THE NEW TWO-PARAMETER EXPONENTIAL-GAMMA-BASED DISTRIBUTION: PROPERTIES AND APPLICATION TO NON-COMMUNICABLE DISEASE DATA

B. I. Sanni [(1)]
S. A. Aderoju [(1,2*)]
R. K. Lamidi [(1)]
A. D. Adeshola [(1)]
A. K. Jimoh [(1)]
E. J. Adeniyi [(1)]

Received: 06/09/2025
Revised: 14/10/2025
Accepted: 15/10/2025

[1] Department of Mathematics and Statistics, Kwara State University, Malete, P.M.B. 1530, Ilorin, Nigeria.
[2] Department of Mathematics and Statistics, Kwara State University, Malete, Ilorin, Nigeria.
*Corresponding Author's Email: Samuel.aderoju@kwasu.edu.ng

THE NEW TWO-PARAMETER EXPONENTIAL-GAMMA-BASED DISTRIBUTION: PROPERTIES AND APPLICATION TO NON-COMMUNICABLE DISEASE DATA

Bello Ishola Sanni
Department of Mathematics and Statistics, Kwara State University,
Malete, , Ilorin, Nigeria.

Samuel Adewale Aderoju
Department of Mathematics and Statistics, Kwara State University,
Malete, , Ilorin, Nigeria.
Statistics Program, BOWEN University,
Iwo, Osun State, Nigeria
Samuel.aderoju@kwasu.edu.ng

Rasheed Kehinde Lamidi
Department of Mathematics and Statistics, Kwara State University,
Malete, Ilorin, Nigeria.

Adediran Dauda Adeshola
Department of Mathematics and Statistics, Kwara State University,
Malete, Ilorin, Nigeria.

Abdulazeez Kayode Jimoh
Department of Mathematics and Statistics, Kwara State University,
Malete, Ilorin, Nigeria.

Ebenezer Juwon Adeniyi
Department of Mathematics and Statistics, Kwara State University,
Malete, Ilorin, Nigeria.

Abstract— This study introduces a novel lifetime distribution, termed the New Two-Parameter Exponential–Gamma (NTPEG) distribution, developed through an additive mixture framework to enhance flexibility in modeling lifetime and survival data. Fundamental structural properties, including moments, hazard rate behavior, and entropy, are derived and discussed. Parameters of the model are estimated via the maximum likelihood method, and simulation studies demonstrate the reliability and efficiency of the estimators. The practical utility of the NTPEG model is illustrated using survival data on non-communicable diseases, particularly among acute bone cancer patients. Comparative analyses with existing extensions of the Lindley distribution reveal that the proposed model provides superior fit and flexibility for complex real-world data. Overall, the NTPEG distribution represents a robust and versatile tool for researchers and practitioners in survival and reliability analysis.

Keywords— Additive mixture; Exponential-Gamma distribution; maximum likelihood estimation; Monte Carlo simulation; non-communicable diseases.

I. INTRODUCTION

A paramount application of statistical methodology lies within health research and its applications [1]. The analysis of cancer data, in particular, often necessitates specialized statistical models to accurately represent survival outcomes. In response to this need, the field of distribution theory has focused on the challenge of developing a diverse suite of probability models to facilitate more robust analysis and exploration of lifetime data [2]. This endeavor has led to the creation, examination, and extension of numerous useful statistical distributions.

The primary impetus for generalizing existing models and developing new distributional classes is to provide a more accurate mathematical representation of lifetime phenomena across numerous disciplines, including public health, medicine, engineering, and biology. Foundational distributions such as the Exponential, Weibull, and Gamma, while fundamental, exhibit limited flexibility for modelling complex real-world data [2]. For instance, while the Exponential distribution is characterized by a constant hazard function, the Weibull distribution can accommodate increasing, decreasing, or constant hazard rates. However, its structural form is incapable of modelling non-monotonic failure rates, such as the bathtub-shaped or unimodal curves frequently observed in practice. Similarly, the Gamma distribution is often hindered by the lack of a closed-form solution for its cumulative distribution function, complicating the derivation of its mathematical properties.

Complex phenomena in areas like human mortality, reliability engineering, and biological studies often demonstrate failure rates that are not monotonic. To address this critical limitation, statisticians have developed several generalized distribution families specifically designed to model both monotonic and non-monotonic hazard rate functions.

Since no single classical distribution can adequately model all datasets from every field, there is growing scholarly interest in constructing more flexible models. This is primarily achieved through two approaches: (i) creating finite mixtures of existing distributions, such as the Lindley, generalized Lindley, and Samade distributions (see [3]–[6]); or (ii) introducing new shape parameters into a baseline model to increase its flexibility, leading to distributions such as the Power Lindley, weighted Lindley, Power Samade, Power Hamza, and Power Generalized Akash (see [7]–[12]).

Among these, the Lindley distribution, originally introduced by Lindley [3] in a fiducial context, has garnered significant interest. It is defined as a two-component mixture of an exponential and a length-biased exponential distribution. Its probability density function (pdf) is given by:

$$f(x;\theta) = (1-P)g_1(x;\theta) + Pg_2(x;\theta)$$

$$= \frac{\theta^2}{(1+\theta)}(1+x)e^{-\theta x}, \quad for\ x, \theta > 0, \qquad (1)$$

where $p = \frac{1}{1+\theta}$ is the mixing proportion while $g_1(x;\theta) = \theta^2 x e^{-\theta x}$ and $g_2(x;\theta) = \theta e^{-\theta x}$

The Lindley distribution serves as a valuable model for analyzing lifetime and reliability data. A comprehensive analysis of its properties was conducted by Ghitany et al. [13], who demonstrated its superiority over the exponential distribution in certain applications, albeit noting its limitations due to its single-parameter formulation. Further investigations by Mazucheli and Achcar [14] explored its utility in modelling competing risks lifetime data. Deniz and

Ojeda [15] introduced a discrete counterpart to this distribution, expanding its relevance to count data analysis in fields such as insurance. The distribution's flexibility has been enhanced through various extensions; for instance, Bakouch et al. [16] derived an extended Lindley distribution and elaborated on its mathematical properties, while Ghitany et al. [17] developed a two-parameter weighted Lindley distribution for application to survival data. The work of Zakerzadeh and Dolati [18] led to a generalized Lindley distribution, broadening the model's applicability. These efforts have spurred the development of numerous other generalizations, as documented in [5], [17], and [18].

This trend of extending lifetime models continues to be an active area of research. For example, Aderoju and Babaniyi [9] built upon the Samade distribution [19] to propose the Power Samade distribution. Similarly, Aderoju and Jolayemi [10] formulated the Power Hamza distribution as an extension of the Hamza distribution, which was previously developed by Aijaz et al. [20]. Other significant contributions include the New Generalized Gamma-Weibull distribution by Aleshinloye et al. [21] and the Power Generalized Akash distribution introduced by Aderoju and Adeniyi [22]. Beyond continuous lifetime modelling, these distributions have been employed as mixing distributions in Poisson mixtures to create novel models for count data. Notable examples include the Poisson-Samade distribution [12], a New Generalized Poisson Mixed Distribution [6], and the zero-inflated generalized Poisson-Sujatha distribution [23].

