THE EG-LINDLEY DISTRIBUTION: A FLEXIBLE MODEL FOR NON-COMMUNICABLE AND CHRONIC DISEASE SURVIVAL ANALYSIS

Samuel Adewale Aderoju[1,2*], Rasheed Kehinde Lamidi[1], Bello Ishola Sanni[1], Abdulazeez Kayode Jimoh[1], Adediran Dauda Adeshola[1], Uchechukwu Kalu[1] and Ebenezer Juwon Adeniyi[1]

Received: June 01, 2025; Revised: August 02, 2025; Accepted: August 04, 2025

Abstract

This study introduces and evaluates the EG-Lindley (EG-L) distribution as a flexible model for analyzing survival data. The EG-L distribution is shown to capture a wide range of hazard rate behaviors - including increasing, decreasing, and bathtub-shaped patterns, making it highly adaptable to biomedical and reliability data analysis. A comprehensive analysis of the statistical properties of the EG-L distribution is presented, including the probability density function, cumulative distribution function, survival function, and hazard function, supported by detailed graphical interpretations. To assess the empirical performance of the EG-L distribution, three real-world datasets were analyzed: (1) bladder cancer patient survival times, (2) acute bone cancer survival data, and (3) head and neck cancer cases treated with radiotherapy and chemotherapy. The proposed model was compared with several existing generalizations of the Lindley distribution - namely, the New Generalized Two-Parameter Lindley Distribution, and the four other Generalized Lindley Distributions. Model performance was evaluated using goodness-of-fit criteria, including the -2log-likelihood, Akaike Information Criterion, Bayesian Information Criterion, Hannan–Quinn Information Criterion, and the Corrected Akaike Information Criterion. Across all datasets, the EG-L distribution consistently achieved the best or most competitive fit, as evidenced by the lowest information criteria values. These results underscore the robustness and adaptability of the EG-L distribution in capturing complex survival patterns and support its application as a valuable tool in the analysis of time-to-event data in clinical and epidemiological research.

Keywords: **EG-Lindley Distribution; Survival Analysis; Hazard Function; Goodness-of-Fit; Biomedical Data Modelling; Maximum Likelihood Method**

[1] *Department of Mathematics and Statistics, Kwara State University, Malete, P.M.B. 1530, Ilorin, Kwara State, Nigeria.*
[2] *Statistics Programme, Bowen University, Iwo, Nigeria*
[*] *Corresponding author*

DOI: https://doi.org/

Suranaree J. Sci. Technol. 33(1):030366(1-11)

Introduction

The analysis of real lifetime data plays a critical role across numerous disciplines such as medicine, genetics, engineering, computer science, biology, statistics, and economics. Among the statistical models employed in this context, the Lindley distribution (Lindley, 1958) stands out as an important continuous probability model. Originally conceived as a Bayesian mixture of exponential and gamma distributions, it has proven to be a valuable alternative to the exponential distribution and serves as a notable counterexample in fiducial statistics.

Over time, the Lindley distribution has undergone extensive study and has been adapted into numerous flexible forms to enhance its capacity for modelling varied data patterns. These modifications aim to capture a wider range of data skewness and distributional shapes. Prominent contributions to this field include those by Zakerzadeh & Dolati (2009), Ghitany *et al*. (2013), Nadarajah *et al*. (2015), Pararai *et al*. (2015), Aderoju (2021), Aderoju and Babalola (2023), Aleshinloye *et al*. (2023), and Aderoju *et al*. (2023). While the inclusion of additional parameters increases a model's adaptability, it also brings complexity to the parameter estimation process, particularly in models comprising three or more parameters, as highlighted by Wolstenholme (2018) and Miller (2012).

The existing Lindley models still exhibit limitations in modelling various types of real-life data. To address this gap, the Extended Generalized Lindley Distributions have been proposed. These new approaches build upon Lindley's foundational models and integrate modern enhancements as well as generalizations from recent literature, such as those by Ahsanullah *et al*. (2017), Irshad (2017), and Hamed & Alzaghal (2021). The Extended Generalized Lindley Distributions are designed to offer both flexibility and robustness, aiming to accommodate a broad spectrum of data behaviors.
In their work, Hamed and Alzaghal (2021) presented the T-Lindley{Y} class of distributions, developed through transformations of quantile functions from various distributions, including the exponential, Weibull, logistic, log-logistic, and Cauchy distributions. A specific member of this class demonstrated outstanding performance in modelling both unimodal and bimodal lifetime data.

Similarly, Gemeay *et al*. (2023) introduced a new two-parameter probability model, formulated as a mixture of exponential and gamma distributions. Their comparative analysis, particularly using COVID-19 datasets, revealed its superiority over several classical models.

Aderoju and Adeniyi (2022) extended the generalized Akash distribution by proposing the Power Generalized Akash Distribution (PGAD). Their empirical findings showed PGAD to be highly adaptable and effective in modelling real lifetime data, as evidenced by favorable BIC, AIC, and 2Log-likelihood criteria.

In another notable contribution, Edith *et al*. (2019) proposed the Two-Parameter Rama Distribution (TPRD), derived as a mixture similar to Lindley's approach. The TPRD exhibited improved goodness-of-fit in numerous applications compared to conventional models.

Shanker (2020) proposed the Two-Parameter New Quasi Sujatha Distribution, formulated through a convex combination of exponential and gamma distributions with varying parameters. This innovative distribution has opened up new perspectives and practical applications in statistical modelling.

Further extending this line of research, Ganaie *et al*. (2020) introduced the Weighted New Quasi Lindley Distribution, a generalized form achieved via a weighted approach. Their empirical analysis, based on real-world datasets, confirmed the model's suitability for lifetime data modelling.

This study aims to introduce a novel, flexible lifetime model known as the EG-Lindley (EG-L) distribution, grounded in a mixture of exponential and gamma distributions. The objective is to enhance modelling flexibility for survival and lifetime data. The structure of the paper is organized as follows: Section 2 outlines the materials and methods, focusing on the derivation and formulation of the EG-L distribution. Section 3 presents the mathematical properties of the distribution. Section 4 addresses parameter estimation using the maximum likelihood approach and simulation studies. Section 5 showcases applications of the proposed model to real datasets. The paper concludes with final remarks in Section 6.

