RT&A, No 1 (72)
Volume 18, March 2023

A Generalization of Lindley Distribution: Characterizations, Methods of Estimation and Applications

A. Mohammed Shabeer[1], Bindu Krishnan[2], K. Jayakumar[3]

•

[1,3] Department of Statistics, University of Calicut, India
[2] Department of Data Science, CS & IT, Jain(Deemed-to-be-University), India
md.shabeer.edv@gmail.com, bindusruthy@gmail.com, jkumar19@rediffmail.com

Abstract

Lindley distribution is a lifetime model with application in survival analysis and reliability theory problems often centred around its increasing hazard rate function and flexibility over exponential distribution. In this paper, we introduce a new generalization of the Lindley distribution referred to as Lindley Truncated Negative binomial (LTNB) distribution. The LTNB model has increasing, decreasing and upside-down bathtub(UBT) shapes for the hazard rate function. Various properties of the LTNB distribution are studied including moments, quantiles, and stochastic ordering. Characterizations of the new distribution are obtained. Maximum likelihood, Cramer-von-Mises, ordinary and weighted least squares methods of estimation are utilized to obtain the estimators of the model parameters. A simulation study is carried out to assess and compare the performance of different estimates. An autoregressive time series model with the LTNB as marginal is developed. The model is fitted to bladder cancer data set to show how the proposed model works in practice.

Keywords: Autoregressive Models, Characterizations, Lindley distribution, Maximum Likelihood, Stochastic Ordering.

1. Introduction

As a counter-example to fiducial statistics, the Lindley distribution, first put forth by [16], is one of the crucial lifetime distributions in the framework of Bayesian statistics. The Lindley distribution has an increasing hazard rate function; however, in real-life situations, the models exhibit non-monotone hazard rate shapes. For example, in the case of a serious illness condition such as cancer, the hazard rate increases and then decreases. Similarly, in the engineering field, the quality of production by untrained workers follows a similar pattern. As a result, many researchers have developed models with non-monotone hazard rates over time. Models with bathtub shapes can be found in [22]. Most often, upside-down bathtub (UBT) shape hazard rate models are used to model medical data, such as patient data for bladder cancer and lung cancer (see, [5] and [15]). In this paper, we introduce a generalized Lindley distribution with increasing, decreasing and UBT shapes for the hazard rate function. A positive random variable X is said to have Lindley distribution with the parameter α, denoted as $L(\alpha)$, if it has the probability density function (pdf)

$$f(x) = \frac{\alpha^2}{\alpha+1}(1+x)e^{-\alpha x}; x > 0, \alpha > 0, \tag{1}$$

The cumulative distribution function (cdf) is given by

$$F(x) = 1 - \frac{\alpha+1+\alpha x}{\alpha+1}e^{-\alpha x}; x \geq 0. \tag{2}$$

Ghitany[8] studied the statistical characteristics of the Lindley distribution and demonstrated its flexibility in modelling lifetime data over exponential distribution. Generalized Lindley distribution was later developed by [20], who also studied its mathematical characteristics. Weighted Lindley distribution was introduced by [9]. An extended Lindley distribution was proposed and investigated by [3]. In 2013, [6] introduced a generalized Lindley distribution. Later Beta Lindley distribution was examined by [18]. The Lindley-Exponential distribution was first proposed by [4]. Inverse Lindley distribution was extended by [23]. The Odd Lindley Burr XII distribution was proposed by [14]. The generalized two-parameter Lindley distribution was introduced by [7]. Unit-Lindley distribution was studied in [2].
The rest of the paper is as follows: The Lindley Truncated Negative binomial (LTNB) distribution is introduced as a compounded distribution in Section 2. Section 3 discuss the statistical properties of the LTNB distribution such as moments, quantile function, skewness, and kurtosis. Characterizations of the LTNB distribution are obtained in section 4.In Section 5, we investigate the stochastic ordering property of LTNB. In section 6, we use the maximum likelihood method, least squares method, weighted least squares, and Cramer-von-Mises-estimator for estimating the parameters of the new distribution. In section 7 we carry out a simulation study to evaluate the performance of Maximum likelihood estimation and other estimation methods. The LTNB distribution is used to model remission times of bladder cancer patients in Section 8. It is demonstrated that the LTNB distribution fits this data set better than the other well-known competitors. In section 9, we develop a first-order autoregressive minification process with the LTNB distribution as the marginal distribution.

2. Lindley Truncated Negative binomial distribution

We propose a new generalization for Lindley distribution with relatively simple expressions for the survival function, hazard rate function, and quantile's. The hazard rate function is more flexible than the previous generalizations and the proposed model has a closed-form for distribution function. In addition, the new distribution fits some real-world data sets better than existing Lindley distribution generalizations.
In 1997, [17] developed a method of adding a tilt parameter to a distribution to extend existing distributions. Many distributions were extended and published in the literature as a result of this technique. As a generalization of the above technique, [21] proposed a novel family of distributions through truncated negative binomial distribution. It's important to remember that these distributions arise as the distribution of random minimum or random maximum. The cdf is given by

$$G(z) = \frac{\gamma^n}{1-\gamma^n}[(\bar{F}(z) + \gamma F(z))^{-n} - 1]; z \in \mathbf{R}, \gamma, \theta > 0, \tag{3}$$

where $F(.)$ is the cdf of baseline distribution . If $F(x)$ follows Lindly with cdf (2) we have Lindley Truncated Negative binomial(LTNB) distribution, LTNB(α, γ, θ), having pdf

$$g(z; \alpha, \gamma, \theta) = \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta\gamma^\theta(1+z)e^{-\alpha z}}{(1-\gamma^\theta)\left[(\alpha+1) - (1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}\right]^{\theta+1}}; z > 0 \quad \alpha, \gamma, \theta > 0. \tag{4}$$

The cumulative distribution function of LTNB(α, γ, θ) is given by

$$G(z; \alpha, \gamma, \theta) = \frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{1-\gamma^\theta}\left[\frac{\alpha+1}{(\alpha+1) - (1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}}\right]^\theta \tag{5}$$

and hence the survival function is

$$\bar{G}(z; \alpha, \gamma, \theta) = \frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1) - (1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}}\right]^\theta - 1\right\} \tag{6}$$

Remark: The LTNB(α, γ, θ) distribution reduces to the Lindley distribution when $\theta = 1$ and $\gamma \longrightarrow 1$.

