STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–28.
Published online in International Academic Press (www.IAPress.org)

Modelling and Reliability Analysis of the Two-Parameter Lindley-Binomial Distribution

Mustafa Neamat Nader [1], Sameera Abdulsalam Othman [2,*], Kurdistan M.Taher Omar [3]

[1,3] *Department of Mathematics, Faculty of Science, University of Zakho, Duhok, Kurdistan*
[2] *Department of Mathematics, College of Basic Education, University of Duhok, 42001 Duhok, Iraq*

Abstract The primary purpose of this research is to describe the two parameter Lindley Binomial (LB2) distribution, a new probability distribution applicable for the proportion data analysis, specifically in the simulation, real data, and reliability analysis setting. The shape of probability mass function and some probabilistic properties of the pro posed distribution, including generating functions are derived. The method of moment, maximum likelihood estimation and the expectation-maximization algorithm are used for parameter estimation. Goodness of fit of the proposed distribution is assessed by using it on real dataset. This research also investigates the age specific prevalence and risk pattern of Hepatitis B virus (HBV) infected within the dataset. It is compared with the binomial, beta binomial, and negative binomial distributions for its performance. The results show that the proposed distribution has some advantages over previous mod els and therefore is advantageous in analyzing proportional data. Additionally, the two parameter Lindley Binomial distribution is fit to the data to evaluate the reliability func tion, hazard rate function, inverted hazard rate function, and mean residual life (MRL) by age group. The findings demonstrate substantial difference of HBV positive between various age demographics with great public health implications.

Keywords Two-Parameter Lindley-Binomial, Reliability, EM algorithm, Reversed Hazard Rate Function

AMS 2010 subject classifications:62G05; 60E05; 62N05; 62P30

DOI: 10.19139/soic-2310-5070-2516

1. Introduction

*B*inomial distribution is a basic probability model that is used frequently as many real-world situations have binary outcomes such as success or failure. For example, it helps in predicting the occurrence of side effects on a new drug being tried in a clinical trial. However, in reality data often exhibit heterogeneity or over dispersion and thus the binomial model cannot be fully utilized. To overcome these drawbacks, compound distributions have been designed by combining the binomial model with any other probability distribution. One such option as a compound binomial distribution is the Lindley distribution, which was first proposed by Lindley in 1958 [1]. The closed form survival and hazard functions for the Lindley distribution make it a very useful tool to analyse lifetime data, especially in areas of survival and reliability investigation. Researchers have expanded the model to be more adaptable by combining it with further distributions. For example, Sankarans (1970) [2] developed the Poisson-Lindley distribution, which is shown to perform well on real world sets of data. Let's say, similarly, the Negative Binomial Lindley distribution, as suggested by Zemani and Ismile (2010) [3], is useful for the over dispersed count data with excess zeros such insurance claims and accident reports. This study focuses on the versatility of the two-parameter variation of the Lindley distribution, which is more flexible in fitting lifetime data.

*Correspondence to: Sameera A. Othman (Email: sameera.othman@uod.ac). Department of Mathematics, College of Basic Education, University of Duhok, (42001) Duhok, Iraq.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Among the expansions, the generalised Poisson-Lindley distribution (Bhati et al., 2015) [4], its use in right skewed data modelling, have been researched. Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) among other model selection criteria have been used other research on Weibull based models including transmuted Weibull distributions [5, 6]. Parameter estimation of the Poisson-Lindley distribution and Lindley based models as well have been investigated using Linear Quantile Moment (LQM) Estimation, Least Squares Estimation (LSE) and Maximum Likelihood Estimation (MLE) [7, 8]. Yet, there has been a dearth of proper focus on bounded discrete counts with overdispersion in the context of epidemiology and reliability testing. One falls back on standard models such as the negative binomial or beta-binomial; however, these may possess limitations in terms of interpretability or involve complex estimation procedures. The newly proposed Two-Parameter Lindley-Binomial (LB2) distribution seeks to overcome these drawbacks by the synthesis of the bound property of the binomial distribution with a comprehensive dispersion framework. The LB2 model has interpretable parameters and is computationally straightforward, and hence is appropriate for real-world modeling situations, including ones dealing with age-specific count data. The contrast is being made to bring out these strengths compared to alternatives.The structure of this study is as follows: Sections 2 and 3 respectively present the methodology and various estimation approaches for the Two Parameter Lindley Binomial Distribution. Section 4 presents application of these techniques and Section 5 presents conclusions summary.

2. Methodology

The theoretical foundation for the Two-Parameters Lindley-Binomial Distribution is presented.

2.1. One-Parameter Lindley Distribution

The following probability density function (pdf) of this one parameter distribution known as Lindley distribution was innovated by Lindley (1958) [1]

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

When β is the shape parameter.

The parameter β for the one-parameter Lindley distribution is a positive real number. Adjusting this shape parameter changes the distribution's general behavior. This distribution is a combination of the exponential (β) and gamma $(2,\beta)$ distributions as following formula:

$$f(x;\beta)=\pi f_1(x)+(1-\pi)f_2(x)$$

with a mixture proportion $\pi=\frac{\beta}{\beta+1}$, $f_1\left(x;\beta\right)=\beta e^{-\beta x}$ and $f_2\left(x;\beta\right)=\beta^2 x e^{-\beta x}$.

Figure (1) in the left shows how the probability mass function (PMF) for the Lindley distribution, defined in Equation (1), while the right shows cumulative distribution function (CDF) which follows Equation (2) changes with different values of β. The graphic shows how the shape of the distribution changes with different values of β.

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

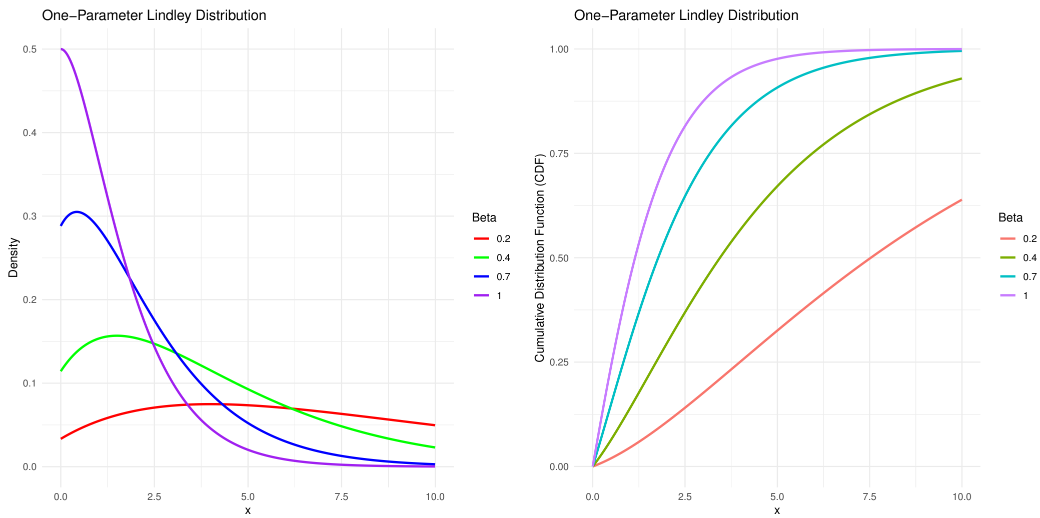

Figure 1. The (PMF) in the left and (cdf) in the right of one-parameter Lindley Distribution for various values of the parameter β.

2.2. *Two-Parameter Lindley Distribution*

The research of Shanker et al. (2013) [9] developed a two-parameter Lindley distribution (LD2) for analyzing waiting and survival time data. Figures (2) in the left shows how the probability mass function (PMF), while the right shows cumulative distribution function (CDF) of this distribution which uses different (α, β) values as shown below.

$$f(x; \alpha, \beta) = \frac{\beta^2}{\beta + \alpha}(1 + \alpha x)e^{-\beta x}; \quad x > 0,\ \beta > 0,\ \alpha > 0 \tag{3}$$

$$F(x; \alpha, \beta) = 1 - \frac{\beta + \alpha + \alpha\beta x}{\beta + \alpha}e^{-\beta x}; \quad x > 0,\ \beta > 0,\ \alpha > 0 \tag{4}$$

The two-parameter Lindley distribution presents its probability density function (pdf) as a combination of exponential (β) and gamma $(2, \beta)$ distributions which takes this form:

$$f(x; \beta) = \pi f_1(x) + (1 - \pi)f_2(x)$$

where $\pi = \frac{\beta}{\beta+\alpha}$, $f_1(x; \beta) = \beta e^{-\beta x}$ and $f_2(x; \beta) = \beta^2 x e^{-\beta x}$

2.3. *Two-parameter Lindley-Binomial distribution*

Deng and Zhang (2024) [10] introduced the Lindley-Binomial distribution, which has two parameters α and β, This distribution combines the Binomial distribution $\left(n, p = e^{-\lambda}\right)$ with the two-parameter Lindley distribution (α, β). Then, the probability mass function (pmf) of X takes the following form:

$$p(X = x) = \int_0^\infty p(x|\lambda)f(\lambda; \alpha, \beta)d\lambda$$

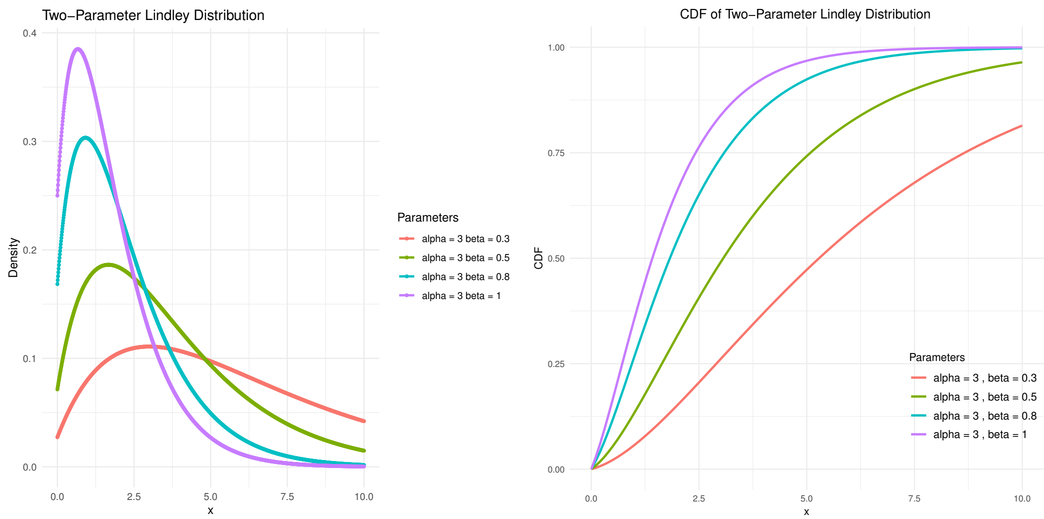

Figure 2. The (PMF) in the left and (cdf) in the right of two-parameter Lindley Distribution for various values of the parameter β.