Materials and Methods

The probability density function (pdf) of a random variable X in a k^{th} component additive mixture distribution is defined as follows (Everitt and Hand, 1981):

$$f(x,\Omega) = \sum_{j=1}^{k} f_j(x,\Omega_j)\omega_j, \quad (1)$$

where $\sum_{j=1}^{k} \omega_j = 1$

In this context, Ω_j refers to the vector containing parameters (α,θ) for the mixture model, while ω_j represents the corresponding mixture proportion.

Theorem 1: Given a random variable X that follows a New Two parameter Exponential-Gamma distribution, which shall henceforth be called EG-Lindley distribution, denoted as $X|\alpha,\theta \sim EGL\ (\alpha,\theta)$, the pdf of X is specified by the following expression

$$f(x;\alpha,\theta) = \frac{\theta^4}{(\theta^3+6)}\left[1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right]e^{-\theta x};\ \ \theta, x > 0\ and\ \alpha > 1\,, \quad (2)$$

where α is a shape parameter and θ is a scale parameter, also $\Gamma(.)$ is the gamma function, for all positive integers, hence,

$$\Gamma(\alpha-1) = \int_0^\infty t^{\alpha-2}\, e^{-t} dt$$

Proof 1: Recall the pdf of a k^{th} component additive mixture distribution (1), therefore, equation (2) represents a two-component mixture comprising exponential (θ) and gamma $(\alpha-1,\ \theta)$ distributions. The mixing proportion, denoted as $P = \frac{\theta^3}{(\theta^3+6)}$, and can be written as:

$$f(x;\alpha,\theta) = Pg_1(x;\theta) + (1-P)g_2(x;\alpha,\theta),$$

where

$g_1(x;\theta) = \theta e^{-\theta x}$ and
$g_2(x;\alpha,\theta) = \frac{\theta^{\alpha-1}x^{\alpha-2}}{\Gamma(\alpha-1)}e^{-x\theta}$.

This gives;

$$f(x;\alpha,\theta) = \left[\frac{\theta^3}{(\theta^3+6)}.\theta e^{-\theta x}\right] + \left[\frac{6}{(\theta^3+6)}.\frac{\theta^{\alpha-1}x^{\alpha-2}}{\Gamma(\alpha-1)}e^{-x\theta}\right]$$

$$f(x;\alpha,\theta) = e^{-\theta x}\left[\left(\frac{\theta^4}{(\theta^3+6)}\right) + \left(\frac{6}{(\theta^3+6)}\frac{\theta^{\alpha-1}}{\Gamma(\alpha-1)}(x^{\alpha-2})\right)\right]$$

$$f(x;\alpha,\theta) = e^{-\theta x}\left[\frac{\theta^4}{(\theta^3+6)} + \frac{6\theta^4\theta^{\alpha-5}x^{\alpha-2}}{(\theta^3+6)\Gamma(\alpha-1)}\right]$$

$$f(x;\alpha,\theta) = \frac{\theta^4}{(\theta^3+6)}\left[1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right]e^{-\theta x}$$

Hence, the probability density function provided in (2) represents the distribution of the EG-L distribution.

Validity check of the New EG-Lindley distribution

Theorem 2: If a random variable X follows a New EG-Lindley distribution denoted as $X|\alpha,\theta \sim EGL(\alpha,\theta)$, then its pdf is valid if it fulfills the following condition:

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Proof 2: Since $x \in R$; and $f_1(x\,;\alpha,\theta)$ to be $\in (0,\infty)$, then

$$\int_0^\infty f(x;\alpha,\theta) = \int_0^\infty \frac{\theta^4}{(\theta^3+6)}\left(1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)e^{-x\theta}dx$$

$$= \int_0^\infty \frac{1}{(6+\theta^3)}\left(\theta^4 + \frac{6\theta(\theta x)^{\alpha-2}}{\Gamma(\alpha-1)}\right)e^{-x\theta}dx$$

$$= \frac{1}{(6+\theta^3)}\left[\int_0^\infty \theta^4 e^{-x\theta}dx + \frac{6\theta}{\Gamma(\alpha-1)}\int_0^\infty (\theta x)^{\alpha-2}\ e^{-x\theta}dx\right]$$

Let $t = \theta x,\ dt = \theta dx,\ dt = \theta dx,\ dx = \frac{dt}{\theta}$

$$= \frac{1}{(6+\theta^3)}\left[\theta^3\int_0^\infty \theta e^{-x\theta}dx + \frac{6}{\Gamma(\alpha-1)}\int_0^\infty t^{\alpha-2}e^{-t}dt\right],$$

where $\int_0^\infty t^{\alpha-2}e^{-t}dt = \Gamma(\alpha-1)$

$$\int_0^\infty f(x;\alpha,\theta) = \frac{1}{(6+\theta^3)}\left[\theta^3(1) + \frac{6\Gamma(\alpha-1)}{\Gamma(\alpha-1)}\right]$$

$$= \frac{1}{(6+\theta^3)}\left[\frac{(\theta^3+6)}{1}\right] = 1$$

Therefore, the probability density function of the random variable X is valid. The corresponding cumulative distribution function (cdf) for the EG-L distribution is given as follows:

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

$$F(x;\alpha,\theta) = \int_{t=0}^{x} \frac{\theta^4}{(\theta^3+6)}\left(1+\frac{6\theta^{\alpha-5}t^{\alpha-2}}{\Gamma(\alpha-1)}\right)e^{-\theta t}\,dt$$

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

Figure 1 represents the set of plots (labeled a-d) that shows the probability density functions for selected values of the shape parameter $\hat{\alpha}$ and scale parameter $\hat{\hat{\theta}}$. The graphs show that:

Increasing the scale parameter $\hat{\hat{\theta}}$ shifts the distribution rightward and increases its spread.

The shape parameter $\hat{\alpha}$ significantly influences the modality and skewness of the distribution.

Moreover, in panel (a) at $\hat{\alpha} = 1.25$: The distribution is heavily right-skewed with peaks near zero. Increasing $\hat{\hat{\theta}}$ flattens the curve and shifts it rightward. In panel (b), at $\hat{\alpha} = 1.5$: The same right-skewed pattern is observed, with a longer tail as $\hat{\hat{\theta}}$ increases. In panel (c), at $\hat{\alpha} = 15$: The distributions begin to show multi-modal characteristics. The mode shifts right as $\hat{\theta}$ increases. In panel (d) at $\hat{\alpha} = 25$: The behavior is similar to (c), with more pronounced shifts and multiple modes as $\hat{\hat{\theta}}$ increases.