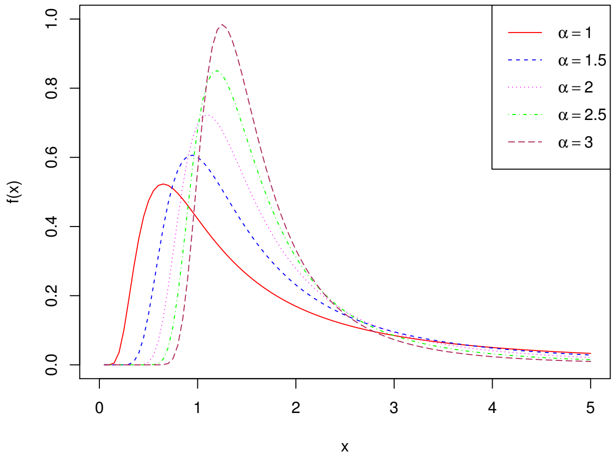

Figure 1: *LTNB*(α, γ, θ) *pdf when* $\gamma = 0.5, \theta = 0.5$.

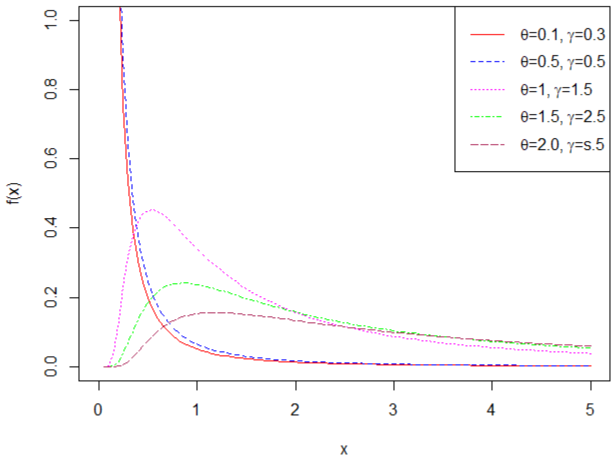

Figure 2: *LTNB*(α, γ, θ) *pdf when* $\alpha = 1$.

The hazard rate function of LTNB distribution is

$$h(z; \alpha, \gamma, \theta) = \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta(1+z)e^{-\alpha z}}{[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]\left[(\alpha+1)^\theta - [(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]^\theta\right]} \quad (7)$$

and the reverse hazard rate function is

$$r(z; \alpha, \gamma, \theta) = \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta(1+z)e^{-\alpha z}}{[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]\left[[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]^\theta - \gamma^\theta(\alpha+1)^\theta\right]}. \quad (8)$$

Figure 3 depicts hazard rate function graphs for various parameter values. The graphs showing decreasing, non-decreasing, and UBT shapes for hazard rate function.

3. Statistical Properties

We discuss about moments, simulation, quantiles, skewness, and kurtosis in this section.

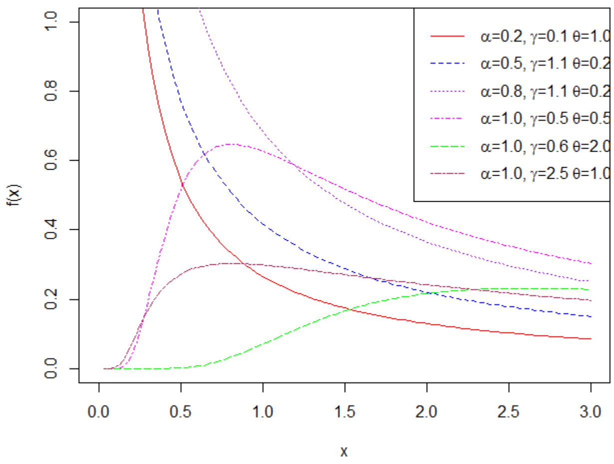

Figure 3: *$LTNB(\alpha, \gamma, \theta)$ hazard rate function for various parameter values.*

3.1. Moments

Let $X \sim LTNB(\alpha, \gamma, \theta)$, then its r^{th} moment with respect to the origin is given by,

$$E[X^r] = \int_0^\infty x^r \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta\gamma^\theta(1+x)e^{-\alpha x}}{(1-\gamma^\theta)\left[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}\right]^{\theta+1}}dx$$
$$= \sum_{k=0}^{\infty} \frac{\alpha^2(1-\gamma)^{k+1}\theta\gamma^\theta}{(1-\gamma^\theta)(\alpha+1)^{k+1}} \binom{k+\theta}{\theta} \int_0^\infty x^r(1+x)e^{-\alpha x(1+k)}(\alpha+1+\alpha x)^k dx.$$