The principal objective of the present study is to introduce the new two-parameter Exponential-Gamma (NTPEG) distribution and to elucidate its statistical properties and practical applications, particularly within the realm of survival analysis. The remainder of this article is organized as follows: Section 2 formally introduces the NTPEG distribution. Its key statistical properties are derived and examined in Section 3. Section 4 details the method of parameter estimation via maximum likelihood. Practical applications of the model are presented in Section 5. Finally, Section 6 provides concluding remarks.

II. The New Two-parameter Exponential-Gamma (NTPEG) Distribution

This study employed the kth component additive mixture distribution method for a random variable X is defined as (see [24] for more):

$$f(x,\theta) = \sum_{j=1}^{k} h_j(x;\theta)p_j \,, \qquad (2)$$

where $\sum_{j=1}^{k} p_j = 1$ and θ represents the parameter vector of the component distribution and p_j denotes the corresponding mixing weight.

Theorem 1. If a random variable X is distributed according to the NTPEG distribution with parameters α and θ, denoted by $X|\alpha,\theta \sim \text{NTPEG}(\alpha,\theta)$, its pdf is expressed as follows:

$$f(X|\alpha,\theta) = \frac{\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}x^{\alpha-1}\right)e^{-\theta x}, \quad for\ x,\theta,\alpha > 0 \qquad (3)$$

Proof 1: Let a random variable X follow a two-component mixture of exponential (θ) {\displaystyle f(x;\lambda)={\begin{cases}\lambda e^{-\lambda x}&x\geq 0,\\0&x<0.\end{cases}}}and gamma (α,θ) distributions with the mixture proportion, $p_1 = \frac{\theta^3}{\theta^3+6}$ and the pdf expressed as

$$f(X;\alpha,\theta) = p_1 h_1(x;\theta) + (1-p_1)h_2(x;\alpha,\theta) \qquad (4)$$

Where

$$h_1(x;\theta) = \theta e^{-\theta x},$$

$$h_2(x;\alpha,\theta) = \frac{\theta^\alpha x^{\alpha-1}e^{-\theta x}}{\Gamma(\alpha)}$$

So,

$$f(x;\alpha,\theta) = \left[\frac{\theta^3}{\theta^3+6}\left(\theta e^{-\theta x}\right)\right] + \left[\frac{6}{\theta^3+6}\left(\frac{\theta^\alpha x^{\alpha-1}e^{-\theta x}}{\Gamma(\alpha)}\right)\right]$$

Consequently, the probability density function for the proposed NTPEG distribution is given by

$$f(x;\alpha,\theta) = \frac{\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}x^{\alpha-1}}{\Gamma(\alpha)}\right)e^{-\theta x}$$

Figure 1 presents the pdf of the NTPEG distribution under different parameter configurations. Panels (a) through (d) demonstrate how the shape and spread of the distribution evolve as the shape parameter $\hat{\alpha}$ and the scale parameter $\hat{\theta}$ vary. Panel (a): when $\hat{\alpha} = 0.25$, the distribution exhibits heavy-tailed behaviour, especially for smaller values of $\hat{\theta}$. As $\hat{\theta}$ increases, the distribution becomes more concentrated around smaller x-values, resulting in a rapid decay of the pdf. Panel (b): for $\hat{\alpha} = 5$, the pdf shows unimodal behaviour, with peaks shifting rightward as $\hat{\theta}$ increases. Higher $\hat{\theta}$ lead to broader distributions with lower peak heights. Panel (c): when $\hat{\alpha} = 10$, the pdf retains unimodal characteristics, but the peak sharpens and shifts further to the right with increasing $\hat{\theta}$. This demonstrates the effect of $\hat{\theta}$ in spreading the distribution over a wider range of x-values. Panel (d): A finer resolution of $\hat{\theta}$ values are considered at $\hat{\alpha} = 10$. As $\hat{\theta}$ increases from 0.15 to 0.45, the peak decreases, and the distribution becomes more dispersed, reflecting increased variance.

These plots illustrate the flexibility of the NTPEG distribution in modelling data with different skewness, tail behaviour, and peak characteristics, making it suitable for diverse real-world applications. Moreover, it is imperative to note the following from the plots in Figure 1: Panel (a): The probability density functions (pdfs) for $\hat{\alpha} = 0.25$ exhibits a monotonically decreasing shape, indicating an unimodal distribution with the peak at $x = 0$. Panel (b), (c), and (d): For larger values of $\hat{\alpha}$ $(e.g.,\ \hat{\alpha} = 5\ or\ \hat{\alpha} = 10$, the pdf displays bimodal behaviour. The curves initially rise to a peak, decrease, and then form a secondary hump, suggesting the presence of two modes.