$$= \int_{0}^{\infty} C_x^n e^{-\lambda x} \sum_{h=0}^{n-x} C_h^{n-x} (-1)^h e^{-\lambda h} \frac{\beta^2}{\beta+\alpha} (1+\alpha x) e^{-\beta\lambda} d\lambda$$

Here the PMF of the Two-parameter Lindley-Binomial distribution (LB2) is given directly in terms of the following equation, obtained by solving this integral.

$$p(X=x) = C_x^n \frac{\beta^2}{\beta+\alpha} \sum_{h=0}^{n-x} C_h^{n-x} (-1)^h \frac{\beta+x+h+\alpha}{(\beta+x+h)^2}; x=0,1,\dots n\ , \quad \beta,\alpha>0 \tag{5}$$

The distribution probability mass function is explained by this equation and its features such as mean, variance, moments, skewness and kurtosis will be examined. This section also includes a discussion of how to generate random numbers from the two parameter Lindley-Binomial distribution and parameter estimation approaches.

The probability mass function (PMF) of the LB2 distribution, defined in Equation (5), involves nested sums which can prove computationally inconvenient especially at large values of To counter this issue, truncation methodology is used for the inner sum, halting the computation when further terms make a negligible contribution to the total result. This method provides a good balance between numerical accuracy and computational speed. To enable practical usability and reprehensibility, we formulated and tested an algorithm in R capable of computing the PMF in this truncation framework.

The cumulative distribution function of (LB2) follows this equation below:

$$F(x) = \sum_{t=0}^{x} p(X=x) = \sum_{t=0}^{x} C_t^n \frac{\beta^2}{\beta+\alpha} \sum_{h=0}^{n-t} C_h^{n-t} (-1)^h \frac{\beta+t+h+\alpha}{(\beta+t+h)^2} \ ; x=0,1,\dots n, \beta,\alpha>0 \tag{6}$$

Figure (3) in the left shows how the probability mass function (PMF), while the right shows cumulative distribution function (CDF) of the (LB2) distribution given in Equations (5) and (6) for various values of α and β. Also, the diagrams show how the forms of these functions change with varying parameter values.

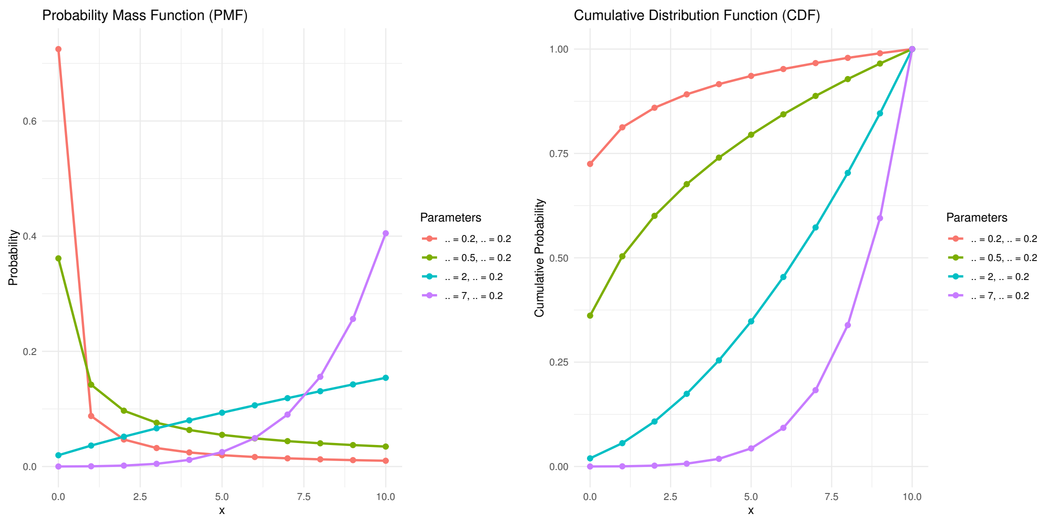

Figure 3. The (PMF) in the left and (cdf) in the rihgt of two-parameter Lindley-Binomial distribution for various β and constant α.

2.3.1 Interpretation of Parameters

The two parameters of the LB2 distribution, represented as α and β, play distinct and meaningful roles in determining the behavior of the distribution:

β: primarily influences the tail shape and behavior of the distribution. Small values of β provide more symmetric and compact probability masses, while large values increase right skewness and stretch out the peak, allowing for heavier tails. This renders β particularly suitable to model phenomena in which risk or frequency accelerates with a driving factor(e.g., age-dependent acceleration of HBV infection risk). α serves as a scale and dispersion parameter. Values of α produce sharper peaks and faster cumulative mass concentration, whereas lower values spread the distribution more uniformly.

To graphically display these effects, Figure (3) in the left shows how the probability mass function (PMF), while the right shows cumulative distribution function (CDF) of the (LB2) distribution given in Equations (5) and (6) for various values of α and β. Also, the diagrams show how the forms of these functions change with varying parameter values. These plots demonstrate how changes in parameter values influence the peak, spread, and skewness, offering helpful insights for interpretation in applications to practical problems such as disease prevalence modeling.

2.3.2 Properties of the Two-Parameter Lindley-Binomial Distribution.

To demonstrate that $\sum\limits_{x=0}^{n} p\left(X=x\right)=1$

The (PMF) of the LB2 distribution is expressed as:

$$p(X=x)=\int_{0}^{\infty} p(x|\lambda) f(\lambda;\alpha,\beta)\, d\lambda$$

$\sum\limits_{x=0}^{n} p\left(X=x\right)$ over all (x) gives:

$$\sum_{x=0}^{n} p(X=x)=\sum_{x=0}^{n}\int_{0}^{\infty} p(x|\lambda) f(\lambda;\alpha,\beta)\, d\lambda$$

Rearranging the sums and integrals:

$$= \sum_{x=0}^{n} C_x^n \left(e^{-\lambda}\right)^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} (1+\alpha\lambda) \ e^{-\beta\lambda} \, d\lambda$$

Using the Binomial theorem, the summing reduces to:

$$= \left(e^{-\lambda}+1-e^{-\lambda}\right)^n \frac{\beta^2}{(\beta+\alpha)} \left[\frac{-(1+\alpha\lambda)}{\beta} \ e^{-\beta\lambda} - \frac{\alpha}{\beta^2} \ e^{-\beta\lambda} \Big|_0^\infty \right]$$

Simplify is given:

$$= \frac{\beta^2}{(\beta+\alpha)} \left[\frac{1}{\beta} + \frac{\alpha}{\beta^2}\right]$$

$$= \frac{\beta^2}{(\beta+\alpha)} \left[\frac{\beta+\alpha}{\beta^2}\right] = 1$$

2.3.3 The two-parameter Lindley-Binomial distribution's r^{th} moment has the following formula:

$$\mu'_r = E(x^r) = \sum_{x=0}^{n} x^r p(X=x) = \sum_{x=0}^{n} x^r \int_0^\infty p(x|\lambda) f(\lambda;\alpha,\beta) \, d\lambda$$

The first four moments about origin are obtained as:

$$\begin{aligned}
\mu'_1 &= E(x) \\
&= \sum_{x=0}^{n} x p(X=x) \\
&= \sum_{x=0}^{n} x \int_0^\infty p(x|\lambda) f(\lambda;\alpha,\beta) d\lambda \\
&= \sum_{x=0}^{n} x \, C_x^n \left(e^{-\lambda}\right)^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} (1+\alpha\lambda) \ e^{-\beta\lambda} \, d\lambda \\
&= n \, e^{-\lambda} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} (1+\alpha\lambda) \, e^{-\beta\lambda} \, d\lambda \\
&= n \ \frac{\beta^2}{(\beta+\alpha)} \int_0^\infty (1+\alpha\lambda) \ e^{-\lambda(\beta+1)} d\lambda \\
&= n \frac{\beta^2}{(\beta+\alpha)} \left[\frac{-(1+\alpha\lambda)}{\beta+1} e^{-\lambda(\beta+1)} - \frac{\alpha}{(\beta+1)^2} e^{-\lambda(\beta+1)} \Big|_0^\infty \right] \\
&= n \frac{\beta^2}{(\beta+\alpha)} \left[\frac{1}{\beta+1} + \frac{\alpha}{(\beta+1)^2} \right] \\
&= n \frac{\beta^2}{(\beta+\alpha)} \left[\frac{\beta+1+\alpha}{(\beta+1)^2} \right] \\
&= \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}
\end{aligned}$$