Similarly, Figure 2 displays the cumulative distribution functions for varying parameters. These illustrate how the distribution accumulates probability mass over the domain. Generally, a smaller-scale parameter $\hat{\hat{\theta}}$ leads to a faster accumulation of probability mass. Larger $\hat{\hat{\theta}}$ values result in heavier tails and slower increases in the cdf. In the left plot ($\hat{\alpha} = 1.5$): The cdf increases rapidly for smaller $\hat{\hat{\theta}}$ values, indicating the concentration of probability near the origin. Larger $\hat{\hat{\theta}}$ values result in gradual accumulation. In the right plot ($\hat{\alpha} = 10$): Similar behavior is observed. As $\hat{\hat{\theta}}$ increases, the distribution spreads wider, and the cdf becomes less steep.

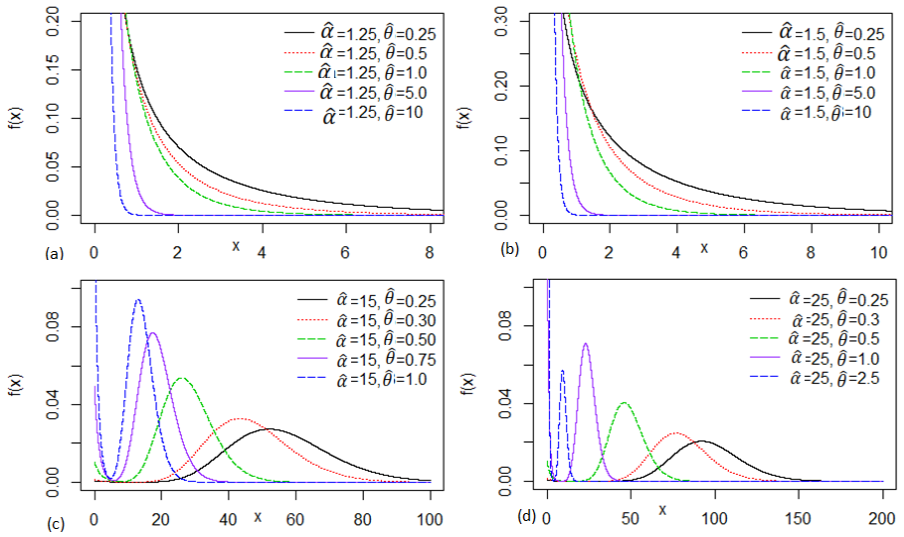


Figure 1. The probability density curves for various selected values of the shape and scale parameters

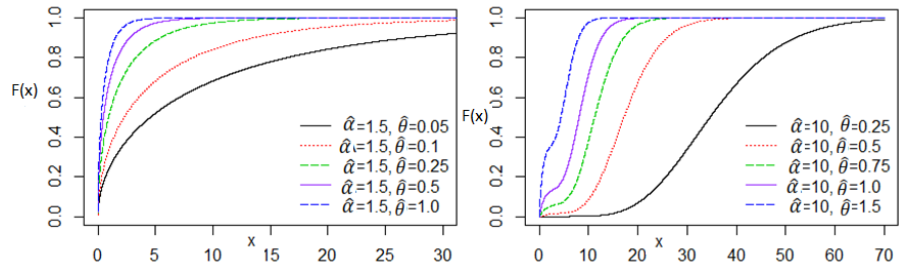


Figure 2. CDF curves of the EG-L distribution for some selected values of parameters

Mathematical properties of the EG-L distribution

This section presents derived mathematical properties of the EG-L distribution, encompassing moments, variance, coefficient of variation, skewness, kurtosis, reliability analysis, entropy, and order statistics.

Moments

$$E(x^r) = \mu_r' = \int_0^\infty x^r f(x)dx \tag{4}$$

$$= \int_{x=0}^{\infty} x^r \frac{\theta^4}{(\theta^3+6)}\left(1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)e^{-\theta x}\,dx$$

$$= \frac{\theta^4}{(\theta^3+6)}\int_{x=0}^{\infty} x^r\left(1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)e^{-\theta x}\,dx$$

$$\mu_r' = \frac{\theta^r\left(\theta^3\Gamma(1+r)+\Gamma(\alpha-1)+6\Gamma(r+\alpha-1)\right)}{(6+\theta^3)\Gamma(\alpha-1)} \tag{5}$$

$$\sigma^2 = \frac{2(-3\alpha+3\alpha^2+\theta^3)}{\theta^2(6+\theta^3)} - \left(\frac{-6+6\alpha+\theta^3}{\theta(6+\theta^3)}\right)^2$$

$$\sigma^2 = \frac{-36+24\theta^3+6\alpha^2\theta^3+\theta^6-18\alpha(-2+\theta^3)}{\theta^2(6+\theta^3)^2} \tag{6}$$

The coefficient of variance is determined by

$$C.V = \frac{\sigma}{\mu_1}$$

$$CV = \frac{\sqrt{(-36+36\alpha+24\theta^3-18\alpha\theta^3+6\alpha^2\theta^3+\theta^6)}}{(-6+6\alpha+\theta^3)(6\theta+\theta^4)} \tag{7}$$

The skewness (S_K) of NGL distribution is

$$S_K = \frac{\mu_3}{\sigma^3}$$

$$S_K = \frac{6\left(-\alpha+\alpha^3+\theta^3\right)}{\theta^3(6+\theta^3)\left(\frac{-36+24\theta^3+6\alpha^2\theta^3+\theta^6-18\alpha(-2+\theta^3)}{\theta^2(6+\theta^3)^2}\right)^{3/2}} \tag{8}$$

The kurtosis (K_s) of EG-L distribution is

$$K_s = \frac{\mu_4}{\sigma^4}$$

where r is a positive integer. The first four moments as $(r = 1; 2; 3; 4)$ may be expressed as follows: from equation (5):

$$M_1 = \frac{6\alpha-6+\theta^3}{\theta(6+\theta^3)}$$

$$M_2 = \frac{2(\theta^3+3\alpha^2-3\alpha)}{\theta^2(6+\theta^3)}$$

$$M_3 = \frac{6(\theta^3+\alpha^3-\alpha)}{\theta^3(6+\theta^3)}$$

$$M_4 = \frac{6\left((\alpha-1)\alpha(\alpha+1)(\alpha+2)+4\theta^3\right)}{\theta^4(6+\theta^3)}$$

