) distribution, Topp-Leone-Exponential (T-L-E) distributions, and Top-plane-Weibull (T-L-W) distribution were estimated via the methods of MLE and MPS. Simulation studies were carried out to estimate the model's parameters and assess the efficacy of the estimators. The study found that the estimates of MLE and MPS were efficient.

More so, Abdulsalam *et al*. (2021) estimate the MPS and ML of the Gompertz Rayleigh Distribution through a simulation study. The result reveals that for all sample Sizes and different actual values of parameters, the MLE is shown to be a better estimator than MPS because of their lower RMSE values, but for parameters at sample sizes from 100−1000, the MPS was better.

There exist extended forms of Exponential distributions; however, the technique of Maximum likelihood estimation is frequently used to estimate their parameters. The parameters of the Odd Gompertz Exponential distribution (OGED) by Kajuru *et al*. (2023) were estimated using only MLE. No known simulation study has specifically compared MLE and MPS within the context of OGED. This research, therefore, fills a methodological gap by assessing their performance across varying sample sizes and parameter values, thus providing insight into their relative advantages in terms of bias, RMSE, and convergence behaviour.

2. Methodology

2.1 The Odd Gompertz Exponential Distribution

Using a sub-family of distribution from Odd Gompertz-G derived by Kajuru *et al*. (2023). The CDF, PDF, and Quantile function of OGED are defined as follows:

$$F_{OGE}(x;\theta,\gamma,\lambda)=1-e^{-\frac{\theta}{\gamma}\left(e^{\gamma\frac{(1-e^{-\lambda x})}{(1-(1-e^{-\lambda x}))}}-1\right)} \quad (1)$$

$$f_{OGE}(x;\theta,\gamma,\lambda)=\theta\lambda e^{-\lambda x}\left(1-(1-e^{-\lambda x})\right)^{-2} e^{\gamma\frac{(1-e^{-\lambda x})}{(1-(1-e^{-\lambda x}))}} e^{-\frac{\theta}{\gamma}\left(e^{\gamma\frac{(1-e^{-\lambda x})}{(1-(1-e^{-\lambda x}))}}-1\right)} \quad (2)$$

$\forall x;\theta,\gamma,\lambda>0$, where θ and γ are the Scale and Shape parameters, respectively

$$x_u=-\frac{1}{\lambda}\left[\log\left(1-\left[\frac{\frac{1}{\gamma}\log\left(1-\frac{\gamma}{\theta}\log(1-u)\right)}{\left[1+\left[\frac{1}{\gamma}\log\left(1-\frac{\gamma}{\theta}\log(1-u)\right)\right]\right]}\right]\right)\right] \quad (3)$$

2.2 Monte Carlo Simulation

In this section, we adopt a Monte Carlo simulation approach to compare and determine the performance of MLE and MPS methods for the OGED distribution parameters. Based on OGED, we use different parameter values and sample sizes (20-1000), and the estimation procedures were compared based on bias and root mean square error (RMSE) of the estimates.

$$\text{Bias}=\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\varphi}_i-\varphi_i\right) \quad (4)$$

$$\text{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\varphi}_i-\varphi_i\right)^2} \quad (5)$$

Where N = Number of replications (e.g, Monte Carlo runs), $\hat{\varphi}$ = Estimates of $\hat{\varphi}$ at the ith replication, $\hat{\varphi}$ = The true parameter value

2.3 Maximum Likelihood Estimation Method

Let $x_1,x_2,...,x_n$ be a random sample of size n from the OGE distribution. Then, the log-likelihood function based on the observed sample for the vector of parameters $(x;\theta,\gamma,\lambda)^T$ is given as

$$\ell(\theta,\gamma,\lambda)=\prod_{i=1}^{n}\ln f_{OGE}(x_i;\theta,\gamma,\lambda) \text{ , where } \ell(\theta,\gamma,\lambda)=L$$

$$L=n\ln\theta+n\ln\lambda+\lambda\sum x_i+\gamma\sum(e^{\lambda x_i}-1)-\frac{\theta}{\gamma}\sum\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right) \quad (6)$$