The second moment is

$$\begin{aligned}
\mu'_2 &= E\left(x^2\right) \\
&= \sum_{x=0}^{n} x^2 p(X=x) \\
&= \sum_{x=0}^{n} x^2 \int_0^{\infty} p(x|\lambda) f(\lambda;\alpha,\beta)\, d\lambda \\
&= \sum_{x=0}^{n} x^2\, C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^{\infty} \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ \ e^{-\beta\lambda}\, d\lambda \\
&= \left[(n^2-n)e^{-2\lambda}+ne^{-\lambda}\right] \int_0^{\infty} \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ e^{-\beta\lambda}\, d\lambda \\
&= \frac{(n^2-n)\beta^2}{(\beta+\alpha)} \int_0^{\infty} (1+\alpha\lambda)\ e^{-\lambda(\beta+2)}\, d\lambda + \frac{n\beta^2}{(\beta+\alpha)} \int_0^{\infty} (1+\alpha\lambda)\ e^{-\lambda(\beta+1)}\, d\lambda \\
&= \frac{(n^2-n)\beta^2}{(\beta+\alpha)} \left[\frac{-(1+\alpha\lambda)}{\beta+2}\ e^{-\lambda(\beta+2)} - \frac{\alpha}{(\beta+2)^2}\ e^{-\lambda(\beta+2)}\Big|_0^{\infty}\right] \\
&+ \frac{n\beta^2}{(\beta+\alpha)} \left[\frac{-(1+\alpha\lambda)}{\beta+1}\ e^{-\lambda(\beta+1)} - \frac{\alpha}{(\beta+1)^2}\ e^{-\lambda(\beta+1)}\Big|_0^{\infty}\right] \\
&= \frac{(n^2-n)\beta^2}{(\beta+\alpha)} \left[\frac{1}{\beta+2} - \frac{\alpha}{(\beta+2)^2}\right] + \frac{n\beta^2}{(\beta+\alpha)} \left[\frac{1}{\beta+1} - \frac{\alpha}{(\beta+1)^2}\right] \\
&= \frac{(n^2-n)\beta^2}{(\beta+\alpha)} \left[\frac{\beta+2+\alpha}{(\beta+2)^2}\right] + \frac{n\beta^2}{(\beta+\alpha)} \left[\frac{\beta+1+\alpha}{(\beta+1)^2}\right] \\
&= \frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}
\end{aligned}$$

Finally, the third and fourth central moments are:

$$\begin{aligned}
\mu_3 &= E\left[(x-\mu'_1)^3\right] \\
&= \sum_{x=0}^{n} (x-\mu'_1)^3 p(X=x) \\
&= \sum_{x=0}^{n} (x-\mu'_1)^3 \int_0^{\infty} p(x|\lambda) f(\lambda;\alpha,\beta) d\lambda \\
&= \sum_{x=0}^{n} (x-\mu'_1)^3\, C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^{\infty} \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ \ e^{-\beta\lambda}\, d\lambda \\
&= \left[x^3 - 3\mu'_1 x^2 + 3{\mu'_1}^2 x - {\mu'_1}^3\right] \sum_{x=0}^{n} C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^{\infty} \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ \ e^{-\beta\lambda}\, d\lambda
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{x=0}^{n} x^3 C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda}\, d\lambda \\
&- 3{\mu'}_1 \sum_{x=0}^{n} x^2 C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda}\, d\lambda \\
&+ 3{\mu'}_1{}^2 \sum_{x=0}^{n} x\, C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda}\, d\lambda \\
&- {\mu'}_1{}^3 \sum_{x=0}^{m} \ C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda}\, d\lambda \\
&= \ {\mu'}_3 - 3{\mu'}_1({\mu'}_2) + 3{\mu'}_1{}^2({\mu'}_1) - {\mu'}_1{}^3(1) \\
&= \ {\mu'}_3 - 3{\mu'}_1({\mu'}_2) + 3{\mu'}_1{}^3 - {\mu'}_1{}^3 \\
&= \ {\mu'}_3 - 3{\mu'}_1({\mu'}_2) + 2{\mu'}_1{}^3 \\
&= \frac{n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)\,(\beta+3)^2} + \frac{3n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2} \\
&- 3\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]\left[\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right] \\
&+ 2\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]^3
\end{aligned}
$$

$$
\begin{aligned}
&{\mu'}_4 = E\left(x^4\right) \\
&= \sum_{x=0}^{n} x^4 p(X=x) \\
&= \sum_{x=0}^{n} x^4 \int_0^\infty p(x|\lambda) f(\lambda;\alpha,\beta) d\lambda \\
&= \sum_{x=0}^{n} x^4 \ C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda}\, d\lambda \\
&= \left[n(n-1)(n-2)(n-3)e^{-4\lambda} + 6n(n-1)(n-2)e^{-3\lambda} + 7n(n-1)e^{-2\lambda} + ne^{-\lambda}\right] \\
&\int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda}\, d\lambda \\
&= \frac{n(n-1)(n-2)(n-3)\beta^2}{(\beta+\alpha)} \int_0^\infty (1+\alpha\lambda) \ e^{-\lambda(\beta+4)}\, d\lambda
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{6n(n-1)(n-2)\beta^2}{(\beta+\alpha)}\int_0^\infty (1+\alpha\lambda)\ e^{-\lambda(\beta+3)}\ d\lambda \\
&+\frac{7n(n-1)\beta^2}{(\beta+\alpha)}\int_0^\infty (1+\alpha\lambda)\ e^{-\lambda(\beta+2)}\ d\lambda+\frac{n\beta^2}{(\beta+\alpha)}\int_0^\infty (1+\alpha\lambda)\ e^{-\lambda(\beta+1)}\ d\lambda \\
&=\frac{n(n-1)(n-2)(n-3)\beta^2}{(\beta+\alpha)}\left[\frac{-(1+\alpha\lambda)}{\beta+4}\ e^{-\lambda(\beta+4)}-\frac{\alpha}{(\beta+4)^2}\ e^{-\lambda(\beta+4)}\Big|_0^\infty\right] \\
&+\frac{6n(n-1)(n-2)\beta^2}{(\beta+\alpha)}\left[\frac{-(1+\alpha\lambda)}{\beta+3}\ e^{-\lambda(\beta+3)}-\frac{\alpha}{(\beta+3)^2}\ e^{-\lambda(\beta+3)}\Big|_0^\infty\right] \\
&+\frac{7n(n-1)\beta^2}{(\beta+\alpha)}\left[\frac{-(1+\alpha\lambda)}{\beta+2}\ e^{-\lambda(\beta+2)}-\frac{\alpha}{(\beta+2)^2}\ e^{-\lambda(\beta+2)}\Big|_0^\infty\right] \\
&+\frac{n\beta^2}{(\beta+\alpha)}\left[\frac{-(1+\alpha\lambda)}{\beta+1}\ e^{-\lambda(\beta+1)}-\frac{\alpha}{(\beta+1)^2}\ e^{-\lambda(\beta+1)}\Big|_0^\infty\right] \\
&=\frac{n(n-1)(n-2)(n-3)\beta^2}{(\beta+\alpha)}\left[\frac{1}{\beta+4}-\frac{\alpha}{(\beta+4)^2}\right]+\frac{6n(n-1)(n-2)\beta^2}{(\beta+\alpha)}\left[\frac{1}{\beta+3}-\frac{\alpha}{(\beta+3)^2}\right] \\
&+\frac{7n(n-1)\beta^2}{(\beta+\alpha)}\left[\frac{1}{\beta+2}-\frac{\alpha}{(\beta+2)^2}\right]+\frac{n\beta^2}{(\beta+\alpha)}\left[\frac{1}{\beta+1}-\frac{\alpha}{(\beta+1)^2}\right] \\
&=\frac{n(n-1)(n-2)(n-3)\beta^2}{(\beta+\alpha)}\left[\frac{\beta+4+\alpha}{(\beta+4)^2}\right]+\frac{6n(n-1)(n-2)\beta^2}{(\beta+\alpha)}\left[\frac{\beta+3+\alpha}{(\beta+3)^2}\right] \\
&+\frac{7n(n-1)\beta^2}{(\beta+\alpha)}\left[\frac{\beta+2+\alpha}{(\beta+2)^2}\right]+\frac{n\beta^2}{(\beta+\alpha)}\left[\frac{\beta+1+\alpha}{(\beta+1)^2}\right]
\end{aligned}
$$

.

$$
\begin{aligned}
&=\frac{n(n-1)(n-2)(n-3)\beta^2(\beta+4+\alpha)}{(\beta+\alpha)(\beta+4)^2}+\frac{6n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)(\beta+3)^2} \\
&+\frac{7n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2}+\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}
\end{aligned}
$$

The variance (σ^2) is obtained as:

$$
\begin{aligned}
Var(x)=\sigma^2&={\mu'}_2-{\mu'}_1^{\,2} \\
&=E(x^2)-[E(x)]^2 \\
&=\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2}+\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}-\left(\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\right)^2 \\
&=\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2}+\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\left[1-\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\right]
\end{aligned}
$$

2.3.4 The two-parameter Lindley-Binomial distribution's r^{th} moment about the mean appears in the following formula:

$$
\mu_r=E\left[(x-{\mu'}_1)^r\right]=\sum_{x=0}^{n}(x-{\mu'}_1)^r p(X=x)=\sum_{x=0}^{n}(x-{\mu'}_1)^r\int_0^\infty p(x|\lambda)f(\lambda;\alpha,\beta)\,d\lambda
$$