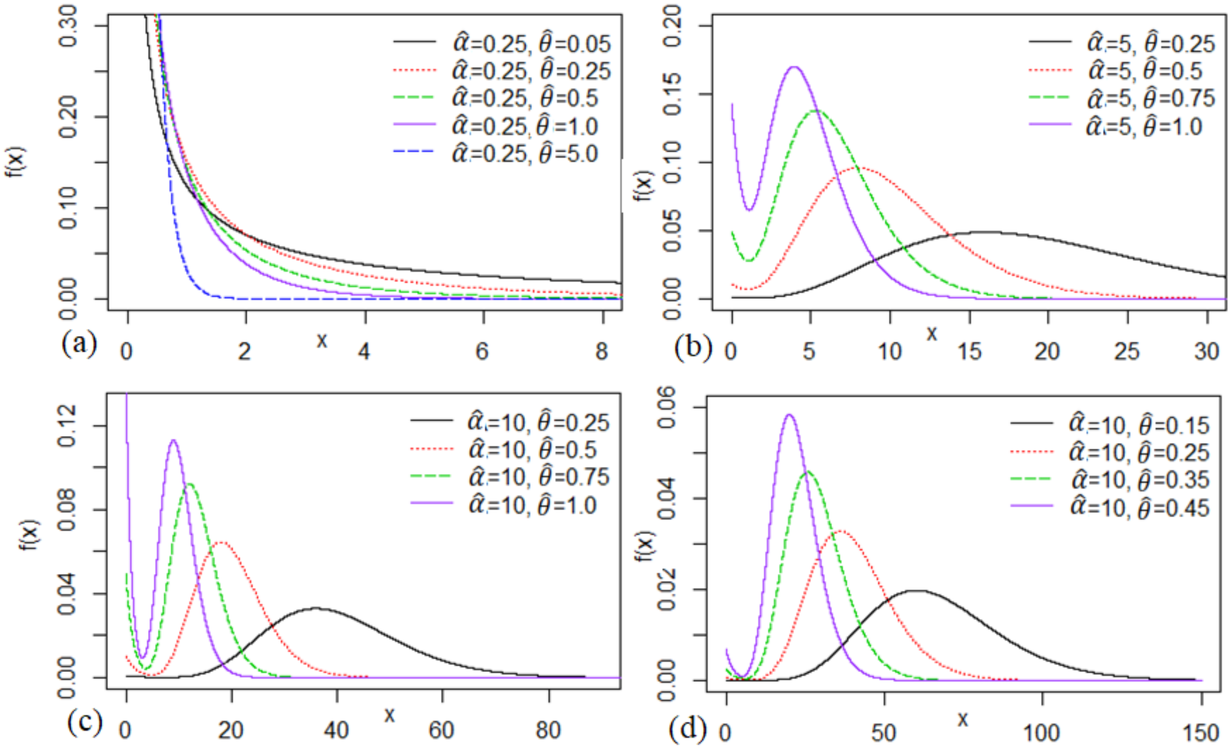

Figure 1. Probability density functions of the proposed NTPEG distribution under different parameter combinations

The cumulative distribution function (cdf) for the proposed NTPEG distribution is derived as follows

$$F(x;\alpha,\theta) = \int_{t=0}^{x} f(t|\theta,\alpha,)\,\mathrm{d}t$$

$$= \int_{t=0}^{x} \frac{\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}t^{\alpha-1}\right)e^{-\theta t}\,\mathrm{d}t$$

$$= \frac{(1-\mathrm{e}^{-x\theta})\theta^3 + \dfrac{6x^{\alpha}\theta^{\alpha}(x\theta)^{-\alpha}(\Gamma(\alpha)-\Gamma(\alpha,x\theta))}{\Gamma(\alpha)}}{6+\theta^3} \quad (5)$$

Figure 2 illustrates the cdf corresponding to different parameter settings of the NTPEG distribution. Panels (a) and (b) demonstrate how the cdf evolves as $\hat{\alpha}$ and the $\hat{\theta}$ change. In panel (a): When $\hat{\alpha} = 10$, the cdf transitions from a gradual increase to a more abrupt step-like shape as $\hat{\theta}$ increases. Higher values of $\hat{\theta}$ correspond to a faster accumulation of probability mass, reflecting a shift toward the left. This behaviour highlights the influence of the scale parameter, $\hat{\theta}$ in controlling the spread and central tendency of the distribution. Similarly, in panel (b), for $\hat{\alpha} = 0.5$, the cdf displays a smoother, more gradual ascent. As $\hat{\theta}$ increases from 0.05 to 1.0, the cdf approaches 1 more rapidly, indicating faster convergence to the upper bound. The curves reveal that smaller $\hat{\alpha}$-values lead to heavier tails and slower accumulation of probability mass at lower x-values.

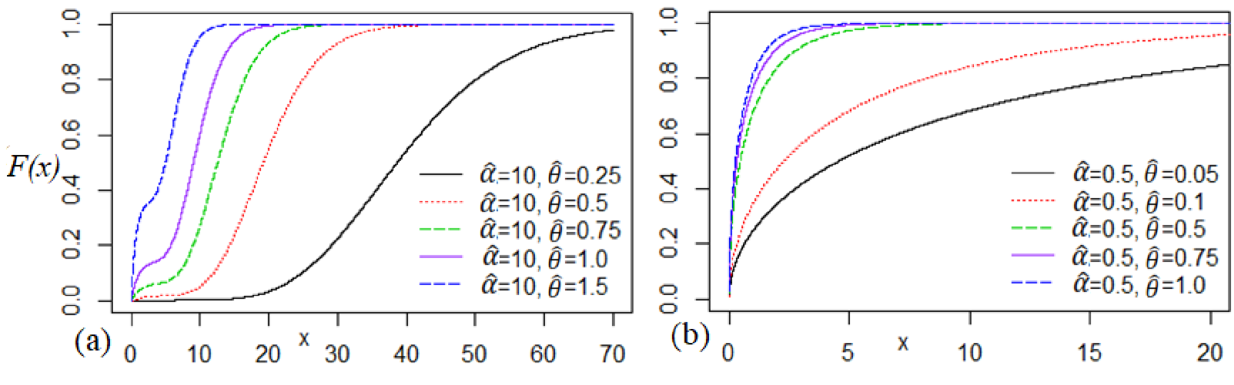

Figure 2. Cumulative distribution function of the NTPEG model under different parameter values

III. MOMENTS AND OTHER MATHEMATICAL PROPERTIES

To derive the rth factorial moment for a random variable X following the NTPEG distribution, we proceed as follows:

$$\mu_r = E(x^r) = \int_0^\infty x^r f(x)dx$$

Therefore,

$$\mu_r' = \int_{x=0}^{\infty} x^r \frac{\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}x^{\alpha-1}}{\Gamma(\alpha)}\right)e^{-\theta x}dx$$

Hence,

$$\mu_r' = \frac{\theta^4}{\theta^3+6}\left[\left(\left(\frac{\Gamma(r+1)}{\theta^{r+1}}\right)+\left\{\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}\left(\frac{\Gamma(r+\alpha)}{\theta^{r+\alpha}}\right)\right\}\right)\right] \quad (6)$$

The initial four raw moments of the distribution are derived as follows:

$$\mu_1' = \frac{6\alpha+\theta^3}{6\theta+\theta^4}$$

$$\mu_2' = \frac{2(3\alpha+3\alpha^2+\theta^3)}{\theta^2(6+\theta^3)}$$

$$\mu_3' = \frac{6(\alpha(1+\alpha)(2+\alpha)+\theta^3)}{\theta^3(6+\theta^3)}$$

$$\mu_4' = \frac{6(4\theta^3\Gamma(\alpha)+\Gamma(4+\alpha))}{\theta^4(6+\theta^3)\Gamma(\alpha)}$$

Key distributional properties were analyzed, including variance (σ^2), coefficient of variation (CV), skewness (γ_1), and kurtosis (γ_2).

$$\sigma^2 = \mu_2' - (\mu_1')^2 = \frac{6\alpha^2\theta^3 - 6\alpha(\theta^3-6)+\theta^3(\theta^3+12)}{\theta^2(6+\theta^3)^2} \quad (7)$$

$$\text{CV} = \frac{\sigma}{\mu_1}$$

$$= \frac{\sqrt{(36\alpha+12\theta^3-6\alpha\theta^3+6\alpha^2\theta^3+\theta^6)}}{(6\alpha+\theta^3)(6\theta+\theta^4)} \quad (8)$$

$$\gamma_1 = \frac{6(\alpha^2(1+\alpha)(2+\alpha)+\theta^3)}{(6\alpha\theta^3+\theta^6)\left(\frac{12\alpha\theta^3-6\alpha^2\theta^3+\theta^6+6\alpha^3(6+\theta^3)}{(6\alpha\theta+\theta^4)^2}\right)^{3/2}} \quad (9)$$

$$\gamma_2 = \frac{6(6\alpha+\theta^3)^3\left(4\theta^3\Gamma(\alpha)+\alpha\Gamma(4+\alpha)\right)}{\left(12\alpha\theta^3-6\alpha^2\theta^3+\theta^6+6\alpha^3(6+\theta^3)\right)^2\Gamma(\alpha)} \quad (10)$$