Variance and Coefficient of Variance

For a random variable X following the EG-L distribution with parameters α and θ, the variance (Var) and coefficient of variation (CV) are given by:

$$\sigma^2 = \mu_2^l - (\mu_1^l)^2$$

$$K_s = \frac{6(6+\theta^3)^3\left((-1+\alpha)\alpha(1+\alpha)(2+\alpha)+4\theta^3\right)}{\left(-36+24\theta^3+6\alpha^2\theta^3+\theta^6-18\alpha(-2+\theta^3)\right)^2} \tag{9}$$

Rényi's entropy

Rényi's Entropy for the EG-L distribution is expressed as

$$R_H(x;\,\alpha,\theta) = \frac{1}{1-w}\log\int_0^\infty [f(x;\,\alpha,\theta)]^w\,dx, w>1$$

$$R_H(x;\,\alpha,\theta) = \frac{1}{1-w}\log\int_0^\infty \left[\frac{\theta^4}{(6+\,\theta^3)}\left(1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)e^{-\theta x}\right]^w dx$$

$$R_H(x;\,\alpha,\theta) = \frac{1}{1-w}\log\int_0^\infty \left[\frac{\theta^{4w}}{(6+\theta^3)^w}\left(1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)^w e^{-\theta xw}\right]^w dx$$

$$R_H(x;\,\alpha,\theta) = \frac{1}{1-w}\log\left[\frac{\theta^{4w}}{\exists(6+\theta^3)^w}\int_0^\infty\left(1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)^w e^{-\theta xw}\right]dx$$

Using the Taylor Series expansion of $(1+x)^w = \sum_{u=0}^{w}\binom{w}{u}x^u$

$$R_H(x;\,\alpha,\theta) = \frac{1}{1-w}\log\left[\frac{\theta^{4w}}{(6+\theta^3)^w}\int_0^\infty\sum_{u=0}^{w}\binom{w}{u}\left(\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)^u e^{-\theta xw}\right]dx$$

$$R_H(x;\,\alpha,\theta) = \frac{1}{1-w}\log\left[\frac{\theta^{4w}}{(6+\theta^3)^w}\sum_{u=0}^{w}\binom{w}{u}\left(\frac{6\theta^{\alpha-5}}{\Gamma(\alpha-1)}\right)^u\int_0^\infty x^{(\alpha-5)u}e^{-\theta xw}\,dx\right]$$

$$Note\ that, \int_0^\infty x^{(\alpha-2)u}e^{-\theta xw}\,dx = \frac{\Gamma(u\alpha-u+2)}{(w\theta)^{u\alpha-u+2}}$$

$$R_H(x;\,\alpha,\theta) = \frac{1}{1-w}\log\left[\frac{\theta^{4w}}{(6+\theta^3)^w}\sum_{u=0}^{w}\binom{w}{u}\left(\frac{6\theta^{\alpha-5}}{\Gamma(\alpha-1)}\right)^u\frac{\Gamma(u\alpha-u+2)}{(w\theta)^{u\alpha-u+2}}\right] \tag{10}$$

Order statistics

Let $X_1, X_2, \ldots X_n$ represent a random sample from the EG-L distribution. The density function of

the k^{th} order statistic, denoted as $F_{x(k)}(x;\ \theta, \alpha)$, can be expressed as follows.

$$F_{x(k)}(x;\alpha,\theta) = \frac{n!}{(k-1)!(n-k)!} f_1(x)[F_1(x)]^{k-1}[1-F_1(x)]^{n-k} \tag{11}$$

where $f_1(x)$ represent the pdf and $F_1(x)$ represent the cdf as defined in equations (2) and (4), respectively. Therefore, the density function of the k^{th} order statistics of the EG-L distribution can be derived as:

$$F_{x(k)}(x;\alpha,\theta) = \frac{n!\ \theta^4}{(k-1)!\,(n-k)!\,(6+\theta^3)} \left(\frac{1+6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right) e^{-\theta x}\left[\frac{\left(1-e^{-x\theta}\right)\theta^3 + 6x^{\alpha}\theta^{\alpha}(x\theta)^{-\alpha}\left(\Gamma(\alpha-1)-\Gamma(\alpha-1,x\theta)\right)}{\Gamma(\alpha-1)} \over 6+\theta^3\right]^{k-1} \left[1-\frac{\frac{\left(1-e^{-x\theta}\right)\theta^3+6x^{\alpha}\theta^{\alpha}(x\theta)^{-\alpha}\left(\Gamma(\alpha-1)-\Gamma(\alpha-1,x\theta)\right)}{\Gamma(\alpha-1)}}{6+\theta^3}\right]^{n-k} \tag{12}$$

where n≠0.

The density functions for the first and n[th]-order statistics can be readily obtained from (12) to respectively yield,

$$F_{x(1)}(x;\alpha,\theta) = \frac{n!\ \theta^4}{(k-1)!\,(n-k)!\,(6+\theta^3)} \left(\frac{1+6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right) e^{-\theta x}\left[1-\frac{\left(1-e^{-x\theta}\right)\theta^3 + 6x^{\alpha}\theta^{\alpha}(x\theta)^{-\alpha}\left(\Gamma(\alpha-1)-\Gamma(\alpha-1,x\theta)\right)}{\Gamma(\alpha-1)} \over 6+\theta^3\right]^{n-1}$$

and

$$F_{x(n)}(x;\alpha,\theta) = \frac{n!\ \theta^4}{(k-1)!\,(n-k)!\,(6+\theta^3)} \left(\frac{1+6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right) e^{-\theta x}\left[\frac{\left(1-e^{-x\theta}\right)\theta^3 + 6x^{\alpha}\theta^{\alpha}(x\theta)^{-\alpha}\left(\Gamma(\alpha-1)-\Gamma(\alpha-1,x\theta)\right)}{\Gamma(\alpha-1)} \over 6+\theta^3\right]^{n-1}$$

Reliability analysis

Reliability analysis of a distribution relies on the survival and hazard rate functions, which are derived through the following steps. In the context of the EG-L distribution, the Survival function S(x) is given by

$$S(x;\alpha,\theta) = 1 - F_1(x;\alpha,\theta)$$

$$S(x;\alpha,\theta) = 1 - \frac{\left(1-e^{-x\theta}\right)\theta^3 + \frac{6x^{\alpha}\theta^{\alpha}(x\theta)^{-\alpha}\left(\Gamma(-1+\alpha)-\Gamma(-1+\alpha,x\theta)\right)}{\Gamma(-1+\alpha)}}{6+\theta^3} \tag{13}$$