Therefore, the second, third and fourth moments about the mean are obtained as:

$$\begin{aligned}
\mu_2 &= E\left[(x-{\mu'}_1)^2\right] \\
&= \sum_{x=0}^{n}(x-{\mu'}_1)^2 p(X=x) \\
&= \sum_{x=0}^{n}(x-{\mu'}_1)^2 \int_0^\infty p(x|\lambda)f(\lambda;\alpha,\beta)d\lambda \\
&= \sum_{x=0}^{n}(x-{\mu'}_1)^2\, C_x^n(e^{-\lambda})^x\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\;(1+\alpha\lambda)\;\; e^{-\beta\lambda}\,d\lambda \\
&= \left[x^2-2{\mu'}_1x+{\mu'}_1{}^2\right]\sum_{x=0}^{n} C_x^n(e^{-\lambda})^x\,\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\;(1+\alpha\lambda)\;\; e^{-\beta\lambda}\,d\lambda \\
&= \sum_{x=0}^{n} x^2\, C_x^n(e^{-\lambda})^x\,\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\;(1+\alpha\lambda)\;\; e^{-\beta\lambda}\,d\lambda \\
&- 2{\mu'}_1\sum_{x=0}^{n} x\, C_x^n(e^{-\lambda})^x\,\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\;(1+\alpha\lambda)\;\; e^{-\beta\lambda}\,d\lambda \\
&+ {\mu'}_1{}^2\sum_{x=0}^{n} C_x^n(e^{-\lambda})^x\,\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\;(1+\alpha\lambda)\;\; e^{-\beta\lambda}\,d\lambda \\
&= \;{\mu'}_2-2{\mu'}_1({\mu'}_1)\;+{\mu'}_1{}^2(1) \\
&= {\mu'}_2-2{\mu'}_1{}^2\;+{\mu'}_1{}^2 \\
&= {\mu'}_2-{\mu'}_1{}^2 \\
&= Var(x) \\
&= \frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2}+\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\left[1-\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]
\end{aligned}$$

$$\begin{aligned}
\mu_3 &= E\left[(x-{\mu'}_1)^3\right] \\
&= \sum_{x=0}^{n}(x-{\mu'}_1)^3 p(X=x) \\
&= \sum_{x=0}^{n}(x-{\mu'}_1)^3 \int_0^\infty p(x|\lambda)f(\lambda;\alpha,\beta)d\lambda \\
&= \sum_{x=0}^{n}(x-{\mu'}_1)^3\, C_x^n(e^{-\lambda})^x\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\;(1+\alpha\lambda)\;\; e^{-\beta\lambda}\,d\lambda \\
&= \left[x^3-3{\mu'}_1x^2+3{\mu'}_1{}^2x-{\mu'}_1{}^3\right]\sum_{x=0}^{n} C_x^n(e^{-\lambda})^x\,\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\;(1+\alpha\lambda)\;\; e^{-\beta\lambda}\,d\lambda
\end{aligned}$$

$$= \sum_{x=0}^{n} x^3 C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda} \, d\lambda$$
$$- 3\mu'_1 \sum_{x=0}^{n} x^2 C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda} \, d\lambda$$
$$+ 3{\mu'_1}^2 \sum_{x=0}^{n} x \, C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda} \, d\lambda$$
$$- {\mu'_1}^3 \sum_{x=0}^{m} C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda} \, d\lambda$$
$$= \ \mu'_3 - 3\mu'_1(\mu'_2) + 3{\mu'_1}^2(\mu'_1) - {\mu'_1}^3(1)$$
$$= \ \mu'_3 - 3\mu'_1(\mu'_2) + 3{\mu'_1}^3 - {\mu'_1}^3$$
$$= \ \mu'_3 - 3\mu'_1(\mu'_2) + 2{\mu'_1}^3$$
$$= \frac{n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)\,(\beta+3)^2} + \frac{3n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}$$
$$- 3\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]\left[\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]$$
$$+ 2\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]^3$$

$$\mu_4 = E\left[(x-\mu'_1)^4\right]$$
$$= \sum_{x=0}^{n} (x-\mu'_1)^4 p(X=x)$$
$$= \sum_{x=0}^{n} (x-\mu'_1)^4 \int_0^\infty p(x|\lambda) f(\lambda;\alpha,\beta) d\lambda$$
$$= \sum_{x=0}^{n} (x-\mu'_1)^4 \, C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda} \, d\lambda$$
$$= \left[x^4 - 4\mu'_1 x^3 + 6{\mu'_1}^2 x^2 - 4{\mu'_1}^3 x + {\mu'_1}^4\right]$$
$$\sum_{x=0}^{n} C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda} \, d\lambda$$
$$= \sum_{x=0}^{n} x^4 \, C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ \ e^{-\beta\lambda} \, d\lambda$$

$$- 4{\mu'}_1\sum_{x=0}^{n} x^3\ C_x^n(e^{-\lambda})^x\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ \ e^{-\beta\lambda}\, d\lambda$$

$$+ 6{{\mu'}_1}^2\sum_{x=0}^{n} x^2\ C_x^n(e^{-\lambda})^x\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ \ e^{-\beta\lambda}\, d\lambda$$

$$- 4{{\mu'}_1}^3\sum_{x=0}^{n} x\ C_x^n(e^{-\lambda})^x\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ \ e^{-\beta\lambda}\, d\lambda$$

$$+ {{\mu'}_1}^4\sum_{x=0}^{n}\ C_x^n(e^{-\lambda})^x\left(1-e^{-\lambda}\right)^{n-x}\int_0^\infty \frac{\beta^2}{(\beta+\alpha)}\ (1+\alpha\lambda)\ \ e^{-\beta\lambda}\, d\lambda$$

$$= \ {\mu'}_4 - 4{\mu'}_1({\mu'}_3) + 6{{\mu'}_1}^2({\mu'}_2) - 4{{\mu'}_1}^3({\mu'}_1) + {{\mu'}_1}^4(1)$$

$$= \ {\mu'}_4 - 4{\mu'}_1({\mu'}_3) + 6{{\mu'}_1}^2({\mu'}_2) - 4{{\mu'}_1}^4 + {{\mu'}_1}^4$$

$$= \ {\mu'}_4 - 4{\mu'}_1({\mu'}_3) + 6{{\mu'}_1}^2({\mu'}_2) - 3{{\mu'}_1}^4$$

$$= \frac{n(n-1)(n-2)(n-3)\beta^2(\beta+4+\alpha)}{(\beta+\alpha)\,(\beta+4)^2} + \frac{6n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)\,(\beta+3)^2}$$
$$+ \frac{7n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2} - 4\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]$$
$$\left[\frac{n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)\,(\beta+3)^2} + \frac{3n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]$$
$$+ 6\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]^2\left[\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right] - 3\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]^4$$

(C.K) and (C.S) represent the coefficients of kurtosis and skewness for the two-parameter Lindley-Binomial distribution that equal:

$$C.K = \frac{\mu_4}{(\mu_2)^2}$$
$$= \frac{E\left[(x-{\mu'}_1)^4\right]}{(\mathrm{var}(x))^2}$$

$$= \frac{\begin{gathered}\frac{n(n-1)(n-2)(n-3)\beta^2(\beta+4+\alpha)}{(\beta+\alpha)\,(\beta+4)^2} + \frac{6n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)\,(\beta+3)^2} \\ + \frac{7n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2} - 4\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right] \\ \left[\frac{n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)\,(\beta+3)^2} + \frac{3n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right] \\ + 6\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]^2\left[\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right] - 3\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}\right]^4\end{gathered}}{\left[\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\left[1-\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\right]\right]^2}$$

And

$$
\begin{aligned}
C.S &= \frac{\mu_3}{\sqrt{(\mu_2)^3}} \\
&= \frac{E\left[(x-\mu'_1)^3\right]}{(\mathrm{var}(x))^{3/2}} \\
&= \frac{\begin{array}{l} \frac{n(n-1)(n-2)\beta^2(\beta+3+\alpha)}{(\beta+\alpha)(\beta+3)^2} + \frac{3n(n-1)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2} \\ -3\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\right]\left[\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\right] \\ +2\left[\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\right] \end{array}}{\left[\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\left[1-\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}\right]\right]^{3/2}}
\end{aligned}
$$

2.3.5 The moment-generating function for the two-parameter Lindley-Binomial distribution as shown below:

$$
\begin{aligned}
M_X(t) = E(e^{tx}) &= \sum_{x=0}^{n} e^{tx} p(X=x) \\
&= \sum_{x=0}^{n} e^{tx} \int_0^\infty p(x|\lambda) f(\lambda;\alpha,\beta) d\lambda \\
&= \sum_{x=0}^{n} e^{tx} C_x^n (e^{-\lambda})^x \left(1-e^{-\lambda}\right)^{n-x} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ e^{-\beta\lambda} \ d\lambda \\
&= (1-e^{-\lambda}+e^{-\lambda}e^t)^n \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ e^{-\beta\lambda} \ d\lambda \\
&= \left[1+e^{-\lambda}(e^t-1)\right]^n \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ e^{-\beta\lambda} \ d\lambda \\
&= \sum_{u=0}^{n} C_u^n (e^t-1)^u e^{-\lambda u} \int_0^\infty \frac{\beta^2}{(\beta+\alpha)} \ (1+\alpha\lambda) \ e^{-\beta\lambda} \ d\lambda \\
&= \sum_{u=0}^{n} C_u^n (e^t-1)^u \frac{\beta^2}{(\beta+\alpha)} \int_0^\infty (1+\alpha\lambda) \ e^{-\lambda(\beta+u)} \ d\lambda \\
&= \sum_{u=0}^{n} C_u^n (e^t-1)^u \frac{\beta^2}{(\beta+\alpha)} \left[\frac{-(1+\alpha\lambda)}{\beta+u} \ e^{-\lambda(\beta+u)} - \frac{\alpha}{(\beta+u)^2} \ e^{-\lambda(\beta+u)}\Big|_0^\infty\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{u=0}^{n} C_u^n (e^t - 1)^u \frac{\beta^2}{(\beta+\alpha)} \left[\frac{1}{\beta+u} + \frac{\alpha}{(\beta+u)^2} \right] \\
&= \sum_{u=0}^{n} C_u^n (e^t - 1)^u \frac{\beta^2}{(\beta+\alpha)} \left[\frac{\beta+u+\alpha}{(\beta+u)^2} \right] \\
&= \sum_{u=0}^{n} C_u^n (e^t - 1)^u \frac{\beta^2(\beta+u+\alpha)}{(\beta+\alpha)\,(\beta+u)^2}
\end{aligned}
$$

2.3.6 Dispersion Index

The dispersion index serves as a valuable statistical measure which shows how clustered or how dispersed the observation set is compared to standard models. The dispersion index for the LB2 distribution provides this result

$$
\begin{aligned}
d &= \frac{Var(x)}{\mu'_1} \\
&= \frac{\mu'_\mu 2 - {\mu'_1}^2}{\mu'_1} \\
&= \frac{E(x^2) - [E(x)]^2}{E(x)} \\
&= \frac{E(x^2)}{E(x)} - E(x) \\
&= \frac{\frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}}{\frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)(\beta+1)^2}} - \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2} \\
&= \frac{(n-1)(\beta+2+\alpha)(\beta+1)^2}{(\beta+1+\alpha)(\beta+2)^2} + 1 - \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2} \qquad (7)
\end{aligned}
$$

Equation (7) cannot be used directly to determine if the LB2 can suitably use under-dispersed data or over-dispersed data or both. The LB2 distribution serves as a fitting method for datasets showing over or under dispersion characteristics which produces plots showing dispersion values based on α and β parameters in Figure (4). The LB2 distribution's dispersion index changes above and below one when different distribution parameters are selected according to Figure (4). The dispersion value decreases towards zero as β increases according to the plot in Figure (5). The plots demonstrate that the LB2 distribution fits data with over-dispersion and under-dispersion when different α and β values are selected.