Taking the partial derivative of (6) with respect to θ,γ,λ, we have

$$\frac{\partial L}{\partial\theta}=\frac{n}{\theta}-\frac{1}{\gamma}\sum\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right) \quad (7)$$

$$\frac{\partial L}{\partial\gamma}=\sum(e^{\lambda x_i}-1)+\frac{\theta}{\gamma^2}\sum\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right)-\frac{\theta}{\gamma}\sum(e^{\lambda x_i}-1)e^{\gamma(e^{\lambda x_i}-1)} \quad (8)$$

$$\frac{\partial L}{\partial \lambda}=\frac{n}{\lambda}+\sum x_i+\gamma\sum x_i e^{\lambda x_i}-\frac{\theta}{\gamma}\sum\left[\gamma x_i e^{\lambda x_i}e^{\gamma(e^{\lambda x_i}-1)}\right] \tag{9}$$

The equations (7), (8), and (9) are nonlinear, cannot be solved analytically, necessitating the use of analytical tools to solve them numerically.

2.4 Maximum Product of Spacing Method

Let $x_1, x_2, ..., x_n$ be a random sample from the OGE distribution having CDF $F(x;\theta,\gamma,\lambda)$ and $x_1, x_2, ..., x_n$ represents the corresponding ordered sample. The spacing

$$\Omega_i = F(x_{(i)}) - F(x_{(i-1)}) \quad for\ i = 1, 2, ..., n+1$$

where $F(x_{(0)}) = 0$ and $F(x_{(n+1)}) = 1$

Therefore,

$$F(x_{(i)};\theta,\gamma,\lambda)=1-e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x(i)}-1)}-1\right)} \tag{10}$$

and

$$F(x_{(i-1)};\theta,\gamma,\lambda)=1-e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda(x_{i-1})}-1)}-1\right)} \tag{11}$$

Thus,

$$\Omega_i=\left[1-e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x(i)}-1)}-1\right)}\right]-\left[1-e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda(x_{i-1})}-1)}-1\right)}\right] \tag{12}$$

The parameter estimates are obtained by maximizing

$$\phi(x;\theta,\gamma,\lambda)=\frac{1}{n+1}\sum_{i=1}^{n}\ \log\Omega_i \tag{13}$$

$$\phi(x;\theta,\gamma,\lambda)=\frac{1}{n+1}\sum_{i=1}^{n}\ \log\left[\left[1-e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x(i)}-1)}-1\right)}\right]-\left[1-e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda(x_{i-1})}-1)}-1\right)}\right]\right] \tag{14}$$

Now $L=\phi(x;\theta,\gamma,\lambda)$

Differentiating L with respect to individual parameters yields the parameter estimates and solving the nonlinear equations.

$$\frac{\partial L}{\partial \theta}=\frac{1}{n+1}\sum_{i=1}^{n+1}\ \frac{1}{\Omega_i}\left[I_1(x_{(i)};\theta,\gamma,\lambda)-I_2(x_{(i-1)};\theta,\gamma,\lambda)\right] \tag{15}$$

$$\frac{\partial L}{\partial \theta}=\frac{e^{\gamma(e^{\lambda x_i}-1)}-1}{\gamma}.e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right)}$$

where $I_1(x_{(i)};\theta,\gamma,\lambda)=\frac{e^{\gamma(e^{\lambda x_i}-1)}-1}{\gamma}.e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right)}$ and $I_2(x_{(i-1)};\theta,\gamma,\lambda)=\frac{e^{\gamma(e^{\lambda x(i-1)}-1)}-1}{\gamma}.e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x(i-1)}-1)}-1\right)}$

$$\frac{\partial L}{\partial \gamma}=\frac{1}{n+1}\sum_{i=1}^{n+1}\ \frac{1}{\Omega_i}\left[J_1(x_{(i)};\theta,\gamma,\lambda)-J_2(x_{(i-1)};\theta,\gamma,\lambda)\right] \tag{16}$$

$$\frac{\partial L}{\partial \gamma}=\left[\frac{\theta}{\gamma^2}(e^{\gamma(e^{\lambda x_i}-1)}-1)-\frac{\theta}{\gamma}\cdot e^{\lambda x_i}-1\cdot e^{\gamma(e^{\lambda x_i}-1)}\right]\cdot e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right)}$$