The hazard rate function, represented by h(x), is similarly derived for the EG-L distribution as

$$h(x;\alpha,\theta) = \frac{f_1(x;\alpha,\theta)}{1-F_1} = \frac{f_1(x;\alpha,\theta)}{S(x;\alpha,\theta)}$$

Using (2) and (3), we have

$$h(x;\alpha,\theta) = \frac{\theta(x\theta)^{-2+\alpha}\left(6x^{\alpha}\theta^{\alpha}+x^2\theta^5\Gamma(-1+\alpha)\right)}{(\theta^3(x\theta)^{\alpha}+e^{x\theta}(-6x^{\alpha}\theta^{\alpha}+6(x\theta)^{\alpha}))\Gamma(-1+\alpha)+6e^{x\theta}x^{\alpha}\theta^{\alpha}\Gamma(-1+\alpha,x\theta)} \tag{14}$$

In Figure 3 (Left Plot), the survival probability S(x) decreases with increasing *x*, as expected. Smaller values of $\hat{\theta}$ (e.g., 0.05, 0.10) produce slower decays, indicating longer survival times. Larger $\hat{\theta}$ values (e.g., 0.75, 1.0) result in steeper declines, meaning shorter survival times. This shows that the scale parameter ($\hat{\theta}$) inversely affects the tail of the survival function. At a lower $\hat{\alpha}$ (like 1.5), the scale parameter plays a dominant role in controlling how quickly survival diminishes. As $\hat{\theta}$ increases, the probability of survival beyond a given time drops significantly.

In Figure 3 (Right Plot), for the higher shape parameter ($\hat{\alpha} = 15$), the survival curves become more compressed. Survival drops off very rapidly for large $\hat{\theta}$ values (e.g., 2.0), indicating extremely short lifetimes. At smaller $\hat{\theta}$ (e.g., 0.10), the survival curve stretches out, showing much longer lifetimes. Compared to the left plot, increasing α leads to sharper transitions in survival, especially when $\hat{\theta}$ is large. At high values of $\hat{\alpha}$, even modest changes in $\hat{\theta}$ produce significant shifts in survival probabilities. This suggests strong sensitivity of survival time to parameter changes in this region, indicating a more peaked distribution with a narrower spread. Generally, both plots confirm the flexibility of the distribution in modelling survival data across short, medium, and long lifetimes. The survival function becomes steeper and shorter-lived as $\hat{\theta}$ increases, and becomes more stretched out as $\hat{\theta}$ decreases. The effect is amplified for higher values of $\hat{\alpha}$, making the distribution suitable for modelling both acute and chronic lifetime data in biosciences.

Figure 4 represents the hazard function *h(x)* of a statistical distribution (likely a generalized gamma distribution or a similar flexible parametric model), under varying values of the scale parameter ($\hat{\theta}$), while keeping the shape parameter ($\hat{\alpha}$) fixed. In Figure 4 (Left Plot: $\hat{\alpha} = 2.5$, varying $\hat{\theta}$), all hazard functions increase monotonically and appear to plateau at different levels. This is indicative of an increasing hazard rate, characteristic of aging or wear-out processes**.** Smaller $\hat{\theta}$ (e.g., 0.25) leads to a slower increase in hazard and a lower asymptotic hazard. Larger $\hat{\theta}$ (e.g., 2.5) produces a steeper increase and a higher hazard plateau. So, as $\hat{\theta}$ increases, the scale of the x-axis stretches, and the hazard function rises more sharply and levels higher. In Figure 4 (Right Plot: $\hat{\alpha} = 5$, varying $\hat{\theta}$), more complex shapes, with some

hazard functions initially decreasing and then increasing (i.e., bathtub-shaped hazard). This suggests that the distribution allows for non-monotonic hazard behavior, useful in reliability analysis (e.g., infant mortality → useful life → wear-out). Small $\hat{\theta}$ (e.g., 0.25) leads to flat or concave hazard functions (slow growth). Intermediate values ($e.g.,\ \hat{\theta} = 1.5$) show a minimum in hazard, then an upward trend. Larger $\hat{\theta}$ (e.g., 5) leads to a slower increasing hazard, with a lower peak compared to the left panel.

Maximum Likelihood Estimation and goodness-of-fit test

The EG-L distribution's parameters are determined through the maximum likelihood estimation technique. The likelihood function for a random sample $X_1, X_2, \ldots X_n$ from the EG-L distribution with parameters α and θ can be expressed as follows.

$$L(\alpha,\theta) = \prod_{i=1}^{n} f_i\,(x_i;\alpha,\theta)$$

This gives

$$L(\alpha,\theta) = \prod_{i=1}^{n} \frac{\theta^4}{(\theta^3+6)}\left[1+\frac{6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right]e^{-\theta x}$$

The log-likelihood function is expressed as

$$F_{x(k)}(x;\alpha,\theta) = \frac{n!\ \theta^4}{(k-1)!\,(n-k)!\,(6+\theta^3)}\left(\frac{1+6\theta^{\alpha-5}x^{\alpha-2}}{\Gamma(\alpha-1)}\right)$$
$$e^{-\theta x}\left[\frac{\left(1-e^{-x\theta}\right)\theta^3+6x^{\alpha}\theta^{\alpha}(x\theta)^{-\alpha}\left(\Gamma(\alpha-1)-\Gamma(\alpha-1,x\theta)\right)}{\Gamma(\alpha-1)}\right]^{k-1}$$

$$\ell(\alpha,\theta) = 4n\log\,\theta - n\log(6+\theta^3) + \sum_{i=1}^{n}\log(\Gamma(\alpha-1)+6\theta^{\alpha-5}x^{\alpha-2}) - nlog\Gamma(\alpha-1) - \theta\sum_{i=1}^{n}x_i \quad (15)$$

By taking partial derivatives of the log-likelihood function (15) with respect to its corresponding parameters, we obtain

$$\frac{\partial\ell(\alpha,\theta)}{\partial\theta} = \frac{4n}{\theta} - \frac{3n\theta^2}{6+\theta^3} - \sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\frac{6(\alpha-5)\theta^{\alpha-6}x_i^{\alpha-2}}{6\theta^{\alpha-5}x_i^{\alpha-2}+\Gamma(\alpha-1)} \quad (16)$$

$$\frac{\partial\ell(\alpha,\theta)}{\partial\alpha} = \sum_{i=1}^{n}\frac{6\theta^{\alpha-5}\log(\theta)x_i^{\alpha-2}+6\theta^{\alpha-5}\log(x_i)x_i^{\alpha-2}+\Gamma'(\alpha-1)}{6\theta^{\alpha-5}x_i^{\alpha-2}+\Gamma(\alpha-1)} \quad (17)$$

To find the Maximum Likelihood Estimates (MLEs) $\hat{\alpha}$ and $\hat{\theta}$, where $\alpha > 1$ and $\theta > 0$, we set equations (16) and (17) to zero and solve them simultaneously. However, due to the complexity of the equations, obtaining analytical expressions for $\hat{\alpha}$ and $\hat{\theta}$ are not feasible. Therefore, we opted to compute the MLEs numerically using R software (2025).