3. Various Estimation Techniques

Numerous estimating methods within the classical paradigm are documented in the statistical literature. However, we will only provide two of these techniques here: moment and maximum likelihood.

3.1. Method of Moments

A random sample consisting of $(x_1, x_2, \ldots, x_n)$ is drawn from the two-parameter Lindley-Binomial distribution with parameters (α, β). The method of moment estimates $\hat{\alpha}$ and $\hat{\beta}$ for parameters α and β are determined by solving these two equations.

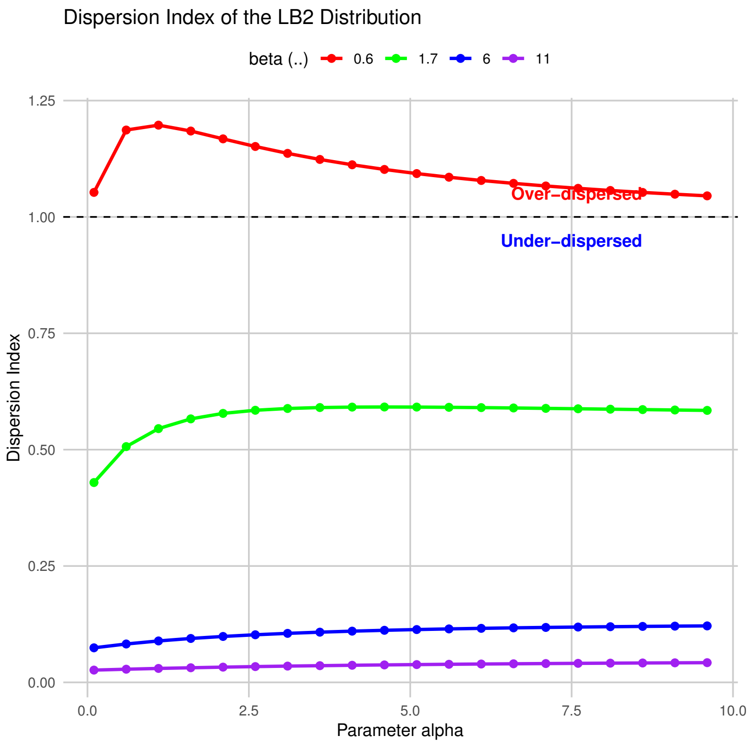

Figure 4. The dispersion of two-parameter Lindley Distribution for various α and constant β.

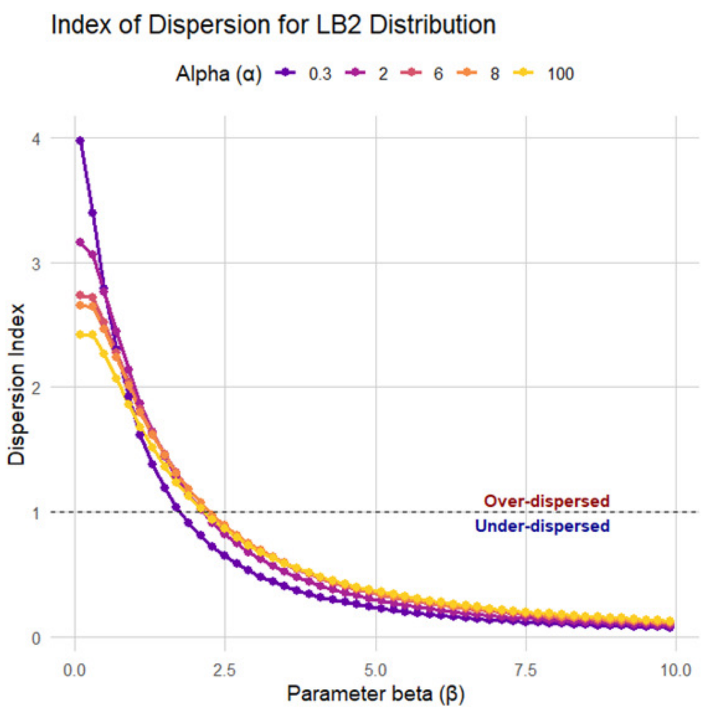

Figure 5. The dispersion of two-parameter Lindley Distribution for various β and constant α.

$$m_1 = \frac{\sum_{i=1}^{k} x_i}{k} = \mu'_1 = E(x) = \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}$$

$$m_2 = \frac{\sum_{i=1}^{k} {x_i}^2}{k} = \mu'_2 = E(x^2) = \frac{(n^2-n)\beta^2(\beta+2+\alpha)}{(\beta+\alpha)\,(\beta+2)^2} + \frac{n\beta^2(\beta+1+\alpha)}{(\beta+\alpha)\,(\beta+1)^2}$$

But there is no easy way to get the formulas of $\hat{\alpha}$ and $\hat{\beta}$ by hand. R package 'nleqslv' can be used to solve these two equations numerically.

3.2. Maximum Likelihood Estimation

Random observations $(x_1,\, x_2,\, \ldots,\, x_n)$ follow the two-parameter Lindley-Binomial distribution with parameters (α, β) when collected from a sample of size n. The likelihood function appears as follows:

$$\begin{aligned}
L\,(\alpha,\beta) &= \prod_{i=1}^{n} p(X_i = x_i) \\
&= \prod_{i=1}^{n} \left\{ C_{x_i}^{n} \frac{\beta^2}{(\beta+\alpha)} \sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2} \right\} \\
\ln\left[L\,(\alpha,\beta)\right] &= \ln\left\{ \prod_{i=1}^{n} \left[C_{x_i}^{n} \frac{\beta^2}{(\beta+\alpha)} \sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2} \right] \right\} \\
&= \sum_{i=1}^{n} \ln \left[C_{x_i}^{n} \frac{\beta^2}{(\beta+\alpha)} \sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2} \right] \\
&= \sum_{i=1}^{n} \left\{ \ln C_{x_i}^{n} + \ln\left[\frac{\beta^2}{(\beta+\alpha)}\right] + \ln\left[\sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2} \right] \right\}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial L}{\partial \alpha} &= \sum_{i=1}^{n} \left[\frac{\frac{-\beta^2}{(\beta+\alpha)^2}}{\frac{\beta^2}{(\beta+\alpha)}} + \frac{\sum_{j=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{1}{(\beta+x_i+h)^2}}{\sum_{h=0}^{n-x_i} C_h^{n-xi} (-1)^h \frac{\theta+x_i+h+\alpha}{(\theta+x_i+h)^2}} \right] \\
&= \sum_{i=1}^{n} \left[\frac{-1}{\beta+\alpha} + \frac{\sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{1}{(\beta+x_i+h)^2}}{\sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2}} \right] = 0 \qquad (8)
\end{aligned}$$

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^{n} \left[\frac{\frac{(\beta+\alpha)(2\beta)-\beta^2(1)}{(\beta+\alpha)^2}}{\frac{\beta^2}{(\beta+\alpha)}} + \frac{\sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \left(\frac{(\beta+x_i+h)^2(1)-(\beta+x_i+h+\alpha)(2(\beta+x_i+h)}{(\beta+x_i+h)^4} \right)}{\sum_{h=0}^{n-x_i} C_h^{n-x_i} (-1)^h \frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2}} \right]$$

$$= \sum_{i=1}^{n}\left[\frac{\frac{\beta(\beta+2\alpha)}{(\beta+\alpha)^2}}{\frac{\beta^2}{(\beta+\alpha)}} + \frac{\sum_{h=0}^{n-x_i} C_h^{n-x_i}(-1)^h\left(\frac{(\beta+x_i+h)[(\beta+x_i+h)-2(\beta+x_i+h+\alpha)]}{(\beta+x_i+h)^4}\right)}{\sum_{h=0}^{n-x_i} C_h^{n-x_i}(-1)^h\frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2}}\right]$$

$$= \sum_{i=1}^{n}\left[\frac{\beta+2\alpha}{\beta(\beta+\alpha)} + \frac{\sum_{h=0}^{n-x_i} C_h^{n-x_i}(-1)^h\left(\frac{-(\beta+x_i+h+2\alpha)}{(\beta+x_i+h)^3}\right)}{\sum_{h=0}^{n-x_i} C_h^{n-x_i}(-1)^h\frac{\beta+x_i+h+\alpha}{(\beta+x_i+h)^2}}\right] = 0 \quad (9)$$

Equations (8) and (9) can be solved using ‘nleqslv’ R package.