A. *Reliability Analysis for the NTPEG distribution*

Reliability measures, such as the survival function and the hazard (failure rate) function, are fundamental for characterizing the NTPEG distribution. These functions provide critical insight into its behavioral properties, reliability patterns, and lifespan characteristics. The survival function, denoted as $S(t)$, defines the probability that a random variable X following the NTPEG distribution exceeds a specific time t, i.e., $S(x) = P(T > x)$. For the NTPEG distribution, this function is derived as follows:

$$S(x;\theta,\alpha) = 1 - F(x;\theta,\alpha)$$

$$S(x;\theta,\alpha) = 1 - \frac{(1-e^{-x\theta})\theta^3 + \frac{6x^\alpha\theta^\alpha(x\theta)^{-\alpha}(\Gamma(\alpha)-\Gamma(\alpha,x\theta))}{\Gamma(\alpha)}}{6+\theta^3} \quad (11)$$

The hazard function is obtained as follows:

$$H(x;\theta,\alpha) = \frac{f(x;\theta,\alpha,x)}{1-F(x;\theta,\alpha,x)} = \frac{f(x;\theta,\alpha,x)}{S(x;\theta,\alpha,x)}$$

$$H(x) = \frac{(x\theta)^\alpha(6x^\alpha\theta^\alpha + x\theta^4\Gamma(\alpha))}{x\left((\theta^3(x\theta)^\alpha + e^{x\theta}(-6x^\alpha\theta^\alpha+6(x\theta)^\alpha))\Gamma(\alpha)+6e^{x\theta}x^\alpha\theta^\alpha\Gamma(\alpha,x\theta)\right)} \quad (12)$$

Figure 3 represents the survival function plots, which illustrate the effect of varying the scale parameter $\hat{\theta}$ while keeping the shape parameter fixed at $\hat{\alpha} = 0.5$. As $\hat{\theta}$ increases from 0.05 to 1.0, the survival probability decays more rapidly, reflecting a faster decline in the likelihood of extreme values. This indicates that higher scale parameter values lead to distributions with lighter tails, emphasizing a more rapid concentration of mass around smaller x-values. For lower $\hat{\theta}$-values, the survival function retains a heavy tail, suggesting a greater probability of larger x-values. This behaviour is characteristic of distributions that model long-term survival or rare events. Figure 3 highlights the flexibility of the proposed model in capturing different survival patterns, making it suitable for applications involving lifetime data, reliability analysis, and event-time modelling.

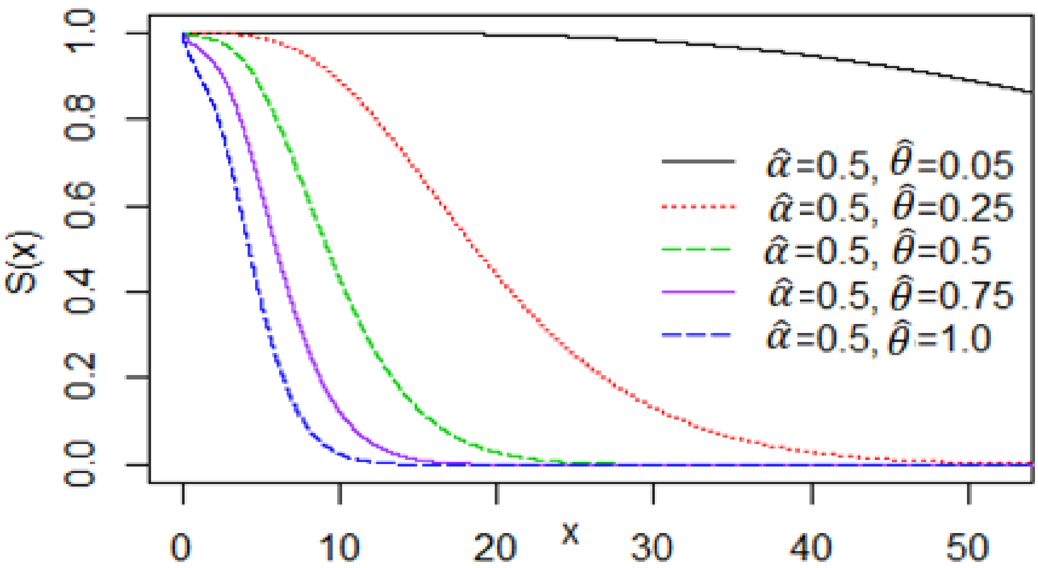

Figure 3: The Survival Function, S(x), plot for Varying Parameter Values of the NTPEG Distribution

Figure 4 presents hazard functions, $h(x)$, for different parameter values of $\hat{\alpha}$ and $\hat{\theta}$. Panel (a): The hazard functions correspond to $\hat{\alpha} = 0.5$ with varying $\hat{\theta}$. All curves exhibit a monotonically decreasing pattern, which suggests a decreasing hazard rate over time. This is characteristic of distributions where the risk of failure (or event occurrence) decreases as time progresses, often observed in early-life failures or infant mortality models. Panel (b): The hazard functions correspond to $\hat{\alpha} = 5$ with different $\hat{\theta}$. Here, the curves exhibit a non-monotonic behaviour, starting with a U-shaped or increasing trend. This pattern suggests an initially decreasing hazard rate followed by an increasing hazard, which is typical in aging or wear-out processes where the risk of failure initially declines and then rises over time. It is important to note that:

i. Higher values of $\hat{\theta}$ lead to higher hazard rates across both panels.
ii. The change in $\hat{\alpha}$ significantly alters the shape of the hazard function, transitioning from a monotonically decreasing hazard ($\hat{\alpha} = 0.5$) to a bathtub-shaped or increasing hazard ($\hat{\alpha} = 5$).
iii. These trends are indicative of different underlying survival distributions, possibly a Weibull or generalized gamma model, depending on the parameterization.