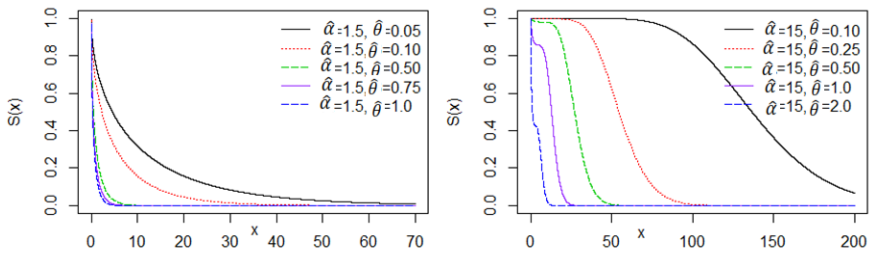


Figure 3. Survival curves of the distribution for some selected values of parameters

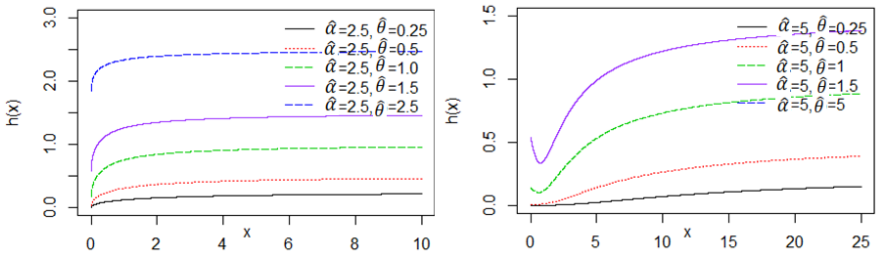


Figure 4. Hazard curves of the distribution for some selected values of parameters

Models' selection criterion

To evaluate how well the proposed distribution performs in comparison to existing distributions, we employ several measures of goodness of fit: -2loglikelihood (-2InL), Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC), and Bayesian Information Criteria (BIC). It is important to note that the most favorable model is the one with the lowest values for -2lnL, AIC, CAIC, and BIC. These statistics can be computed as follows.

$$AIC = -2log\ell(\hat{\theta}) + 2p$$

$$BIC = -2log\ell(\hat{\theta}) + p\log(n)$$

$$CAIC = -2\log[\ell(\hat{\theta})] + p[\log(n) + 1],$$

where p is the number of parameters in the model and ℓ is the loglikelihood.

Simulation Studies

In this paragraph, we analyze the behaviours of $\hat{\alpha}_{MLE}$ and $\hat{\theta}_{MLE}$ of the EG-LD distribution using a brief simulation. The simulation is carried out by generating random samples (with 1000 iterations), say n = 50, 100,..., 1000, from the EG-LD distribution using the inverse cdf method.

Simulations were conducted for two combinations: $\alpha = 5.0, \theta = 0.5\ and\ \alpha = 15.0, \theta = 1.0$. We used two statistical methods to compare the performance of $\hat{\alpha}_{MLE}$ and $\hat{\theta}_{MLE}$. These measures comprised the bias and mean square error (MSE), with mathematical formulations provided, respectively, by

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

and

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

The aforementioned assessment criteria are computed for $\hat{\alpha}_{MLE}$ also. Table 1 shows numerical simulation findings. The simulation of the EG-LD distribution shows that as n tends to infinity,

1. The values of $\hat{\alpha}_{MLE}$ and $\hat{\theta}_{MLE}$ become more stable and closer to their genuine values.

Table 1. The numerical illustration of the Simulation Study of the EG-LD distribution

	n	Parameters	MLEs	MSEs	Biases
α = 5 θ = 0.5	50	α	5.1427	0.7293	0.1427
		θ	0.5153	0.0111	0.0153
	100	α	5.0913	0.3963	0.0913
		θ	0.5102	0.0060	0.0102
	150	α	5.0697	0.2600	0.0697
		θ	0.5079	0.0040	0.0079
	200	α	5.0617	0.2044	0.0617
		θ	0.5077	0.0032	0.0077
	300	α	5.0235	0.1245	0.0235
		θ	0.5025	0.0019	0.0025
	500	α	5.0159	0.0820	0.0159
		θ	0.5018	0.0012	0.0018
	700	α	5.0132	0.0573	0.0132
		θ	0.5012	0.0009	0.0012
	900	α	5.0091	0.0433	0.0091
		θ	0.5022	0.0006	0.0022
	1000	α	5.0027	0.0397	0.0027
		θ	0.5009	0.0006	0.0009
α = 15 θ = 1.0	50	α	15.0909	2.6006	0.0909
		θ	1.0058	0.0117	0.0058
	100	α	14.9888	1.3434	-0.0111
		θ	0.9979	0.0062	-0.0021
	150	α	15.0275	0.8596	0.0275
		θ	1.0033	0.0040	0.0033
	200	α	15.0362	0.6232	0.0362
		θ	1.0026	0.0030	0.0026
	300	α	14.9898	0.4026	-0.0101
		θ	0.9994	0.0019	-0.0005
	400	α	15.0130	0.3021	0.0130
		θ	1.0010	0.0014	0.0010
	500	α	15.0036	0.2515	0.0036
		θ	1.0003	0.0012	0.0003
	700	α	14.9938	0.1628	-0.0061
		θ	0.9995	0.0007	-0.0005
	900	α	14.9943	0.1205	-0.0056
		θ	0.9998	0.0005	-0.0002
	1000	α	14.9969	0.1061	-0.0030
		θ	0.9999	0.0004	-0.0001

2. MSEs for $\hat{\alpha}_{MLE}$ and $\hat{\theta}_{MLE}$ dropped to zero.
3. Biases for $\hat{\alpha}_{MLE}$ and $\hat{\theta}_{MLE}$ decreased to zero.

Therefore, the MLEs become very accurate and unbiased as the sample size increases, confirming the asymptotic properties of MLE.