3.2. Simulation, Quantiles and Median

In order to generate random numbers from a LTNB(α, γ, θ) distribution, we use

$$X = \frac{-1}{\alpha} - 1 - \frac{1}{\alpha}W_{-1}(-e^{-(\alpha+1)}\frac{(\alpha+1)}{1-\gamma}(1-\gamma((1-\gamma^\theta)Y+\gamma^\theta)^{-1/\theta})) \tag{9}$$

where $W_{-1}(.)$ is the negative Limbert W function and $Y \sim$ U(0,1).
The qth quantile of the LTNB(α, γ, θ) is

$$X = \frac{-1}{\alpha} - 1 - \frac{1}{\alpha}W_{-1}(-e^{-(\alpha+1)}\frac{(\alpha+1)}{1-\gamma}(1-\gamma((1-\gamma^\theta)q+\gamma^\theta)^{-1/\theta})), \tag{10}$$

and thus the median of LTNB(α, γ, θ) is,

$$X = \frac{-1}{\alpha} - 1 - \frac{1}{\alpha}W_{-1}(-e^{-(\alpha+1)}\frac{(\alpha+1)}{1-\gamma}(1-\gamma((1-\gamma^\theta)\frac{1}{2}+\gamma^\theta)^{-1/\theta})).$$

3.3. Skewness and Kurtosis

According to [13], the distribution skewness can be calculated using the below equation

$$S = \frac{Q\left(\frac{3}{4}\right) - 2Q(\frac{1}{2}) + Q(\frac{1}{4})}{Q(\frac{3}{4}) - Q(\frac{1}{4})}$$

and according to [19], the kurtosis of the LTNB distribution is as follows

$$K = \frac{Q\left(\frac{7}{8}\right) - Q(\frac{5}{8}) + Q(\frac{3}{8}) - Q(\frac{1}{8})}{Q(\frac{6}{8}) - Q(\frac{2}{8})},$$

RT&A, No 1 (72)
Volume 18, March 2023

where $Q(.)$ is the quantile function of X as defined by (10). These metrics are less susceptible to outliers.

4. Characterizations of LTNB distribution

Many characterization results, established in various ways, can be found in the literature. In this section, the ratio of truncated moments and the hazard rate function are used to characterize the LTNB distribution.

4.1. Characterizations based on two truncated moments

This section discusses how to characterise the LTNB distribution using the ratio of two truncated moments. In our first characterization, we apply the Theorem 1 below established in [10].

Theorem 1. Let $(\Omega, \mathcal{F}, P)$ be a given probability space and let $H = [d, e]$ be an interval for some $d < e$ $\,(d = -\infty, e = \infty$might as well be allowed). Let $X : \Omega \to H$ be a continuous random variable with the distribution function F and let q_1 and q_2 be two real functions on H such that

$$E[q_2(X)|X \geq x] = E[q_1(X)|X \geq x]\zeta(x), x \in H,$$

is defined with some real function ζ. Assume that $q_1, q_2 \in C^1(H), \zeta \in C^2(H)$ and F is twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $\zeta q_1 = q_2$ has no real solution in the interior of H. Then F is uniquely determined by the functions q_1, q_2 and ζ, particularly

$$F(x) = \int_a^x C|\frac{\zeta'(u)}{\zeta(u)q_1(u) - q_2(u)}|exp(-s(u))du,$$

where the function s is a solution of the differential equation $s' = \frac{\zeta' q_1}{\zeta q_1 - q_2}$ and C is the normalization constant, such that $\int_H dF = 1$.

Theorem 2. Let X be a positive real valued continuous random variable with pdf $f(x)$. Denote $q_1(x) = (1-(1-\gamma)\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x})^{\theta+1}$ and $q_2(x) = q_1(x)\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}$ for $x > 0, \alpha > 0, \theta > 0, \gamma > 0$. Then, the random variable X has pdf (4) if and only if the function $\zeta(.)$ defined in Theorem 1 is of the form

$$\zeta(x) = \frac{1}{2}\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}$$

Proof. Suppose the random variable X has the pdf (4). Then

$$(1-F(x))E\left[q_1(X)|X \geq x\right] = C\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}, \quad x > 0$$

and

$$(1-F(x))E\left[q_2(X)|X \geq x\right] = \frac{C}{2}\left(\frac{1+\alpha+\alpha x}{\alpha+1}\right)^2 e^{-2\alpha x}, \quad x > 0$$

where $C = \left(\frac{\gamma^\theta\theta(1-\gamma)}{1-\gamma^\theta}\right)$.Further,

$$\zeta(x)q_1(x) - q_2(x) = -\frac{1}{2}q_1(x)\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x} < 0, \;\; x > 0.$$

Conversely, if ζ is of the above form, then

$$s'(x) = \frac{\zeta'(x)q_1(x)}{\zeta(x)q_1(x) - q_2(x)} = \frac{\alpha^2(1+x)}{1+\alpha+\alpha x}, \;\; x > 0.$$

Thus

$$s(x) = -log(\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}), \;\; x > 0.$$

Now from Theorem 1, result follows.

■

4.2. Characterization based on the hazard rate function

For $\theta = 1$, we characterise the LTNB distribution using the hazard rate function.