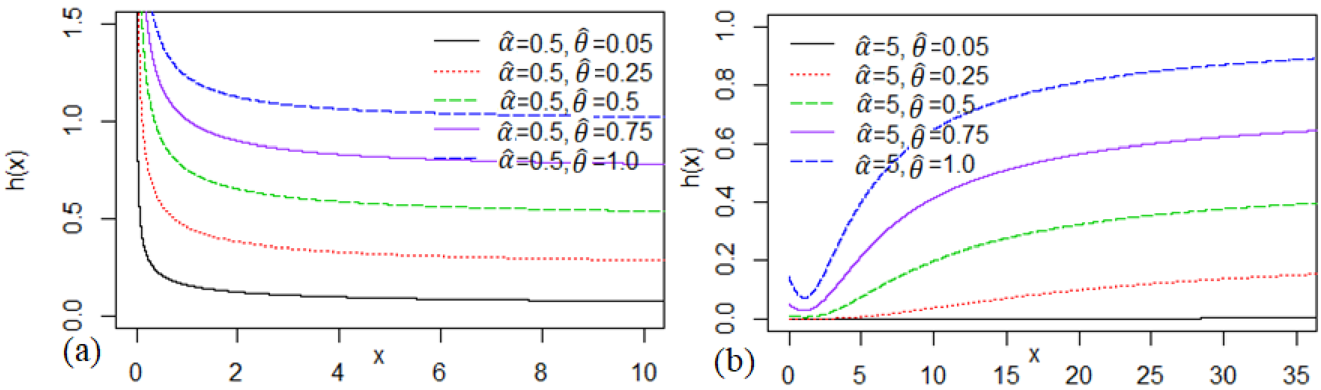

Figure 4: The hazard function of the NTPEG distribution

B. Order Statistics

The theory of order statistics originated from the need to model real-world phenomena that involve the analysis of extreme values, such as minima and maxima, as detailed in [25, 26]. This section derives and examines several properties of the order statistics for the proposed NTPEG distribution.

The probability density function for the i^{th} order statistic, denoted by $f_{x(i)}(x)$, is given by the following expression:

$$f_{x(i)}(x) = \frac{n!}{(i-1)!\,(n-i)!} f(x)[F(x)]^{i-1}[1-F(x)]^{n-i} \quad (13)$$

Consequently, the expression for the i^{th} order statistic from the NTPEG distribution is derived as follows:

https://doi.org/10.20428/jst.v30i12.3314

$$f_{x(i)}(x) = \frac{n!}{(i-1)!\,(n-i)!}\frac{\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}x^{\alpha-1}\right)e^{-\theta x}\left[\frac{(1-e^{-x\theta})\theta^3+\frac{6x^\alpha\theta^\alpha(x\theta)^{-\alpha}(\Gamma(\alpha)-\Gamma(\alpha,x\theta))}{\Gamma(\alpha)}}{6+\theta^3}\right]^{i-1}\left[1-\frac{(1-e^{-x\theta})\theta^3+\frac{6x^\alpha\theta^\alpha(x\theta)^{-\alpha}(\Gamma(\alpha)-\Gamma(\alpha,x\theta))}{\Gamma(\alpha)}}{6+\theta^3}\right]^{n-i},$$

For a sample size $n > 0$, the probability density functions for the minimum (first) and maximum (nth) order statistics of the NTPEG distribution are derived as follows:

$$f_{x(1)}(x) = \frac{n\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}x^{\alpha-1}\right)e^{-\theta x}\left[1-\frac{(1-e^{-x\theta})\theta^3+\frac{6x^\alpha\theta^\alpha(x\theta)^{-\alpha}(\Gamma(\alpha)-\Gamma(\alpha,x\theta))}{\Gamma(\alpha)}}{6+\theta^3}\right]^{n-1} \quad (14)$$

and

$$f_{x(n)}(x) = \frac{n\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}x^{\alpha-1}\right)e^{-\theta x}\left[\frac{(1-e^{-x\theta})\theta^3+\frac{6x^\alpha\theta^\alpha(x\theta)^{-\alpha}(\Gamma(\alpha)-\Gamma(\alpha,x\theta))}{\Gamma(\alpha)}}{6+\theta^3}\right]^{n-1} \quad (15)$$

C. Rényi's Entropy

Rényi entropy [27] provides a generalized measure of the uncertainty associated with a random variable *X*. For a given parameter, it is defined as:

$$RE_p = \frac{1}{1-p}\log\int_0^\infty [f(x;\alpha,\theta]^p dx, \qquad p>0, p\neq 1 \quad (16)$$

$$RE_p = \frac{1}{1-p}\log\int_0^\infty\left[\frac{\theta^4}{\theta^3+6}\left(1+\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}x^{\alpha-1}\right)e^{-\theta x}\right]^p dx$$

$$=\frac{1}{1-p}\log\left[\frac{\theta^{4p}}{(\theta^3+6)^p}\int_0^\infty\left(1+\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}x^{\alpha-1}\right)^p e^{-\theta px}\,dx\right]$$

Recall Taylor's series expansion of $(1+x)^p = \sum_{k=0}^{p}\binom{p}{k}x^k$

RE_p

$$=\frac{1}{1-p}\log\left[\frac{\theta^{4p}}{(\theta^3+6)^p}\int_0^\infty\sum_{k=0}^{p}\binom{p}{k}\left(\frac{6\theta^{\alpha-4}}{\Gamma(\alpha)}x^{\alpha-1}\right)^k e^{-\theta px}\,dx\right]$$

$$=\frac{1}{1-p}\log\left[\frac{\theta^{4p}}{(\theta^3+6)^p}\sum_{k=0}^{p}\binom{p}{k}\frac{(6\theta^{\alpha-4})^k}{(\Gamma(\alpha))^k}\int_0^\infty(x)^{(\alpha-4)k}e^{-\theta px}\,dx\right]$$

$$Note\ that,\ \int_0^\infty(x)^{(\alpha-4)k}e^{-\theta px}\,dx = \frac{\Gamma(\alpha k-k+4)}{(\theta p)^{\alpha k-k+4}}$$

$\therefore RE_p$

$$=\frac{1}{1-p}\log\left[\frac{\theta^{4p}}{(\theta^3+6)^p}\sum_{k=0}^{p}\binom{p}{k}\frac{(6\theta^{\alpha-4})^k}{(\Gamma(\alpha))^k}\frac{\Gamma(\alpha k-k+4)}{(\theta p)^{\alpha k-k+4}},\right]$$

the following expression is obtained for the Rényi entropy of the proposed NTPEG distribution:

$$=\frac{1}{1-p}\log\left[\frac{\theta^{4p}}{(\theta^3+6)^p}\sum_{k=0}^{p}\binom{p}{k}\frac{(\theta^{-(3k+4)})6^k}{(\Gamma(\alpha))^k}\frac{\Gamma(\alpha k-k+4)}{(p)^{\alpha k-k+4}},\right]$$

IV. PARAMETER ESTIMATION

The estimation of unknown parameters from observed data is a fundamental task in statistical inference. Warahena-Liyanage and colleagues [2] conducted a comparative analysis of six prevalent estimation techniques: Anderson–Darling (AD), maximum product of spacings (MPS), weighted least squares (WLS), least squares (LS), and the minimum distance estimators Cramér–von Mises (CVM) and maximum likelihood (MLE). Their simulation study, based on Monte Carlo methods, demonstrated the superior performance of the MLE approach in obtaining parameter estimates. This finding aligns with established statistical literature, justifying our adoption of the maximum likelihood method for the present analysis.