Applications

Three datasets were used to assess the applicability of the EG-L distribution. In this evaluation, we compared the goodness-of-fit of the proposed EG-L distribution with several existing distributions, which include the New Generalized Two-Parameter Lindley Distribution (NGTPLD) introduced by (Ekhosuehi *et al*. 2018), the Generalized Lindley Distribution 1 (GLD1) introduced by (Nadarajah *et al*., 2011), the Generalized Lindley Distribution 2 (GLD2) introduced by (Zakerzadeh and Dolati 2009), the Generalized Lindley Distribution 3 (GLD3) introduced by (Shanker and Mishra, 2013), which they called quasi-Lindley distribution, and the Generalized Lindley Distribution 4 (GLD4) introduced by (Abouammoh *et al*. 2015). Data Set 1: The data set represents the remission times (in months) of a random sample of 128 bladder cancer patients reported in Lee and Wang (2003). Data Set 2: The data set represents the survival times (in days) of 73 patients who were diagnosed with acute bone cancer (Mansour *et al*., 2020). Data Set 3: The dataset reported by Makkar (2013) represents the survival times of a group of patients suffering from Head and Neck cancer disease and treated using a combination of radiotherapy and chemotherapy (RT+CT).

The results of the EG-L distribution and other competing distributions for the datasets are displayed in Tables 2 - 4. The EG-L distribution shows the smallest -2log-likelihood compared to the other models considered, indicating that the proposed model offers the most suitable fit for the given datasets among all models. It's important to note that choosing the model with the lowest -2log-likelihood is also synonymous with selecting the optimal model based on AIC, BIC, and CAIC when models have an equal number of parameters. Essentially, the model with the lowest information criteria is considered the best.

Table 2. Parameter estimations and goodness of fit test statistics for dataset 1: bladder cancer patients (n=128). See Lee and Wang (2003) for complete data

Models	Parameter estimates (s.e)	-2logL	AIC	AICC	HQIC	BIC
EG-LD	$\hat{\alpha} = 2.1723$ $\hat{\theta} = 0.1252$	826.7358	830.7358	830.8318	833.0534	836.4399
NGTPLD	$\hat{\alpha} = 1.1891$ $\hat{\theta} = 0.1247$	826.8288	830.8288	830.9248	833.1463	836.5328
GLD1	$\hat{\alpha} = 0.7336$ $\hat{\theta} = 0.1648$	832.5718	836.5719	836.6679	838.8894	842.2759
GLD2	$\hat{\alpha} = 0.1734$ $\hat{\theta} = 0.1253$ $\hat{\beta} = 668.84$	826.7568	832.7568	832.9504	836.2332	841.3129
GLD3	$\hat{\alpha} = 663.69$ $\hat{\theta} = 0.1069$	828.6838	832.6839	832.7799	835.0015	838.3879
GLD4	$\hat{\alpha} = 1.5690$ $\hat{\theta} = 0.1533$	832.3136	836.3136	836.4096	838.6312	842.0177

Table 3. Parameter estimations and goodness of fit test statistics for dataset 2: Acute bone cancer dataset (n= 73). See Mansour *et al.* (2020) for complete data

Models	Parameter estimates (s.e)	-2logL	AIC	AICC	HQIC	BIC
EG-LD	$\hat{\alpha} = 1.7456$ $\hat{\theta} = 0.1986$	334.5374	338.5374	338.7088	340.3629	343.1183
NGTPLD	$\hat{\alpha} = 0.7463$ $\hat{\theta} = 0.2087$	335.3726	339.3726	339.544	341.1982	343.9535
GLD1	$\hat{\alpha} = 0.5088$ $\hat{\theta} = 0.2877$	356.7712	360.7712	360.9426	362.5967	365.3521
GLD2	$\hat{\alpha} = 0.7457$ $\hat{\theta} = 0.1987$ $\hat{\beta} = 0.0013$	334.5312	340.5312	340.8791	343.2696	347.4026
GLD3	$\hat{\alpha} = 6808.97$ $\hat{\theta} = 0.2663$	339.1788	343.1788	343.3503	345.0045	347.7598
GLD4	$\hat{\alpha} = 1.3788$ $\hat{\theta} = 0.3049$	358.8346	362.8346	363.0059	364.6601	367.4154

Table 4. Parameter estimations and goodness of fit test statistics for dataset 3: Head and Neck cancer disease (RT+CT) (n=44). See Makkar *et al*. (2013) for the complete data

Models	Parameter estimates (s.e)	-2logL	AIC	AICC	HQIC	BIC
EG-LD	$\hat{\alpha} = 2.0075$ $\hat{\theta} = 0.0045$	564.0109	568.0109	568.3036	569.3342	571.5793
NG2PLD	$\hat{\alpha} = 1.0499$ $\hat{\theta} = 0.0047$	564.0242	568.0242	564.3169	569.3475	571.5926
GLD1	$\hat{\alpha} = 0.5063$ $\hat{\theta} = 0.0059$	566.0782	570.0781	570.3708	571.4014	573.6465
GLD2	$\hat{\alpha} = 0.0901$ $\hat{\theta} = 0.0049$ $\hat{\beta} = 1.6228$	564.3254	570.3254	570.9254	572.3104	575.678
GLD3	$\hat{\alpha} = 39.139$ $\hat{\theta} = 0.0046$	564.0464	568.0464	568.3391	569.3697	571.6148
GLD4	$\hat{\alpha} = 1.0645$ $\hat{\theta} = 0.0048$	564.3916	568.3917	568.6844	569.715	571.9601

Conclusion

In this study, a new lifetime distribution known as the EG-Lindley (EG-L) distribution was developed, with a specific focus on its applicability to survival data modelling. A detailed investigation of its statistical properties, including the probability density function (pdf), cumulative distribution function (cdf), moments, survival function, and hazard rate function, was performed. Graphical illustrations were employed to demonstrate the flexibility and interpretability of the model parameters, particularly how changes in the shape (α) and scale (θ) parameters influence the distribution's behavior. Survival and hazard function plots revealed that the EG-L distribution can exhibit both increasing and decreasing hazard rates, which makes it suitable for modelling a broad range of biomedical and reliability data. The capacity to capture early failure rates and long-term survival simultaneously strengthens its practical appeal, especially in clinical and epidemiological studies. To empirically validate the performance of the EG-L distribution, three real-world datasets were analyzed: bladder cancer patients ($n = 128$), acute bone cancer patients ($n = 73$), and head and neck cancer patients undergoing RT+CT ($n = 44$). The goodness-of-fit performance of the EG-L model was benchmarked against several existing generalizations of the Lindley distribution, including the New Generalized Two-Parameter Lindley Distribution, and four other variants generalized Lindley distribution. Across all datasets, the EG-L distribution consistently yielded the lowest values for -2log-likelihood, as well as the Akaike Information Criterion, Bayesian Information Criterion, Corrected AIC, and Hannan-Quinn Information Criterion, either absolutely or competitively. This dominance in model selection metrics strongly supports the assertion that the EG-L distribution provides a superior fit for the analyzed survival data, making it a compelling alternative for practitioners dealing with time-to-event data. As such, it holds promise for broader application in the modelling of lifetimes in medical, biological, and industrial reliability studies. Future research could extend the EG-L distribution to regression frameworks such as AFT. The model also shows promise for personalized medicine and reliability studies.