Theorem 3. Let X be a positive real valued continuous random variable with hazard rate function $h(x)$. The random variable X has LTNB distribution, for $\theta = 1$, if and only if $h(x)$ satisfies the differential equation

$$h^{'}(x) - \frac{h(x)}{(1+x)(1+\alpha+\alpha x)} = \frac{\alpha^4(1-\gamma)(1+x)^2 e^{-\alpha x}}{(1+\alpha+\alpha x)(1-\frac{(1-\gamma)(1+\alpha+\alpha x)e^{-\alpha x}}{1+\alpha})^2}, \quad x > 0.$$

Proof. If X has pdf (4) , the above differential equation clearly holds.
Now suppose the differential equation holds then,

$$\frac{d}{dx}\left\{\frac{h(x)(1+\alpha+\alpha x)}{1+x}\right\} = \alpha^2 \frac{d}{dx}\left\{\frac{1}{1-(\frac{(1-\gamma)(1+\alpha+\alpha x)e^{-\alpha x}}{1+\alpha})}\right\}$$

which implies,

$$h(x) = \frac{\alpha^2(1+x)}{(1+\alpha+\alpha x)} \frac{1}{1-(\frac{(1-\gamma)(1+\alpha+\alpha x)e^{-\alpha x}}{1+\alpha}}$$

which is the hazard rate function of LTNB distribution.

■

5. Stochastic Ordering

Let X and Y be two random variables with distribution functions of F_1 and F_2, respectively, with corresponding pdfs of f_1 and f_2. Then it is stated that X is less than Y in ,
i) stochastic order (denoted as $X \leq_{st} Y$) if $F_1(x) \geq F_2(x)$ for all x;
ii) hazard rate order (denoted as $X \leq_{hr} Y$) if $(1 - F_1(x)/(1 - F_2(x))$ is decreasing in $x \geq 0$;
ii) likelihood ratio order (denoted as $X \leq_{lr} Y$) if $f_1(x)/f_2(x)$ is decreasing in $x \geq 0$;
iv) reverse hazard rate order (denoted as $X \leq_{rhr} Y$) if $F_1(x)/F_2(x)$ is decreasing in $x \geq 0$.
The four orderings have a relationship with one another; for further information, see [24],

$$X \leq_{lr} Y \implies X \leq_{hr} Y \implies X \leq_{rhr} Y \implies X \leq_{st} Y. \tag{11}$$

For $\gamma_2 > \gamma_1$,
let $X \sim LTNB(\alpha, \gamma_1, \theta)$ and $Y \sim LTNB(\alpha, \gamma_2, \theta)$. Then

$$\frac{f_X(y)}{f_Y(y)} = \frac{\gamma_1^\theta(1-\gamma_2^\theta)(1-\gamma_1)\left[(\alpha+1)-(1-\gamma_2)(\alpha+1+\alpha z)e^{-\alpha y}\right]^{\theta+1}}{\gamma_2^\theta(1-\gamma_1^\theta)(1-\gamma_2)\left[(\alpha+1)-(1-\gamma_1)(\alpha+1+\alpha y)e^{-\alpha y}\right]^{\theta+1}}$$

Since $\gamma_2 > \gamma_1$,

$$\frac{d}{dx}\left[\frac{f_X(y)}{f_Y(y)}\right] = \frac{\gamma_1^\theta(1-\gamma_2^\theta)(1-\gamma_1)\alpha^2 e^{-\alpha y}(1+y)(\alpha+1)(\gamma_1-\gamma_2)}{(1-\gamma_2)\gamma_2^\theta(1-\gamma_1^\theta)}$$

$$\times \frac{\left[(\alpha+1)-(1-\gamma_2)(\alpha+1+\alpha y)e^{-\alpha y}\right]^\theta}{\left[(\alpha+1)-(1-\gamma_1)(\alpha+1+\alpha y)e^{-\alpha y}\right]^{\theta+2}},$$

$$< 0.$$

$\implies f_X(y)/f_Y(y)$ is decreasing in y.
That is $X \leq_{lr} Y$. From (11), the remaining ordering follows.

6. Different methods of estimation

In this section we use maximum likelihood estimation, least squares, weighted least squares and cramer-von-Mises estimation methods to estimate the parameters of the LTNB distribution.

6.1. Maximum likelihood Estimation

Let $x_1, x_2, ..., x_n$ be an observed random sample from LTNB distribution with unknown parameter vector $\nu = (\alpha, \gamma, \theta)^T$. Then the log-likelihood function is

$$\begin{aligned} log\ \ell &= 2nlog\alpha + nlog\theta + nlog(1-\gamma) + n\theta log\gamma + n\theta log(\alpha+1) - nlog(1-\gamma^\theta) - \alpha\sum_{i=1}^n x_i \\ &+ \sum_{i=1}^n log(1+x_i) - (\theta+1)\sum_{i=1}^n log\left[(\alpha+1) - (1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}\right] \end{aligned} \tag{12}$$

The partial derivatives are given by

$$\begin{aligned} \frac{\partial log\ \ell}{\partial\alpha} &= \frac{2n}{\alpha} + \frac{n\theta}{\alpha+1} - \sum_{i=1}^n x_i \\ &-(\theta+1)\sum_{i=1}^n\left[\frac{1-(1-\gamma)e^{-\alpha x_i}(1-\alpha x_i-\alpha x_i^2)}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}}\right], \end{aligned} \tag{13}$$

$$\begin{aligned} \frac{\partial log\ \ell}{\partial\gamma} &= \frac{-n}{1-\gamma} + \frac{n\theta}{\gamma} + \frac{n\theta\gamma^{\theta-1}}{1-\gamma^\theta} \\ &-(\theta+1)\sum_{i=1}^n\left[\frac{(\alpha+1+\alpha x_i)e^{-\alpha x_i}}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}}\right], \end{aligned} \tag{14}$$

$$\begin{aligned} \frac{\partial log\ \ell}{\partial\theta} &= nlog\gamma + \frac{n}{\theta} + nlog(\alpha+1) + \frac{n\gamma^\theta log\gamma}{1-\gamma^\theta} \\ &-\sum_{i=1}^n log\left[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}\right] \end{aligned} \tag{15}$$

Set the score vector to zero,

$$U(\nu) = \left(\frac{\partial log\ell}{\partial\alpha}, \frac{\partial log\ell}{\partial\gamma}, \frac{\partial log\ell}{\partial\theta}\right)^T.$$

$U(\nu) = 0$, and solve them simultaneously to obtain the ML estimators $\hat{\alpha}$, $\hat{\gamma}$, and $\hat{\theta}$. These equations can be numerically solved using statistical software using iterative techniques like the Newton-Raphson algorithm as they cannot be solved analytically. All of the second order derivatives exist for the three-parameter LTNB distribution as well.