where $J_1(x_{(i)};\theta,\gamma,\lambda)=\left[\frac{\theta}{\gamma^2}(e^{\gamma(e^{\lambda x_i}-1)}-1)-\frac{\theta}{\gamma}\cdot e^{\lambda x_i}-1\cdot e^{\gamma(e^{\lambda x_i}-1)}\right]\cdot e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right)}$

and $J_2(x_{(i-1)};\theta,\gamma,\lambda)=\left[\frac{\theta}{\gamma^2}(e^{\gamma(e^{\lambda x_{(i-1)}}-1)}-1)-\frac{\theta}{\gamma}\cdot e^{\lambda x_{(i-1)}}-1\cdot e^{\gamma(e^{\lambda x_{(i-1)}}-1)}\right]\cdot e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_{(i-1)}}-1)}-1\right)}$

$$\frac{\partial L}{\partial \lambda}=\frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{\Omega_i}\left[K_1(x_{(i)};\theta,\gamma,\lambda)-K_2(x_{(i-1)};\theta,\gamma,\lambda)\right] \tag{17}$$

$$\frac{\partial L}{\partial \lambda}=\theta x e^{\lambda x_i}\cdot e^{\gamma(e^{\lambda x_i}-1)}\cdot e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right)}$$

where $K_1(x_{(i)};\theta,\gamma,\lambda)=\theta x e^{\lambda x_i}\cdot e^{\gamma(e^{\lambda x_i}-1)}\cdot e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_i}-1)}-1\right)}$

and $K_2(x_{(i-1)};\theta,\gamma,\lambda)=\theta x e^{\lambda x_{(i-1)}}\cdot e^{\gamma(e^{\lambda x_{(i-1)}}-1)}\cdot e^{-\frac{\theta}{\gamma}\left(e^{\gamma(e^{\lambda x_{(i-1)}}-1)}-1\right)}$

The MPS are obtained by setting equations (15), (16), and (17) to zero and solving these equations simultaneously. Thus, these cannot be solved analytically, necessitating the use of analytical tools to solve them numerically.

3. Results and Discussion

This section evaluates the reliability of the OGED through Monte Carlo simulations, where the parameter estimates, their biases, and root mean square errors are obtained using the MLE and MPS methods.

3.1 Simulation Analysis

The quantile function of the OGE distribution defined in equation (3) was used to generate 1000 replicates with different sample sizes of n = 20, 50, 100, 250, 500, and 1000 from the OGE distribution at the chosen values $(\theta,\gamma,\lambda)=(0.5,1.5,0.5)$. From the 1000 replicates, the estimates, bias, and RMSE were computed and presented in Table 3.1

Table 3.1: Result of the simulated data from the OGED using MLE and MPS

	MLE			MPS		
Sample size	Estimates	Bias	RMSE	Estimates	Bias	RMSE
20	0.4874	0.0126	0.2599	0.6449	0.1449	0.3578
	1.5449	0.0449	0.3251	1.4654	-0.0346	0.3742
	0.5287	0.0287	0.0790	0.4786	-0.0214	0.0746
50	0.5049	0.0049	0.1704	0.5883	0.0883	0.2162
	1.5438	0.0438	0.2686	1.5391	0.0391	0.2930
	0.5065	0.0065	0.0518	0.4766	-0.0234	0.0558
100	0.5069	0.0069	0.1237	0.5554	0.0554	0.1450
	1.5426	0.0426	0.2308	1.5528	0.0528	0.2320
	0.4942	0.0058	0.0385	0.4800	-0.0200	0.0415
250	0.5067	0.0067	0.0802	0.5301	0.0301	0.0887
	1.5433	0.0433	0.1952	1.5613	0.0613	0.1893
	0.4956	0.0044	0.0319	0.4847	-0.0153	0.0331
500	0.5066	0.0066	0.0572	0.5231	0.0231	0.0633
	1.5420	0.0420	0.1536	1.5559	0.0559	0.1597
	0.4947	0.0053	0.0261	0.4871	-0.0129	0.0287
1000	0.5064	0.0064	0.0434	0.5174	0.0174	0.0467
	1.5342	0.0342	0.1322	1.5514	0.0514	0.1412
	0.4953	0.0047	0.0225	0.4893	-0.0107	0.0242

Table 3.1 presents the results obtained from the Monte Carlo simulation study of the OGE distribution. The results revealed that both estimation techniques were consistent, as the Bias and RMSE decreased at different sample sizes of 20 to 1000. However, for all sample sizes and different real values of parameters, the MLE proved to be a better estimator than MPS because the estimates converged more to the true parameter values and had a lower RMSE across all sample sizes.