3.3. Expectation Maximization (EM) algorithm

The **EM** algorithm for the LB2 distribution to estimate the parameters (α, β) consists of the following two-step

1) Expectation (E-Step) is

a. Input $X_i = [X_1, X_2, \cdots X_n]$, the observation data $X_i\left(n,\ e^{\lambda_i}\right)$
b. fixe α^0, β^0 Initial parameter estimates
c. Fixe the convergence condition $\in = 10^{-6}$
d. Determine the number of iterations: t
e. Define the latent variable λ_i
f. Calculate the $E(\lambda_i|X_i\alpha^t, \beta^t)$, using Lindley distribution prior

$$E[\lambda_i|X_i, \alpha^t, \beta^t] = \sum_{h=0}^{m-X_i}\binom{m-X_i}{h}(-1)^h\frac{\alpha}{1+(X_i+h)\beta}$$
$$E[W_i|X_i, \alpha^t, \beta^t] = \sum_{h=0}^{m-X_i}\binom{m-X_i}{h}(-1)^h\frac{2\beta+(X_i+h)\alpha\beta^2}{(1+(X_i+h)\beta)^3}$$

2) Maximization (M-Step) is
i. Update α and β using:

$$\alpha^{t+1} = \frac{n}{\sum_{i=1}^{n}\frac{1}{X_i+\alpha^t+\ \beta^t}}$$
$$\beta^{t+1} = \frac{\sum_{i=1}^{n}(X_i+\alpha^t)}{n}$$

ii. find $\Delta_\alpha = \left|\alpha^{t+1}-\alpha^t\right| < \in$, and $\Delta_\beta = \left|\beta^{t+1}-\beta^t\right| < \in$
iii. If max $(\Delta_\alpha,\ \Delta_\beta) < \in$ stop and return $\hat{\alpha} = \alpha^{t+1}$, $\hat{\beta} = \beta^{t+1}$, Otherwise, go to apply step d to ii until get converge and replace $t = t+1$

3.4. Analysis of Reliability

The Survival Function (also known as the complementary cumulative distribution function) for the $LB2$ distribution is expressed as follows: The Survival Function $(S(x))$ indicates the probability that a random variable

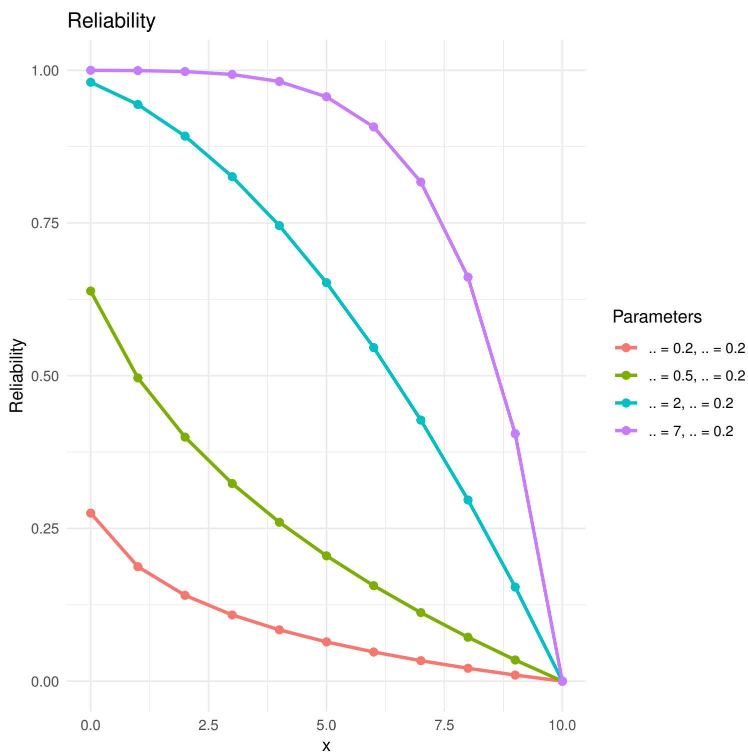

Figure 6. The survival analysis (Reliability Function) of the LB2 various values of the parameters β and α.

X from the LB2 distribution will exceed a certain value x. Figure (6) shows the reliability. Function for different (α, β) values.

$$S(x) \; = \; 1 - F(x) \tag{10}$$

Here, $F(x)$ is the cumulative distribution function as given in Equation (6).

3.4.1 Hazard Rate Function

The Hazard Rate Function for the $LB2$ distribution, represented as $h(x)$, is defined as the instantaneous rate of failure at time x. It is calculated as the ratio of the probability density function to the survival function:

$$h(x) \; = \frac{p(X = x)}{S(x)}$$

In this equation, $p(X = x)$ is the probability density function as provided in Equation (5) and $S(x)$ In Equation (10). Now, let's formulate the equations for the Survival Function and Hazard Function of the $LB2$ distribution: Figure (7) illustrates the Hazard Rate Function of the $LB2$ distribution with specified values for the parameters β, and α.

3.4.2 Reversed Hazard Rate Function

The reversed hazard rate function $(RHRF)$ of the $LB2$ distribution, denoted as $\lambda^*(x)$, can be expressed as the reciprocal of the distribution function's derivative with respect to x. Mathematically, it is defined as [reference]:

$$\lambda^*(x) = \frac{P(X = x)}{F(x)}$$

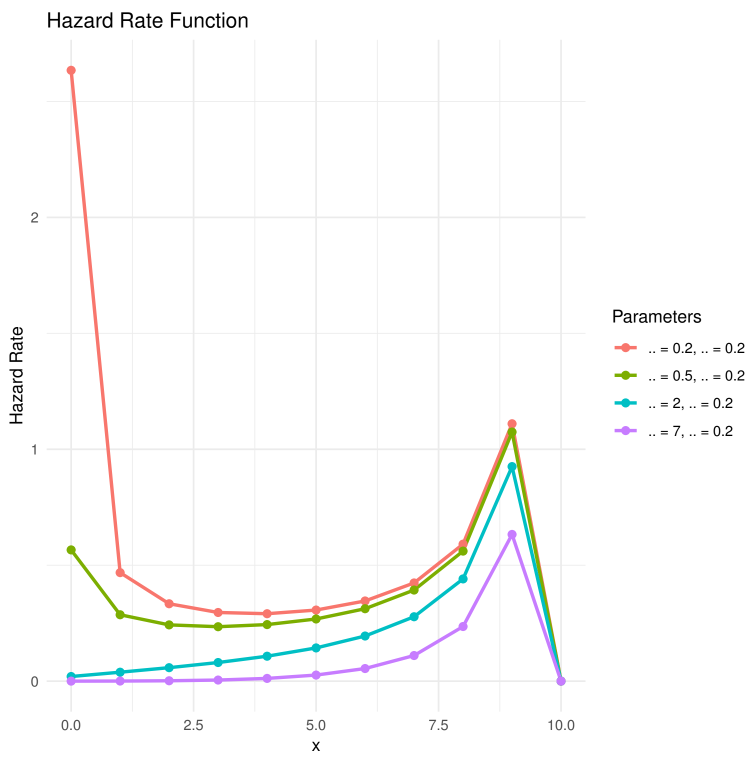

Figure 7. The Hazard Rate Function of the LB2 Distribution.

where $P(X = x)$ is the probability density function (PDF) and $F(x)$ is the distribution function of the $LB2$ distribution provided in equations (5) and (6). Therefore, substituting the expressions from equations (5) and (6) into the reversed hazard rate function. The Reversed Hazard Rate Function, illustrated in Figure (8), varies based on the values of the parameters (α, β).

3.4.3 Mean Residual Life (MRL)

The Mean Residual Life (MRL) for a lifetime random variable x in the context of survival analysis and reliability theory is defined as:

$$\text{MRL}\,(t) = E(X - t \mid X > t)$$

$$MRL(t, \alpha, \beta) = \frac{1}{S(t)} \left[\sum_{j \mid t_j \geq t} t_j f(t_j) \right] - t \qquad \text{where } t > 0$$

Where:

- t is a specific time point for which you want to calculate the MRL.
- (α, β) are the parameter of the $LB2$ distribution. $R(t)$ is the survival function, which represents the probability that the random variable x is greater than or equal to t with a given α, β.
- Time to failure is represented by x.

The MRL allows one to derive valuable information about the expected remaining lifetime of a system or process at a particular time, taking into account its history of survival. For $LB2$ system models, it is a fundamental concept

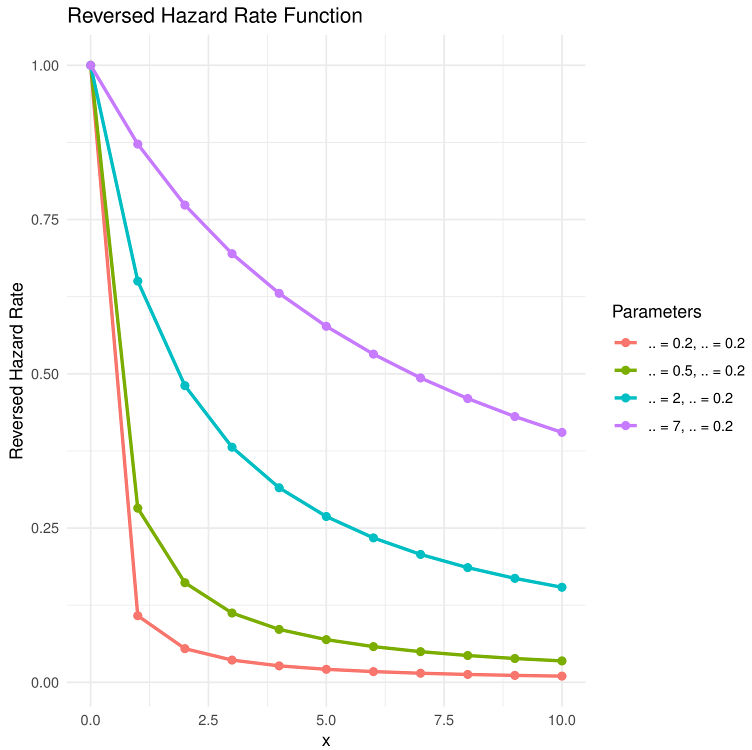

Figure 8. The Reversed Hazard Rate Function for the $LB2$ Distribution varies with the parameters (α, β).

in reliability engineering for determining the longevity and performance. Figure (9) shows the Mean Residual Life (MRL), which depends on values of the parameters (α, β).