Let $\Phi = (\alpha,\theta)$ be the parameter vector of the proposed model. For a given random sample of size n, denoted by the order statistics $X_{(1)}, X_{(2)}, \dots X_{(n)}$, drawn from the NTPEG distribution, the corresponding log-likelihood function can be formulated as follows:

$$\ell_n(\Phi) = 4n\log(\theta) - n\log(\theta^3+6) + \sum_{i=1}^{n}\log(\Gamma(\alpha) + 6\theta^{\alpha-4}x_i^{\alpha-1}) - n\log\Gamma(\alpha) - \theta\sum_{i=1}^{n}x_i \quad (17)$$

Differentiating partially with respect to the relevant parameters gives the following expressions:

$$\frac{\partial \ell_n(\Phi)}{\partial \theta} = \frac{n(24+\theta^3)}{\theta(6+\theta^3)} - \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \frac{6(\alpha-4)\theta^{\alpha-5}x_1^{\alpha-1}}{\Gamma(\alpha)+6\theta^{\alpha-4}x_i^{\alpha-1}} \quad (18)$$

$$\frac{\partial \ell_n(\Phi)}{\partial \alpha} = -\log n\Gamma(\alpha)\Gamma(0,\alpha) + \sum_{i=1}^{n} \frac{\Gamma(\alpha)\Gamma(0,\alpha)+6\theta^{-4+\alpha}\mathrm{Log}(\theta)x_1^{-1+\alpha}+6\theta^{-4+\alpha}\mathrm{Log}(x_i)x_1^{-1+\alpha}}{\Gamma(\alpha)+6\theta^{-4+\alpha}x_i^{-1+\alpha}} \quad (19)$$

The maximum likelihood estimates, $\hat{\theta}$ and $\hat{\alpha}$, are solutions to the score equations derived from the likelihood function. The absence of an explicit analytical solution necessitates the use of numerical optimization techniques. In this study, the parameters are estimated by applying the Newton-Raphson algorithm to this nonlinear system.

A. Simulations' Studies.

A comprehensive simulation study was conducted to evaluate the performance of the MLEs for the parameters of the proposed NTPEG distribution. The simulation design involved generating samples of varying sizes (n = 50, 100, 150, 200, 250, 300, 350, 400, 450, 500) across different parameter combinations using the R statistical computing environment [28]. For each sample size, the process was replicated *N* = 1000 times using a fixed set of true parameter values, specifically α = 5.0 and θ = 0.5. The performance of the estimators, denoted as $\hat{\alpha}_{MLE}$ and $\hat{\theta}_{MLE}$, was assessed using two standard statistical criteria: the average bias and the Mean Squared Error (MSE). The results of this analysis are summarized in Table 1. The formulae for these performance metrics are given respectively by:

$$\mathrm{Bias}(\hat{\alpha}_{\mathrm{MLE}}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(\hat{\alpha}_{\mathrm{i}} - \alpha)$$

$$\mathrm{Bias}(\hat{\theta}_{\mathrm{MLE}}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(\hat{\theta}_{\mathrm{i}} - \theta)$$

and

$$\mathrm{MSE}(\hat{\alpha}_{\mathrm{MLE}}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(\hat{\alpha}_{\mathrm{i}} - \alpha)^2$$

$$\mathrm{MSE}(\hat{\theta}_{\mathrm{MLE}}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}(\hat{\theta}_{\mathrm{i}} - \theta)^2$$

Table 1: Simulation results for the NTPEG distribution parameter estimates with true values α = 5 and θ = 0.5

N	$\hat{\alpha}$	$\hat{\theta}$	$Bias_{\hat{\alpha}}$	$Bias_{\hat{\theta}}$	$MSE_{\hat{\alpha}}$	$MSE_{\hat{\theta}}$	$SD_{\hat{\alpha}}$	$SD_{\hat{\theta}}$
50	5.3323	0.5358	0.3323	0.0358	1.3525	0.0145	1.1151	0.1151
100	5.1893	0.5193	0.1893	0.0193	0.6044	0.0066	0.7544	0.0792
300	5.0346	0.5042	0.0346	0.0042	0.1794	0.0020	0.4224	0.0443
500	5.0201	0.5024	0.0201	0.0024	0.1046	0.0011	0.3229	0.0330
700	5.0207	0.5025	0.0207	0.0025	0.0757	0.0008	0.2745	0.0285
900	5.0138	0.5016	0.0138	0.0016	0.0574	0.0006	0.2394	0.0250
950	5.0155	0.5018	0.0155	0.0018	0.0546	0.0006	0.2333	0.0242
1000	5.0196	0.5016	0.0196	0.0016	0.0506	0.0005	0.2241	0.0230

A comprehensive Monte Carlo simulation study was conducted to assess the performance and asymptotic behavior of the Maximum Likelihood Estimators (MLEs) for the parameters of the NTPEG distribution. The study was performed across a range of sample sizes ($n = 50\ to\ 1000$) with N = 1000 replications for each size, using true parameter values of $\alpha = 5.5$ and $\beta = 1.5$. The estimation algorithm demonstrated excellent stability, with a success rate exceeding 99.6% across all sample sizes, indicating robust convergence of the numerical optimization routine. The key performance metrics, such as bias, mean squared error (MSE), and standard deviation (SD), are presented in Table 1. The results conclusively demonstrate the desirable properties of the MLEs:

i. **Consistency:** The estimators for both parameters exhibit strong consistency. As the sample size n increases, the mean estimates converge systematically toward their true values. This is evidenced by the monotonic decrease in bias. For instance, the bias for α reduced from 0.3323 ($n = 50$) to 0.0196 ($n = 1000$), and for β from 0.0358 to 0.0016. By $n = 450$, the bias for both parameters is negligible, effectively approaching zero.

ii. **Efficiency and Precision:** The Mean Squared Error (MSE) for both parameters decreases monotonically and substantially as the sample size grows. The MSE for α drops from 1.3525 ($n = 50$) to 0.0506 ($n = 1000$), representing a 96.3% reduction. Similarly, the MSE for β decreases from 0.0145 to 0.0005, a 96.6% reduction. This rapid decline in MSE confirms that the estimators are efficient, with estimation precision improving markedly with larger samples.

iii. **Asymptotic Normality:** The observed behavior of the estimators is consistent with the theoretical property of asymptotic normality. The standard deviation (SD) of the estimates, which represents the empirical standard error, decreases predictably as the sample size increases. The SD for α falls from 1.115 to 0.224, and for β from 0.115 to 0.023. This confirms that the sampling distributions of the estimators become more concentrated around the true parameter values as more data becomes available.