6.2. Least squares and weighted least squares estimators

Let $t_1 < t_2 < t_3 < ... < t_n$ be the n ordered random sample from any distribution with cdf $F(t)$. Then we have,

$$E[F(t_i)] = \frac{i}{n+1}.$$

The least squares method minimizes

$$P_{LSE}(\alpha,\gamma,\theta) = \sum_{i=1}^n\left(F(t_i) - \frac{i}{n+1}\right)^2 \tag{16}$$

with respect to the unknown parameters. Here the least squares estimates are obtained by minimizing the following equation with respect to α, γ, θ.

$$P_{LSE}(\alpha,\gamma,\theta) = \sum_{i=1}^n\left(\frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{1-\gamma^\theta}\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - \frac{i}{n+1}\right)^2. \tag{17}$$

Weighted least squares estimates of α, γ, θ are obtained by minimizing the following equation with respect to α, γ, θ.

$$P_{WLSE}(\alpha,\gamma,\theta) = \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left(\frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{1-\gamma^\theta} \left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}} \right]^\theta - \frac{i}{n+1} \right)^2 . \tag{18}$$

6.3. Cramer-von-Mises-estimator(CME)

CME is obtained by minimizing the following equation with respect to α, γ, θ. Here $F(.)$ is the distribution function of LTNB distribution given by (5) .

$$P_{CME}(a,b,\gamma,\theta) = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(t_i) - \frac{2i-1}{2n} \right)^2 . \tag{19}$$

7. Simulation

Monte Carlo simulation is used to evaluate the performance of the maximum likelihood estimation procedure for estimating the LTNB parameters. From the LTNB model we generates samples of sizes $n = 100, 200, 300, 400,$ and 500 for various combinations of $\alpha, \gamma,$ and θ. We ran the simulation 1000 times to determine the MLE's and MSE's of the parameter estimates. Table 1 displays the results. We can observe that the ML estimates are consistent. To investigate the efficiency of least squares estimators, we took samples from the LTNB distribution with $n = 100, 200, 300, 400,$ and 500. We repeated the simulation 1000 times and computed the estimates and corresponding MSE's for the same set of parameter values using three methods. Table 2 displays the results. Table 2 shows that least squares and CME estimators perform similarly and both methods are giving smaller MSE's compared to weighted least squares estimates.

8. Data Analysis

Table 3 contains data on the remission times (in months) of a random sample of 128 bladder cancer patients provided by [15]. Table 4 displays the descriptive statistics for the data.

We will now look at the Total Time on Test (TTT) plot, a graphical method for determining the shape of the data's hazard rate function. The empirical TTT plot is defined as,

$$G(r/n) = \left(\sum_{i=1}^{r} x_{(1)} + (n-r)x_{(r)} \right) / \sum_{i=1}^{n} x_{(i)}, \quad r = 1, 2, .., n$$

where $x_{(i)}$ denote the i^{th} order statistic of the sample. The shape of the hazard rate function would be increasing, decreasing, bathtub-shaped, and UBT if the TTT transform is concave, convex, convex then concave, and concave then concave (see [1]). Figure 3 depicts the data's TTT plot. The hazard rate function is clearly UBT.

The Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion(HQIC) are s the goodness-of-fit metrics that we take into consideration,

$$\begin{aligned} AIC &= 2k - 2log\, \hat{\ell}, \\ AICC &= \frac{2kn}{n-k-1} - 2log\, \hat{\ell}, \\ BIC &= klog\, n - 2log\, \hat{\ell}, \\ HQIC &= 2klog(log\, n) - 2log\, \hat{\ell}, \end{aligned} \tag{20}$$

Table 1: *MLE's and their corresponding MSE's of LTNB distribution parameters.*

α	γ	θ	n	$\hat{\alpha}(MSE(\hat{\alpha}))$	$\hat{\gamma}(MSE(\hat{\gamma}))$	$\hat{\theta}(MSE(\hat{\theta}))$
.8	.2	1.5	100	.9196(.1598)	.2055(.0266)	1.5251(.7527)
			200	.8551(.0888)	.1989(.0158)	1.5342(.7232)
			300	.8813(.1184)	.2001(.0193)	1.5428(.6812)
			400	.8347(.0506)	.2015(.0091)	1.5034(.6691)
			500	.8171(.0382)	.1986(.0064)	1.5098(.5061)
.5	.3	2	100	.5698(.0539)	.2824(.0449)	1.9985(1.2211)
			200	.5298(.0313)	.2726(.0309)	1.9952(.9892)
			300	.5202(.0230)	.2836(.0245)	2.0922(.7610)
			400	.5122(.0192)	.2856(.0222)	2.0962(.5948)
			500	.5087(.0142)	.2942(.0164)	2.0962(.5484)
.3	.2	1.5	100	.3403(.0229)	.2075(.0309)	1.4718(.6438)
			200	.3179(.0111)	.2088(.0201)	1.6118(.4881)
			300	.3141(.0087)	.2047(.0149)	1.5339(.2959)
			400	.3109(.0065)	.2026(.0112)	1.4966(.2363)
			500	.3078(.0049)	.2073(.0092)	1.5387(.1401)

where $\hat{\ell}$ is the likelihood function evaluated at the maximum likelihood estimates, the sample size is n, and k is the number of parameters.
We fitted the LTNB distribution to the data and compared it to the Lindley distribution with pdf (1) and the New Generalized Lindley distribution (NGLD) with pdf