Data Availability: The data used to support the findings of this study are within the references cited.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

Funding: This research received no external funding.

Ethical approval: Not applicable.

Informed consent: Not applicable.

References

Abouammoh, A. M., Alshangiti, A. M., & Ragab, I. E. (2015). A new generalized Lindley distribution. *Journal of Statistical Computation and Simulation*, *85*(18), 3662-3678. https://doi.org/10.1080/00949655.2014.995101

Aderoju, S. (2021). Samade probability distribution: Its properties and application to real lifetime data. *Asian Journal of Probability and Statistics*, *14*(1), 1-11.

Aderoju, S. A., & Babaniyi, O. (2023). Power Samade distribution: Its properties and application to real lifetime data. *Nigerian Journal of Science and Environment*, *21*(1), 237-248.

Aderoju, S. A., Aleshinloye, N. I., Taiwo, B. L., & Sanni, B. I. (2023). A new lifetime distribution and its application to cancer data. *Journal of Biostatistics and Epidemiology*, *9*(4), 451-460.

Aderoju, S., & Adeniyi, I. (2022). On power generalized Akash distribution with properties and applications. *Journal of Statistical Modelling & Analytics*, *4*(1).

Ahsanullah, M., Ghitany, M. E., & Al-Mutairi, D. K. (2017). Characterization of the Lindley distribution by truncated moments. *Communications in Statistics-Theory and Methods*, *46*(12), 6222-6227.

Aleshinloye, N. I., Aderoju, S. A., Abiodun, A. A., & Taiwo, B. L. (2023). A new generalized Gamma–Weibull distribution and its applications. *Al-Bahir Journal for Engineering and Pure Sciences*, *2*(2), 90-100.

Edith, U. U., Ebele, T. U., & Henrietta, A. I. (2019). A two-parameter Rama distribution. *Earthline Journal of Mathematical Sciences*, *2*(2), 365-382.

Ekhosuehi, N., Opone, F., & Odobaire, F. (2018). A new generalized two-parameter Lindley distribution. *Journal of Data Science*, *16*(3), 549-566.

Everitt, B. S., & Hand, D. J. (1981). Mixtures of normal distributions. In *Finite mixture distributions* (pp. 25-57). Springer. https://doi.org/10.1007/978-94-009-5897-5_2

Ganaie, R. A., Rajagopalan, V., & Rather, A. A. (2020). On weighted two-parameter quasi-Shanker distribution with properties and its applications. *International Journal of Statistics and Reliability Engineering*, *7*(1), 1-12.

Gemeay, A. M., Halim, Z., Abd El-Raouf, M. M., Hussam, E., Abdulrahman, A. T., Mashaqbah, N. K., Aljohani, H. M., & Makumi, N. (2023). General two-parameter distribution: Statistical properties, estimation, and application on COVID-19. *PLOS ONE*, *18*(2), e0281474. https://doi.org/10.1371/journal.pone.0281474

Ghitany, M. E., Al-Mutairi, D. K., Balakrishnan, N., & Al-Enezi, L. J. (2013). Power Lindley distribution and associated inference. *Computational Statistics & Data Analysis*, *64*, 20-33. https://doi.org/10.1016/j.csda.2013.02.026

Hamed, D., & Alzaghal, A. (2021). New class of Lindley distributions: Properties and applications. *Journal of Statistical Distributions and Applications*, *8*, Article 8.

Irshad, M. R. (2017). New extended generalized Lindley distribution: Properties and applications. *Statistica*, *77*(1), 33-52.

Lee, E. T., & Wang, J. W. (2003). *Statistical methods for survival data analysis* (3rd ed.). Wiley.

Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society: Series B (Methodological)*, *20*(1), 102-107. https://doi.org/10.1111/j.2517-6161.1958.tb00278.x

Makkar, P., Srivastava, P. K., Singh, R. S., & Upadhyay, S. K. (2014). Bayesian survival analysis of head and neck cancer data using lognormal model. *Communications in Statistics-Theory and Methods*, *43*(2), 392-407. https://doi.org/10.1080/03610926.2012.664233

Mansour, M., Yousof, H. M., Shehata, W. A., & Ibrahim, M. (2020). A new two-parameter Burr XII distribution: Properties, copula, different estimation methods, and modelling acute bone cancer data. *Journal of Nonlinear Science and Applications*, *13*(5), 223-238. https://doi.org/10.22436/jnsa.013.05.01

Miller, J. A. (2012). Species distribution models: Spatial autocorrelation and non-stationarity. *Progress in Physical Geography*, *36*(5), 681-692. https://doi.org/10.1177/0309133312442522

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

Nadarajah, S., Cordeiro, G. M., & Ortega, E. M. (2015). The Zografos-Balakrishnan-G family of distributions: Mathematical properties and applications. *Communications in Statistics-Theory and Methods*, *44*(1), 186-215.

Pararai, M., Oluyede, B. O., & Warahena-Liyanage, G. (2015). Kumaraswamy Lindley-Poisson distribution: Theory and applications. *Asian Journal of Mathematics and Applications*, 2015, ama0229.

R Core Team. (2025). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.

Shanker, R. (2020). A new quasi-Sujatha distribution. *Statistics in Transition New Series*, *21*(3), 53-71. https://doi.org/10.21307/stattrans-2020-044

Shanker, R., & Mishra, A. (2013). A quasi-Lindley distribution. *African Journal of Mathematics and Computer Science Research*, *6*(4), 64-71.

Wolstenholme, L. C. (2018). *Reliability modelling: A statistical approach*. Routledge. https://doi.org/10.1201/9780203740958

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