V. APPLICATIONS

This section evaluates the goodness-of-fit of the proposed two-parameter Exponential-Gamma distribution (NTPEGD) through its application to empirical datasets. The performance of the NTPEGD is compared against several established competing models using standard information criteria. These include the Akaike Information Criterion (AIC) [29], the Corrected Akaike Information Criterion (AICC) [30], the Hannan-Quinn Information Criterion (HQIC) [31], and the Bayesian Information Criterion (BIC) [32]. The competing models selected for this comparative analysis encompass a range of generalized Lindley distributions, namely:

- The New Generalized Two-Parameter Lindley Distribution (GLD1) by Ekhosuehi et al. [4],
- The Generalized Lindley Distribution (GLD2) by Abouammoh et al. [5],
- The Generalized Lindley Distribution (GLD3) by Shankar and Mishra [33],
- The Generalized Lindley Distribution (GLD4) by Nadarajah et al. [34],
- The Generalized Lindley Distribution (GLD5) by Zakerzadeh and Dolati [35].

The mathematical formulae for the aforementioned model selection criteria are provided below:

$$AIC = 2k - 2\ell_n(\Phi)$$

$$AICC = AIC + \frac{2k(k-1)}{n-k-1}$$

$$HQIC = 2\ell_n(\Phi)\log(n) - 2\ell_n(\Phi)$$

$$BIC = k\log(n) - 2\ell_n(\Phi)$$

Dataset 1: These data represent the remission times (in months) of a random sample of 128 bladder cancer patients reported by Lee and Wang [36].

Dataset 2: The data represent the survival times (in days) of 73 patients who were diagnosed with acute bone cancer, Mansour et al. [37].

Table 2: Parameter estimates and goodness-of-fit statistics for the proposed and competing models applied to Dataset 1

Models	**Parameter estimates**	**-2logL**	**AIC**	**AICC**	**HQIC**	**BIC**
NTPEGD	$\hat{\alpha} = 1.1726$ $\hat{\theta} = 0.1252$	**826.7358**	**830.7358**	**830.8318**	**833.0534**	**836.4399**
GLD1	$\hat{\alpha} = 1.1891$ $\hat{\theta} = 0.1247$	826.8288	830.8288	830.9248	833.1463	836.5328
GLD2	$\hat{\alpha} = 1.5690$ $\hat{\theta} = 0.1533$	832.3136	836.3136	836.4096	838.6312	842.0177
GLD3	$\hat{\alpha} = 663.69$ $\hat{\theta} = 0.1069$	828.6838	832.6839	832.7799	835.0015	838.3879
GLD4	$\hat{\alpha} = 0.7336$ $\hat{\theta} = 0.1648$	832.5718	836.5719	836.6679	838.8894	842.2759
GLD5	$\hat{\alpha} = 0.1734$ $\hat{\theta} = 0.1253$ $\hat{\beta} = 668.84$	826.7568	832.7568	832.9504	836.2332	841.3129

As detailed in Table 2, the proposed NTPEGD model demonstrated superior performance for Dataset 1, achieving the lowest values across all goodness-of-fit criteria (-2log-likelihood = 826.736, AIC = 830.736, AICC = 830.832, HQIC = 833.053, BIC = 836.440) when compared to five competing models (GLD1-GLD5). It outperformed all other two-parameter models and, crucially, also outperformed the three-parameter GLD5 model, whose additional complexity ($\beta = 668.8$) failed to provide a better fit and was heavily penalized by the information criteria. This consistent dominance confirms that the NTPEGD provides the most parsimonious and optimal fit for the data.

Table 3: Parameter estimates and goodness-of-fit statistics for the proposed and competing models applied to Dataset 2

Models	Parameter estimates	-2logL	AIC	AICC	HQIC	BIC
NTPEGD	$\hat{\alpha} = 0.7455$ $\hat{\theta} = 0.1985$	334.5374	338.5374	338.7088	340.3629	343.1183
GLD1	$\boldsymbol{\hat{\alpha} = 0.7463}$ $\boldsymbol{\hat{\theta} = 0.2087}$	**335.3726**	**339.3726**	**339.544**	**341.1982**	**343.9535**
GLD2	$\hat{\alpha} = 1.3788$ $\hat{\theta} = 0.3049$	358.8346	362.8346	363.0059	364.6601	367.4154
GLD3	$\hat{\alpha} = 6808.97$ $\hat{\theta} = 0.2663$	339.1788	343.1788	343.3503	345.0045	347.7598
GLD4	$\hat{\alpha} = 0.5088$ $\hat{\theta} = 0.2877$	356.7712	360.7712	360.9426	362.5967	365.3521
GLD5	$\hat{\alpha} = 0.7457$ $\hat{\theta} = 0.1987$ $\hat{\beta} = 0.0013$	334.5412	340.5412	340.8791	343.2696	347.4026

Table 3 represents the goodness-of-fit test of the NTPEGD and other competing models for the second dataset. The proposed NTPEGD model provides the optimal fit for the acute bone cancer survival data (Dataset 2), achieving the lowest values across all information criteria (AIC = 338.54, BIC = 343.12). It demonstrates a superior balance of goodness-of-fit and parsimony, outperforming both simpler models like GLD1 (ΔAIC = 0.84) and more complex alternatives like the three-parameter GLD5 (ΔAIC = 2.00), whose additional parameter failed to yield a meaningfully better likelihood. This strong performance confirms the NTPEGD's effectiveness in modelling this survival time distribution.

VI. CONCLUSION

This study proposed the novel NTPEG distribution, developed through a kth-component additive mixture of Exponential and Gamma components to enhance flexibility in modeling lifetime data. The statistical properties were systematically derived, and parameter estimation via maximum likelihood was shown, through simulation, to yield consistent and efficient estimators. Empirical applications to non-communicable disease data demonstrated that the NTPEG model provides superior fit compared to several existing Lindley-type and related distributions, achieving an optimal balance between flexibility and parsimony.

Looking ahead, the NTPEG distribution offers a promising foundation for further methodological and applied research. Potential directions include exploring Bayesian estimation procedures, developing regression extensions for covariate-dependent survival modeling, and applying the model in broader contexts such as reliability engineering, epidemiology, and financial risk assessment. Such extensions would further validate and expand the usefulness of the NTPEG framework across diverse real-world applications.

REFERENCES

[1]Klakattawi HS (2022) Survival analysis of cancer patients using a new extended Weibull distribution. PLoS ONE 17(2): e0264229. https:// doi.org/10.1371/journal.pone.0264229

[2]Warahena-Liyanage, G.; Oluyede, B.; Moakofi, T.; Sengweni, W. (2023). The New Exponentiated Half Logistic-Harris-G Family of Distributions with Actuarial Measures and Applications. Stats 2023, 6, 773–801. https://doi.org/ 10.3390/stats6030050

[3]Lindley, D.V., (1958), Fiducial Distributions and Bayes' Theorem, Journal of the Royal Statistical Society Series B (Methodological), 102-107.