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

Beta Lindely(BL) distribution having pdf,

$$f(x) = \frac{\theta^2(\theta+1+\theta x)^{\beta-1}(1+x)e^{-\theta\beta x}}{B(\alpha,\beta)(\theta+1)^{\beta}}\left(1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right)^{\alpha-1},$$

new generalized two-parameter Lindley distribution (NG2PLD) distribution having pdf,

$$f(x) = \frac{\theta^2}{\theta+1}\left(1+\frac{\theta^{\alpha-2}x^{\alpha-1}}{\Gamma(\alpha)}\right)e^{-\theta x}$$

transmuted Lindley distribution(TLD) having pdf,

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

and exponential distribution(ED) having pdf,

$$f(x) = \alpha e^{-\alpha x}$$

Table 2: *LSE, WLSE, CVME and corresponding MSEs of parameters of LTNB distribution.*

α γ θ	n	$\hat{\alpha}(MSE(\hat{\alpha}))$	$\hat{\gamma}(MSE(\hat{\gamma}))$	$\hat{\theta}(MSE(\hat{\theta}))$
Least squares	100	.6262(.0015)	.2178(.0007)	1.5386(.0043)
.6 .2 1.5	200	.6206(.0009)	.2151(.0005)	1.5408(.0045)
	300	.6173(.0005)	.2141(.0004)	1.5408(.0038)
	400	.6162(.0005)	.2136(.0003)	1.5437(.0039)
	500	.6150(.0004)	.2122(.0002)	1.5441(.0039)
Weighted Least squares	100	.6194(.1305)	.1308(.0288)	1.3082(2.3691))
.6 .2 1.5	200	.6062(.0893)	.1474(.0186)	1.3082(1.7783)
	300	.5902(.0641)	.1595(.0134)	1.3271(.7774)
	400	.5870(.0514)	.1660(.0102)	1.3618(.6277)
	500	.5938(.0478)	.1729(.0089)	1.3942(.5434)
CVM	100	.6261(.0015)	.2178(.0008)	1.5386(.0042)
.6 .2 1.5	200	.6206(.0009)	.2151(.0006)	1.5440(.0046)
	300	.6172(.0005)	.2141(.0004)	1.5408(.0038)
	400	.6162(.0005)	.2135(.0004)	1.5434(.0041)
	500	.6150(.0004)	.2122(.0002)	1.5441(.0039)

The table 5 lists the parameter estimates and goodness of fit statistics for bladder cancer patient data.

The LTNB distribution is more suitable for these data since the values of $-log\hat{l}$, AIC, AICC, BIC, and HQIC for the LTNB distribution are lower than those of the other competing models. Figure 4 displays the fitted densities.

9. Time series Models with LTNB marginals

Several researchers have developed and studied time series models with non-Gaussian marginals (see, for example, [11], [12], and [25]). The following definition is required in order to create the time series model with LTNB marginal distribution

Definition 9.1. A positive real valued random variable X is said to have Marshall-Olkin Lindley Truncated Negative binomial distribution and write $X \stackrel{d}{=}$MOLTNB$(\vartheta, \alpha, \gamma, \theta)$ if it has the survival function

$$\bar{F}_X(x) = \frac{1}{1+\frac{1}{\vartheta}\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^{\theta}-\gamma^{\theta}(\alpha+1)^{\theta}}{\gamma^{\theta}[(\alpha+1)^{\theta}-((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^{\theta}]}\right]}. \tag{21}$$

Theorem 4. The first order auto regressive (AR(1)) process given by

$$X_n = \begin{cases} \varepsilon_n & \text{w.p} \quad \delta \\ min(X_{n-1}, \varepsilon_n) & \text{w.p} \quad 1-\delta \end{cases} \tag{22}$$

Table 3: *Remission times of Bladder Cancer Patient Data.*

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.20	2.23	3.52	4.98
6.97	9.02	13.29	0.40	2.26	3.57	5.06	7.09	9.22	13.80	25.74	0.50
2.46	3.64	5.09	7.26	9.47	14.24	25.82	0.51	2.54	3.70	5.17	7.28
9.74	14.76	26.31	0.81	2.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64
3.88	5.32	7.39	10.34	14.83	34.26	0.90	2.69	4.18	5.34	7.59	10.66
15.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01	1.19	2.75
4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33	5.49	7.66	11.25	17.14
79.05	1.35	2.87	5.62	7.87	11.64	17.36	1.40	3.02	4.34	5.71	7.93
11.79	18.10	1.46	4.40	5.85	8.26	11.98	19.13	1.76	3.25	4.50	6.25
8.37	12.02	2.02	3.31	4.51	6.54	8.53	12.03	20.28	2.02	3.36	6.76
12.07	21.73	2.07	3.36	6.93	8.65	12.63	22.69				

Table 4: *Descriptive Statistics of Cancer data.*

Min.	Q_1	Median	Mean	Q_3	Max.	Var.
0.08	3.348	6.395	9.366	11.838	79.05	110.425

where $0 < \delta < 1; n \geq 1$, defines a stationary AR(1) minification process with LTNB(α, γ, θ) as marginal distribution if and only if ε_n's are i.i.d MOLTNB$(\delta^{-1}, \alpha, \gamma, \theta)$ with $X_0 \stackrel{d}{=} CL(\alpha, \gamma, \theta)$.