4. Application and Analysis

4.1. Simulation study

In this section, it presents an experimental technique to assess when the distribution model $(LB2)$ is correct and reliable to explain the phenomenon of interest. In particular, simulation is often used to produce exactly and reliably backed up results in carefully conceived experiments which are necessary for progress in science. This work evaluates the performance of the $LB2$ distribution model using a simulation-based method that is both rational and practical. Simulation experiment setting: Simulation experiment is run with six different sample sizes $(10,\ 25,\ 50,\ 100,\ 150,\ and\, 500)$. We use two stages to construct two parameter $LB2$. First, produce λ_i, $i\ = 1, 2 \cdots, n$ from the L_2 distribution. The rslindley function from (LindleyR(v1.1.0)) can be used. The word slindley is a term referring to Shanker's L_2 distribution [9]. Second To generate $X_i\ , i = 1, 2 \cdots, n$, where $X_i \sim Bin\left(n, e^{-\lambda_i}\right)$. Within the following ranges, the parameters were systematically adjusted: $\alpha =\ [0.3, 0.6]$ and $\beta =\ [0.3, 0.6, 1]$. The ranges were chosen to represent several distribution shapes and scales including skewness and kurtosis that are commonly seen in environmental and reliability investigations. Simulation Procedure: To establish statistical robustness, each simulation scenario was run 1000 times with different sample sizes and parameter sets. The simulation was run on [R version 4.3.3] with ggplot2(v3.5.1) and dplyr(v1.2.1), which includes tools for random number generation, parameter estimation, and goodness-of-fit testing. The performance of moments, maximum likelihood estimators $(MLEs)$, and Em method estimators was assessed by computing mean squared error (MSE) for each scenario. Lower (MSE) values suggest improved estimator performance. The findings,

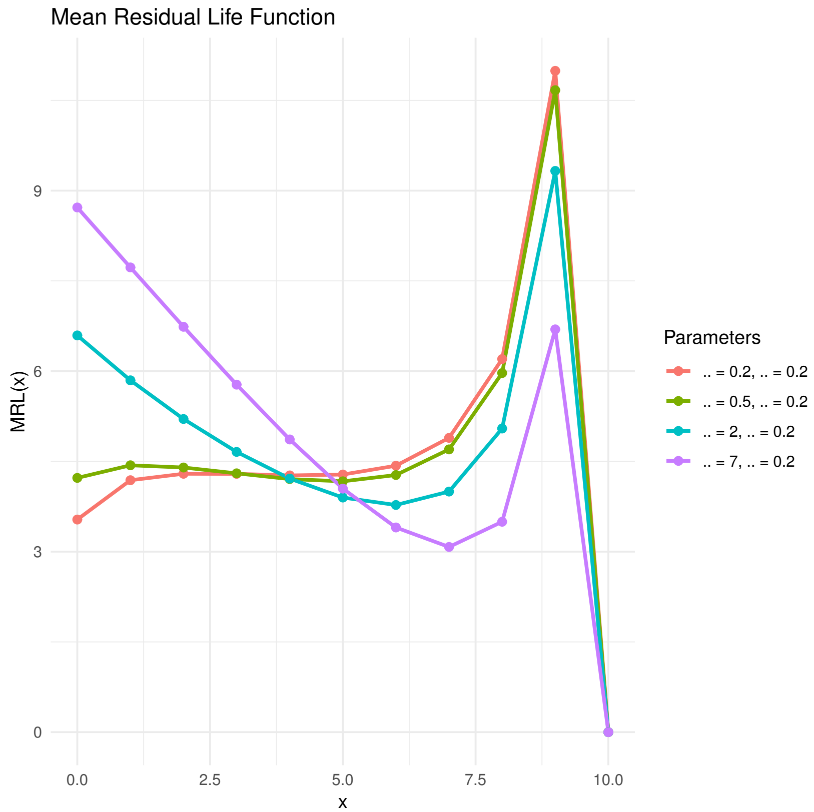

Figure 9. Mean Residual Life (MRL) for the $LB2$ distribution with (α, β) are varying parameters.

summarised in Tables (1),(2), and (3). In the Table (1) presents the results for Method of Moments estimators. For example, Table (1) demonstrates that with an increase in sample size, both bias and mean squared error diminish, signifying enhanced parameter estimation. Minor samples demonstrate more biases, especially for β when $\beta = 1$. Estimates stabilize for large sample size when greater than or equal to 100, demonstrating minimal bias and low mean squared error, so validating the consistency of the estimation approach. Hence, the Method of Moments performs well for estimating parameters. And then the following formula calculates the mean square errors across every possible set of parameter combinations

$$MSE = \frac{1}{k} \sum_{i=1}^{k} \left(\hat{\alpha}_i - \alpha\right)^2$$

And

$$Bias\left(\alpha\right) = E\left(\hat{\alpha}\right) - \alpha$$

This means that the elements $\hat{\alpha}_i$ represent estimators whereas α denotes the actual parameter values.

Table (2) presents the results for Maximum Likelihood Estimation. For instance, some parameter estimates in these results match closely with the actual values of the underlying parameters. The analyzed mean square error (MSE) values demonstrate proximity to zero. And, as well as the bias of the estimators is near to zero. So, maximum likelihood estimation is fairly good for parameter estimation. In Table (3) the show the findings of EM

Algorithm estimation. The findings indicate that if the sample size n will be small or large is not important, so the

Table 1. Results of the moments method.

Sample Size (n)	α (true)	β (true)	Estimate(α) (Bias(Estimate(α)))	Estimate(β) (Bias(Estimate(β)))	MSE (Estimate (α))	MSE (Estimate (β))
	0.3	0.3	0.1431 (-0.1569)	0.3083 (0.0083)	0.4139	0.0228
		0.6	0.2414 (-0.0586)	0.6172 (0.0172)	0.2738	0.3187
10		1	0.2101 (-0.0899)	0.5074 (-0.4926)	0.3881	6.2313
	0.6	0.3	0.2533 (-0.3467)	0.2658 (-0.0342)	0.5962	0.0195
		0.6	0.4213 (-0.1787)	0.5689 (-0.0311)	0.3249	0.0822
		1	0.4223 (-0.1777)	0.7717 (-0.2283)	0.3254	1.7067
	0.3	0.3	0.2503 (-0.0497)	0.3147 (0.0147)	0.1991	0.0123
		0.6	0.2975 (-0.0025)	0.6446 (0.0446)	0.1505	0.0511
25		1	0.2632 (-0.0368)	0.6365 (-0.3635)	0.2889	4.8793
	0.6	0.3	0.4739 (-0.1261)	0.2924 (-0.0076)	0.2215	0.0119
		0.6	0.5414 (-0.0586)	0.6036 (0.0036)	0.1720	0.0463
		1	0.5587 (-0.0413)	0.9865 (-0.0135)	0.1667	0.2528
	0.3	0.3	0.2913 (-0.0087)	0.3142 (0.0142)	0.1138	0.0077
		0.6	0.3254 (0.0254)	0.6437 (0.0437)	0.0996	0.0362
50		1	0.3320 (0.0320)	1.0013 (0.0013)	0.1384	1.1222
	0.6	0.3	0.5417 (-0.0583)	0.3019 (0.0019)	0.1321	0.0074
		0.6	0.5779 (-0.0221)	0.6145 (0.0145)	0.1125	0.0301
		1	0.5918 (-0.0082)	1.0282 (0.0282)	0.1162	0.0955
	0.3	0.3	0.3119 (0.0119)	0.3115 (0.0115)	0.0646	0.0048
		0.6	0.3287 (0.0287)	0.6293 (0.0293)	0.0550	0.0194
100		1	0.3339 (0.0339)	1.0405 (0.0405)	0.0673	0.3402
	0.6	0.3	0.5906 (-0.0094)	0.3093 (0.0093)	0.0818	0.0056
		0.6	0.6155 (0.0155)	0.6300 (0.0300)	0.0767	0.0236
		1	0.6298 (0.0298)	1.0649 (0.0649)	0.0825	0.0805
	0.3	0.3	0.3085 (0.0085)	0.3076 (0.0076)	0.0428	0.0032
		0.6	0.3160 (0.0160)	0.6175 (0.0175)	0.0350	0.0119
150		1	0.3212 (0.0212)	1.0440 (0.0440)	0.0385	0.0453
	0.6	0.3	0.6096 (0.0096)	0.3118 (0.0118)	0.0648	0.0046
		0.6	0.6233 (0.0233)	0.6300 (0.0300)	0.0609	0.0197
		1	0.6299 (0.0299)	1.0581 (0.0581)	0.0668	0.0677
	0.3	0.3	0.3048 (0.0048)	0.3026 (0.0026)	0.0104	0.0006
		0.6	0.3070 (0.0070)	0.6053 (0.0053)	0.0084	0.0021
500		1	0.3098 (0.0098)	1.0139 (0.0139)	0.0091	0.0083
	0.6	0.3	0.6246 (0.0246)	0.3100 (0.0100)	0.0289	0.0023
		0.6	0.6253 (0.0253)	0.6199 (0.0199)	0.0273	0.0091
		1	0.6358 (0.0358)	1.0483 (0.0483)	0.0354	0.0369

Table 2. Results of maximum likelihood estimation.

Sample Size (n)	α (true)	β (true)	Estimate(α) (Bias(Estimate(α)))	Estimate(β) (Bias(Estimate(β)))	MSE (Estimate (α))	MSE (Estimate (β))
10	0.3	0.3	1.0000 (0.7000)	0.4967 (0.1967)	0.4900	0.0560
		0.6	0.5827 (0.2827)	0.7129 (0.1129)	0.3034	0.1026
		1	0.5008 (0.2008)	1.0733 (0.0733)	0.3342	0.4356
	0.6	0.3	1.0000 (0.4000)	0.4518 (0.1518)	0.1600	0.0414
		0.6	0.7387 (0.1387)	0.6654 (0.0654)	0.1619	0.0614
		1	0.6021 (0.0021)	1.0070 (0.0070)	0.2319	0.2545
25	0.3	0.3	0.9968 (0.6968)	0.4601 (0.1601)	0.4863	0.0338
		0.6	0.3723 (0.0723)	0.6419 (0.0419)	0.2090	0.0440
		1	0.3958 (0.0958)	1.1205 (0.1205)	0.2018	0.2098
	0.6	0.3	0.8059 (0.2059)	0.3605 (0.0605)	0.1323	0.0098
		0.6	0.6603 (0.0603)	0.6572 (0.0572)	0.1417	0.0390
		1	0.5758 (-0.0242)	1.0643 (0.0643)	0.1666	0.1721
50	0.3	0.3	0.9834 (0.6834)	0.4635 (0.1635)	0.4719	0.0329
		0.6	0.3329 (0.0329)	0.6179 (0.0179)	0.0959	0.0220
		1	0.3414 (0.0414)	1.0240 (0.0240)	0.1045	0.0680
	0.6	0.3	0.7739 (0.1739)	0.3555 (0.0555)	0.1215	0.0081
		0.6	0.6460 (0.0460)	0.6487 (0.0487)	0.1028	0.0306
		1	0.5880 (-0.0120)	1.0610 (0.0610)	0.1141	0.1169
100	0.3	0.3	0.5594 (0.2594)	0.3767 (0.0767)	0.1928	0.0129
		0.6	0.2829 (-0.0171)	0.5981 (-0.0019)	0.0313	0.0081
		1	0.2962 (-0.0038)	1.0080 (0.0080)	0.0324	0.0257
	0.6	0.3	0.7497 (0.1497)	0.3486 (0.0486)	0.0978	0.0067
		0.6	0.6193 (0.0193)	0.6297 (0.0297)	0.0756	0.0234
		1	0.6341 (0.0341)	1.0910 (0.0910)	0.0835	0.0959
150	0.3	0.3	0.5428 (0.2428)	0.3705 (0.0705)	0.1740	0.0122
		0.6	0.3090 (0.0090)	0.6079 (0.0079)	0.0251	0.0062
		1	0.3094 (0.0094)	1.0112 (0.0112)	0.0239	0.0183
	0.6	0.3	0.7451 (0.1451)	0.3455 (0.0455)	0.0858	0.0059
		0.6	0.6318 (0.0318)	0.6324 (0.0324)	0.0614	0.0199
		1	0.6446 (0.0446)	1.0756 (0.0756)	0.0664	0.0721
500	0.3	0.3	0.3133 (0.0133)	0.3052 (0.0052)	0.0088	0.0004
		0.6	0.2983 (-0.0017)	0.6006 (0.0006)	0.0065	0.0015
		1	0.3047 (0.0047)	1.0048 (0.0048)	0.0069	0.0057
	0.6	0.3	0.6597 (0.0597)	0.3179 (0.0179)	0.0380	0.0023
		0.6	0.6301 (0.0301)	0.6085 (0.0085)	0.0278	0.0052
		1	0.6351 (0.0351)	1.0199 (0.0199)	0.0339	0.0224