4. Conclusion and Recommendation

This study has clearly demonstrated the value of using simulation-based approaches to assess the performance of parameter estimation methods, particularly within compound probability models such as the OGED. Through extensive Monte Carlo simulations, the study compared the efficiency of MLE and MPS, thereby providing robust evidence of their relative strengths. The findings reveal that both MLE and MPS are consistent estimators, as shown by the steady decline in bias and RMSE with increasing sample size. This consistency implies that both methods are reliable in the long run, converging towards the true parameter values as more data become available. However, MLE achieved lower RMSE values and faster convergence to the true parameters compared to MPS. This indicates that MLE not only provides more accurate estimates but also uses available data more efficiently, delivering precise results with relatively fewer deviations. Given the superior performance of MLE in this study, it is recommended as the preferred method for estimating the parameters of the OGED, particularly in situations where computational resources and sample sizes are adequate. Nevertheless, MPS remains a feasible alternative, especially in scenarios where the likelihood function is difficult to optimize or suffers from numerical instability.

REFERENCES

Abdulsalam, H. A., Bashiru, S. O., Isa, A. M., & Ojirobe, Y. A. (2021). Maximum Product of Spacing Parameter Estimation of Gompertz Rayleigh Distribution and Application to Rainfall Datasets. *Asian Journal of Probability and Statistics*, 14(2):41-59.

Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). *The design of simulation studies in medical statistics*. *Statistics in Medicine,* 25(24), 4279–4292.

Carpenter, J. R., Kenward, M. G., & White, I. R. (2007). Sensitivity analysis after multiple imputation under missing at random: A weighting approach. *Statistical Methods in Medical Research,* 16(3), 259–275.

Casella, G. and Berger, R. (2002). Statistical Inference. Pacific Grove, CA: Duxbury, 2nd edition.

Cheng, R. C. H., & Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society*: Series B, 45(3), 394–403.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics.

Habu, L., Usman, A., Sadiq, I. A., & Abdullahi, U. A. (2024). Estimation of Extension of Topp- Leone Distribution using Two Different Methods: Maximum Product Spacing and Maximum Likelihood Estimate. *UMYU Scientifica,* 3(2), 133-138.

Kajuru, J. Y., Dikko, H. G., Mohammed, A. S., and Fulatan, A. I. (2023). Odd Gompertz-G family of distribution, its properties and applications. *Fudma Journal of Sciences,* 7(3):351-358.

Kajuru, J. Y. (2025). Development of Odd Gompertz-G Families for Generating Compound Distributions: Properties and Applications. (Doctoral dissertation, Ahmadu Bello University, Zaria).

Kendall, M., & Stuart, A. (1979). *The Advanced Theory of Statistics* (Vol. 2). Charles Griffin.

Kozubowski, T. J., & Panorska, A. K. (2005). *A note on the estimation of the Laplace distribution*. *Communications in Statistics Simulation and Computation,* 34(4), 1069–1078.

Lemonte, A. J. (2013). A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. *Computational Statistics Data Analysis,* 62, 149-170.

Maurya, S. K., Kaushik, A., Singh, S. K., & Singh, U. (2016). A new class of distribution having decreasing, increasing, and bathtub-shaped failure rates. *Communications in Statistics - Theory and Methods,* 46(20), 10359–10372.

Oguntunde, P. E. (2017). Generalisation of the Inverse Exponential Distribution: Statistical Properties and Applications (Doctoral dissertation, Covenant University).

Ranneby, B. (1984). The maximum spacing method: An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, 11(2), 93–112.

Rasheed, B. A., Adnan, R., Saffari, S. E., & DanoPatil, K. (2014). Robust weighted least squares estimation of regression parameters in the presence of outliers and heteroscedastic errors. *Jurnal Teknologi (Sciences & Engineering)*, *71*(1).

Singh, U., Singh, S. K., & Singh, R. K. (2014). A comparative study of traditional estimation methods and the maximum product spacing method in generalized inverted exponential distribution. *Journal of Statistics Applications & Probability*, 3(2), 153.