Proof. If $\{X_n\}$ is stationary with LTNB(α, γ, θ) marginals, then

$$\begin{aligned}
\bar{F}_{\varepsilon_n}(x) &= \frac{\bar{F}_X(x)}{\delta + (1-\delta)\bar{F}_X(x)} \\
&= \frac{\frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - 1\right\}}{\delta + (1-\delta)\frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - 1\right\}} \\
&= \frac{1}{1+\delta\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}. \qquad (23)
\end{aligned}$$

That is, ε_n's are i.i.d MOLTNB$(\delta^{-1}, \alpha, \gamma, \theta)$.

Conversely, if ε_n's are i.i.d MOLTNB$(\delta^{-1}, \alpha, \gamma, \theta)$ with $X_0 \stackrel{d}{=} CL(\alpha, \gamma, \theta)$, then,

$$\begin{aligned}
\bar{F}_{X_1}(x) &= \delta\bar{F}_{\varepsilon_1}(x) + (1-\delta)\bar{F}_{\varepsilon_1}(x)\bar{F}_{X_0}(x) \\
&= \delta\left\{\frac{1}{1+\delta\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}\right\} + \\
&\quad (1-\delta)\left\{\frac{1}{1+\delta\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}\right\} \\
&\quad \left\{\frac{1}{1+\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}\right\} \\
&= \frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - 1\right\}.
\end{aligned}$$

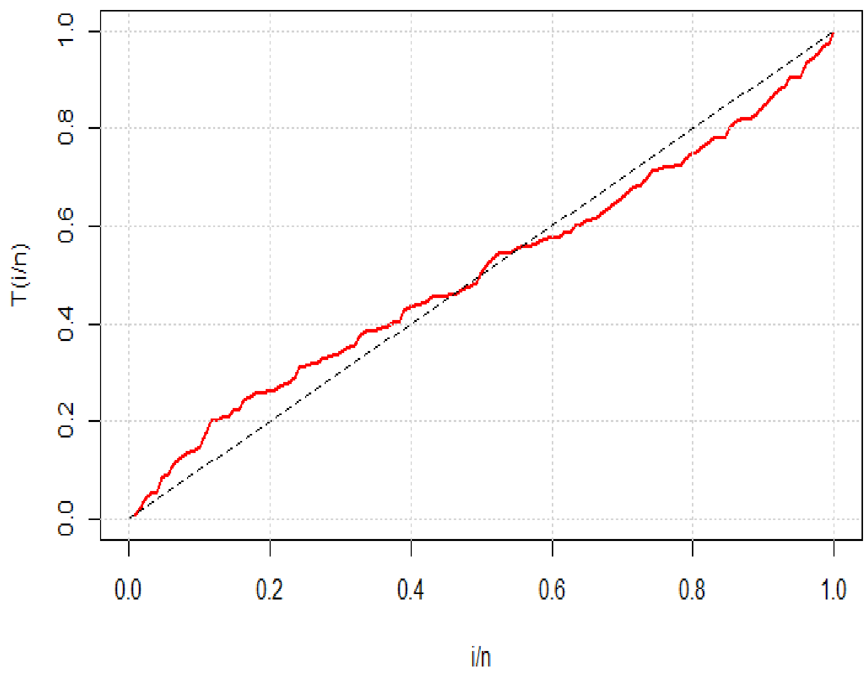

Figure 4: *The empirical TTT plot of the data.*

Table 5: *Parameter estimates and goodness of fit statistics for various models fitted to the data.*

Model	Estimates	$-log\hat{l}$	AIC	AICC	BIC	HQIC
LTNB	$\hat{\alpha} = 0.0617, \hat{\gamma} = 0.1218, \hat{\theta} = 1.5407$	409.23	824.47	824.67	833.03	827.95
NGLD	$\hat{\alpha} = 4.679, \hat{\beta} = 1.324, \hat{\theta} = 0.180$	412.75	831.50	831.69	840.06	834.98
BL	$\hat{\alpha} = 1.340, \hat{\beta} = 0.065, \hat{\theta} = 1.861$	412.80	831.60	831.80	840.16	835.08
NG2PLD	$\hat{\alpha} = 0.8303, \hat{\theta} = 0.2942$	413.37	830.73	830.83	836.44	833.05
TLD	$\hat{\theta} = 0.156, \hat{\lambda} = -0.617$	415.15	834.31	834.41	840.01	836.62
LD	$\hat{\alpha} = 0.1960$	419.53	841.06	841.09	843.89	842.20
ED	$\hat{\alpha} = 0.097$	426.76	855.53	855.56	858.37	856.68