[4]Ekhosuehi. N., Opone, F. and Odobaire, F. (2018). A New Generalized Two Parameter Lindley Distribution. *Journal of Data Science*. 16(3); 549-566 , DOI: 10.6339/JDS.201807_16(3).0006

[5]Abouammoh, A.M., Arwa M., Alshangiti and Ragab I. E., (2015): A new generalized Lindley distribution, Journal of Statistical Computation and Simulation, DOI: 10.1080/00949655.2014.995101

[6]Aderoju, S.A. (2020). A New Generalized Poisson Mixed Distribution and its application. *Applied Mathematical Sciences*, Vol. 14, 2020, no. 5, 229 – 234; https://doi.org/10.12988/ams.2020.914185

[7]C. Chesneau, L. Tomy and J. Gillariose, (2020). The inverted modified Lindley distribution, Journal of Statistical Theory and Practice, 14, 1-17.

[8]Aderoju, S.A. and Babaniyi, O. (2023): Power Samade distribution: its properties and application to real lifetime data. *Nigerian Journal of Science and Environment (NJSE)*. 21 (1): 237–248.

[9]Aderoju, S.A. and Jolayemi, E.T. (2022): Power Hamza distribution and its applications to model survival time. *Journal of the Nigerian Statistical Association*, 34 (1): 1-10.

[10]Aleshinloye, N. I., Aderoju, S.A., Abiodun, A. A. and Taiwo, B. L. (2023) "A New Generalized Gamma-Weibull Distribution and its Applications," *Al-Bahir Journal for Engineering and Pure Sciences*. 2 (2): 92-100.

[11]Aderoju, S. A., Adeniyi, I., Olaifa, J.B., Olaosebikan, A. (2023): On Poisson-Samade Distribution: Its Applications in Modelling Count Data. *Earthline Journal of Mathematical Sciences.* 12 (2): 255-270. Published by *Earthline Publishers, India.* Available online at

[12]Aderoju, S.A., Dauda, K.A. and Olaifa, J.B. **(**2025). Development of New Flexible Lifetime Model: its Associated Inferences and Applications to Cancer and Covid-19 Data. *Journal of Science and Technology*. 30 (8). DOI:https://doi.org/10.20428/jst.v30i8.3024.

[13]Ghitany, M. E., Atieh, B. and Nadarajah, S., (2008), Lindley Distribution and its Applications, Mathematics and Computers in Simulation, 78(4), 493-506.

[14]Mazucheli, J. and Achcar, J. A. (2011). The Lindley distribution applied to competing risks lifetime data. *Computer methods and programs in biomedicine, 104(2), 188-192*.

[15]Deniz, E. G. and Ojeda, E. C. (2011). The discrete Lindley distribution-properties and applications. *Journal of Statistical Computation and Simulation, 81(11): 1405–1416*.

[16]Bakouch, H. S., Al-Zahrani, B. M., Al-Shomrani, A. A., Marchi, V. A., and Louzada, F. (2012). An extended Lindley distribution. *Journal of the Korean Statistical Society, 41, 75-85.*

[17]Ghitany, M. E., Alqallaf, F., Al-Mutairi, D. K., and Husain, H. A. (2011). A two-parameter weighted Lindley distribution and its applications to survival data. *Mathematics and Computers in simulation, 81(6), 1190-1201*

[18]Zakerzadeh, H. and Dolati, A., (2009), Generalized Lindley Distribution, Journal of Mathematical Extension 3, 13-25.

[19]Aderoju, S.A. (2021). Samade Probability Distribution: Its Properties and Application to Real Lifetime Data. *Asian Journal of Probability and Statistics*. 14 (1): 1-11.

[20]Aijaz, A., Jallal, M., Qurat, S., Ain, U.I. and Tripath, R. (2020). The Hamza distribution with statistical properties and applications, Asian Journal of Probability and Statistics,8(1),28 – 42.DOI: 10.9734/AJPAS/2020/v8i130196

[21]Aleshinloye, N. I., Aderoju, S.A., Abiodun, A. A. and Taiwo, B. L. (2023) "A New Generalized Gamma-Weibull Distribution and its Applications," *Al-Bahir Journal for Engineering and Pure Sciences*. 2 (2): 92-100.

[22]Aderoju, S.A. and Adeniyi, I. (2022a). On Power Generalized Akash Distribution with Properties and Applications. *Journal of Statistical Modelling &Amp; Analytics (JOSMA)*. 4 (1): 1-13.

[23]Aderoju, S. A**.,** and Adeniyi, I. (2022b). On zero inflated generalized Poisson-Sujatha distribution and its applications. *Statistica Applicata - Italian Journal of Applied Statistics*, *34*(1), 7–22. https://doi.org/10.26398/IJAS.0034-001

[24]Everitt, B.S., Hand, D.J. (1981). Mixtures of normal distributions. In: *Finite Mixture Distributions*. Monographs on Applied Probability and Statistics. Springer, Dordrecht. pp. 25–57. https://doi.org/10.1007/978-94-009-5897-5_2

[25]Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1992). *A First Course in Order Statistics*. Wiley.

[26]Balakrishnan and Cohen (1991): Balakrishnan, N., & Cohen, A. C. (1991). *Order Statistics & Inference: Estimation Methods*. Academic Press.

[27]Rényi, A, (1961). On measure of entropy and information. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability 1, University of California Press, Berkeley. 547-561.

[28]R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

[29]Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

[30]Hurvich, C. M., and Tsai, C.L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2), 297–307. https://doi.org/10.1093/biomet/76.2.297

[31]Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2), 190–195. https://www.jstor.org/stable/2985032

[32]Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. https://doi.org/10.1214/aos/1176344136

[33]Shanker, R., & Mishra, A. (2013). A two-parameter Lindley distribution. *Statistics in Transition New Series*, 14(1), 45–56.

[34]Nadarajah, S., Bakouch, H. S., & Tahmasbi, R. (2011). A generalized Lindley distribution. *Sankhya B*, 73(2), 331–359.

[35]Zakerzadeh, H., & Dolati, A. (2009). Generalized Lindley distribution. *Journal of Mathematical Extension*, 3(2), 13–25.

[36]Lee, E. T., & Wang, J. (2003). *Statistical Methods for Survival Data Analysis* (3rd ed.). Wiley-Interscience.

[37]Mansour, M. A., Bakouch, H. S., & Al-Babtain, A. (2020). A new two-parameter Burr XII distribution: Properties, copula, different estimation methods, and modelling acute bone cancer data. *Journal of King Saud University – Science*, 32(1), 390–398.