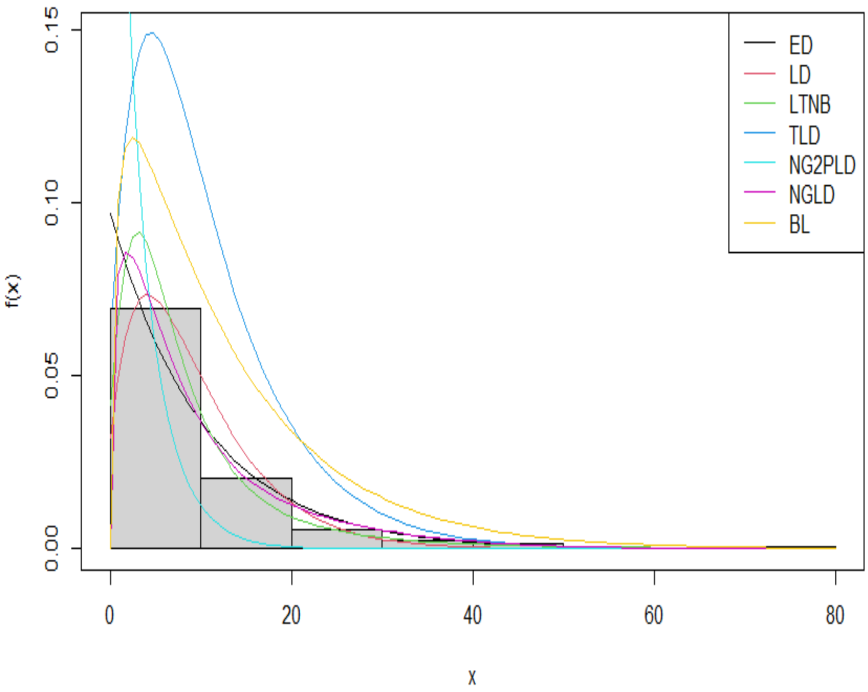

Figure 5: *Estimated pdf*

That is, $X_1 \stackrel{d}{=} LTNB(\alpha, \gamma, \theta)$.
If we assume that $X_{n-1} \stackrel{d}{=} LTNB(\alpha, \gamma, \theta)$, then by induction, we can establish that $X_n \stackrel{d}{=} LTNB(\alpha, \gamma, \theta)$. Hence the process $\{X_n\}$ is stationary with LTNB marginals.

Funding

The first author wishes to thank the University Grants Commission, New Delhi for the financial assistance in the form of Junior Research Fellowship.

References

[1] Aarset, M. V. (1987). How to identify bathtub hazard rate. *IEEE Transactions on Reliability* , 36: 106-108.
[2] Bapat S. R. and Bhardwaj R.(2021). On an inflated Unit-Lindley distribution. *Journal of Statistical Research* , 55: 299-311.
[3] Bakouch H., Al-Zahrani B., Al-Shomrani A., Marchi V. and Louzad F.(2012). An extended Lindley distribution. *Journal of the Korean Statistical Society* ,41: 75–85.
[4] Bhati D., Malik M. A. and Vaman H. J.(2015). Lindley-Exponential distribution: Properties and applications. *Metron* ,73: 335–357.
[5] Bennett S. (1983). Log-logistic regression models for survival data. *Applied Statistics*, 32: 165.
[6] Elbatal I., Merovci F. and Elgarhy M.(2013). A new generalized Lindley distribution. *Mathematical Theory and Modeling* , 3: 30-47.
[7] Ekhosuehi N., Opone F. and Odobaire F.(2018). New generalised two parameter Lindley distribution.*Journal of Data Science* ,16: 549-566.

[8] Ghitany M.E., Atieh B., and Nadarajah S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation* , 78: 493-506.

[9] 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: 1190-1201.

[10] Glanzel (1987). A characterization theorem based on truncated moments and its application to some distribution families. *Mathematical statistics and probability theory, (Bad Tatzmannsdorf, 1986)*, B: 75-84.

[11] Jayakumar K. and Pillai R.N (1993). The first-order autoregressive Mittag–Leffler process. *Journal of Applied Probability* , 30: 462-466.

[12] Jayakumar K., Bindu Krishnan, and Hamedani G.G (2020). On a new generalization of Pareto distribution and its applications. *Communications in Statistics: Computation and Simulation* , 49: 1264-1284.

[13] Kenny, J. F. and Keeping, E., Mathematics of Statistics D. Van Nostrand Company, Princeton. 1962.

[14] Korkmaz C. M., Yousof H.M., Rasekhi M. and Hamedani G.G.(2018). The Odd Lindley Burr XII Model. Bayesian Analysis, Classical Inference and Characterizations. *Journal of Data Science* , 16: 327-354.

[15] Lee E. T. and Wang J. Statistical Methods for Survival Data Analysis, New York, Wiley, 2003.

[16] Lindley D.V. (1958). Fiducial distributions and Bayes' theorem . *Journal of the Royal Statistical Society B* , 20: 102-107.

[17] Marshall A.W and Olkin I.A (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika* , 84: 641-652.

[18] Merovci, F. and Sharma, V.K.(2014). The beta Lindley distribution: Properties and Applications . *Journal of Applied Mathematics* , 2014: 1-10.

[19] Moors J. J. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society, Series D*, 37: 25–32.

[20] Nadarajah S., Bakouch H.S., and Tahmasbi R. (2011). A Generalized Lindley distribution . *Sankhya B* ,73: 331-359.

[21] Nadarajah S., Jayakumar K. and Ristic M. M (2013). A new family of lifetime models. *Journal of Statistical Computation and Simulation* , 83: 1389-1404.

[22] Rajarshi S. and Rajarshi M. B. (1988). Bathtub distributions: A review . *Communications in Statistics - Theory and Methods* , 17: 2597-2621.

[23] Sharma V. K. and Khandelwal P.(2017). On the extension of the Inverse Lindley distribution *Journal of Data Science* , 15: 205-220.

[24] Shaked M., and Shanthikumar J.G. Stochastic orders, Springer, New York, 2007.

[25] Yeh H.C., Arnold B.C. and Robertson C.A. (1988). Pareto processes *Journal of Applied probability* , 25: 291-301.