results show parameter estimates that are near to the real values of the parameters. And also, the mean square error (MSE) values are near zero. And, as well as the bias of the estimators are near to zero. EM algorithm compare with other of method of moments and maximum likelihood estimation is much better, accurate and appropriate for estimation of parameters.

4.2. Real Data Applications

This section shows how the $LB2$ Distribution can be used in practice by fitting it to a real-world data set.
Data Set: Incidence of Hepatitis B virus-positive of Azadi Teaching Gastroenterology and Hepatology Center (GHC) in Duhok, Kurdistan Region, Iraq
The data set in Table (5) shows the hepatitis B virus-positive incidence rates recorded at Azadi Teaching Gastroenterology and Hepatology Center (GHC) from 2017 through 2024 in Duhok, Kurdistan Region, Iraq.The dataset includes 742 individuals across 87 age-specific groups, covering the full range from childhood to older adulthood. The number of positive cases is represented by the variable x_i while the total number of trials is indicated by n_i. Table (4) demonstrates the goodness of fit results for the two-parameter Lindley Binomial distribution when compared against the Binomial distribution and negative-binomial distribution and Beta-Binomial distribution. Based on this result, two parameters Lindley–Binomial distribution is found to have the highest log-likelihood $(L = 163.0421)$ and lowest Akaike Information criterion $AIC = -322.0841$ as well as smallest $\chi^2 = 60.425$ compared with Binomial distribution, negative-binomial and Beta–Binomial distribution. Moreover, the following functions of Lindley-binomial and Beta-Binomial distributions are much better for the fitting this data compared with the Negative-Binomial and binomial distributions: Additionally, the Lindley-Binomial and Beta-Binomial distributions' probability mass functions are provided by:

$$p\left(X=x\right)=C_x^n\frac{\beta^2}{\beta+\alpha}\sum_{h=0}^{n-x}C_h^{n-x}(-1)^h\frac{\beta+x+h+\alpha}{\left(\beta+x+h\right)^2}$$

where $\hat{\alpha} = 0.1000$ and $\hat{\beta} = 0.4008$
and

$$p\left(X=x\right)=C_x^n\frac{B\left(x+c,n-x+d\right)}{B\left(c,d\right)}$$

where $\hat{c} = 1.4564$ and $\hat{d} = 3.0994$ These findings suggest that the Lindley-Binomial and Beta-Binomial distributions, rather than the Binomial and Negative-Binomial distributions, offer a better model for the data set. Previously, the method of moments which is typically less effective than other estimating techniques was used to estimate the parameters for the Beta-Binomial model. Consequently, for modelling this dataset, the two-parameter Lindley-Binomial distribution is a superior substitute for the Binomial distribution. . Figure (10) is the residual Q-Q plot, which tests for normality of the residual distribution. The residuals are mostly on the line of reference, suggesting a good fit, but the deviations of the upper quantiles reveal mild skewness or heavy tails. Figure (11) presents the variation in dispersion index across four distributions: Binomial,Beta-Binomial, Lindley-Binomial, and Negative-Binomial. The Lindley-Binomial is more flexible and has varying levels of dispersion across probability values, suggesting it has the capability of accommodating various levels of over- and under-dispersion (compared to the more limited Binomial,Beta-Binomial (red) and Negative-Binomial). The Negative-Binomial curve exhibits the least variability and hence the least flexibility for dispersion. The results reinforce the utility of LB2 in modeling real-world count data, and overdispersion and underdispersoin can matter. We then used reliability analysis on actual data. Figure shows the results of the reliability analysis.

The figure (12) illustrates the data indicates notable age-specific trends in HBV positive. The spike in hazard and reversed hazard rates around age 25 suggests a critical period of increased risk. Public health policies must prioritize targeted interventions for this demographic to mitigate HBV transmission. In order to sustain high levels of immunity in the community, early preventive and vaccination programs are crucial, as seen by the declining

Table 3. Results of expectation maximization algorithm

Sample Size (n)	α (true)	β (true)	Estimate(α) (Bias(Estimate(α)))	Estimate(β) (Bias(Estimate(β)))	MSE (Estimate (α))	MSE (Estimate (β))
10	0.3	0.3	0.3406 (0.0406)	0.3365 (0.0365)	0.1519	0.0249
		0.6	0.3433 (0.0433)	0.6992 (0.0992)	0.1573	0.1443
		1	0.3446 (0.0446)	1.2219 (0.2219)	0.1620	0.5642
	0.6	0.3	0.4311 (-0.1689)	0.2863 (-0.0137)	0.1894	0.0185
		0.6	0.4884 (-0.1116)	0.6184 (0.0184)	0.1821	0.1105
		1	0.5101 (-0.0899)	1.0841 (0.0841)	0.1884	0.3925
25	0.3	0.3	0.3232 (0.0232)	0.3259 (0.0259)	0.1200	0.0124
		0.6	0.3336 (0.0336)	0.6665 (0.0665)	0.1193	0.0617
		1	0.3486 (0.0486)	1.1482 (0.1482)	0.1303	0.2349
	0.6	0.3	0.5069 (-0.0931)	0.3003 (0.0003)	0.1449	0.0120
		0.6	0.5337 (-0.0663)	0.6195 (0.0195)	0.1402	0.0547
		1	0.5573 (-0.0427)	1.0601 (0.0601)	0.1423	0.1729
50	0.3	0.3	0.3194 (0.0194)	0.3180 (0.0180)	0.0863	0.0074
		0.6	0.3342 (0.0342)	0.6494 (0.0494)	0.0886	0.0403
		1	0.3470 (0.0470)	1.1074 (0.1074)	0.0947	0.1368
	0.6	0.3	0.5467 (-0.0533)	0.3056 (0.0056)	0.1115	0.0078
		0.6	0.5642 (-0.0358)	0.6206 (0.0206)	0.1041	0.0339
		1	0.5790 (-0.0210)	1.0508 (0.0508)	0.1065	0.1105
100	0.3	0.3	0.3229 (0.0229)	0.3128 (0.0128)	0.0565	0.0046
		0.6	0.3304 (0.0304)	0.6307 (0.0307)	0.0538	0.0209
		1	0.3357 (0.0357)	1.0658 (0.0658)	0.0555	0.0730
	0.6	0.3	0.5869 (-0.0131)	0.3101 (0.0101)	0.0769	0.0057
		0.6	0.5991 (-0.0009)	0.6266 (0.0266)	0.0697	0.0232
		1	0.6089 (0.0089)	1.0590 (0.0590)	0.0723	0.0790
150	0.3	0.3	0.3131 (0.0131)	0.3080 (0.0080)	0.0402	0.0030
		0.6	0.3156 (0.0156)	0.6171 (0.0171)	0.0342	0.0119
		1	0.3186 (0.0186)	1.0397 (0.0397)	0.0359	0.0432
	0.6	0.3	0.6035 (0.0035)	0.3112 (0.0112)	0.0601	0.0045
		0.6	0.6124 (0.0124)	0.6278 (0.0278)	0.0558	0.0192
		1	0.6248 (0.0248)	1.0644 (0.0644)	0.0602	0.0694
500	0.3	0.3	0.3056 (0.0056)	0.3027 (0.0027)	0.0103	0.0006
		0.6	0.3070 (0.0070)	0.6052 (0.0052)	0.0084	0.0022
		1	0.3083 (0.0083)	1.0113 (0.0113)	0.0080	0.0070
	0.6	0.3	0.6196 (0.0196)	0.3083 (0.0083)	0.0260	0.0020
		0.6	0.6188 (0.0188)	0.6159 (0.0159)	0.0234	0.0077
		1	0.6268 (0.0268)	1.0387 (0.0387)	0.0279	0.0296

Table 4. Results for the Hepatitis B virus-positive in Duhok dataset.

Distribution	Parameters	Log-Likelihood	AIC	Chi-Squared	p-value
Binomial	$\hat{p} = 0.3293$	-194.4131	390.8261	207.7336	2.866685e-12
Beta-Binomial	$\hat{c} = 1.4564$ $\hat{d} = 3.0994$	162.3379	-320.6758	61.502	0.00000
Lindley-Binomial	$\hat{\alpha} = 0.1000$ $\hat{\beta} = 0.4008$	163.0421	-322.0841	60.425	0.00000
Negative-Binomial	$\hat{p} = 1.0354$	-190.9346	385.8692	205.6898	0.00000

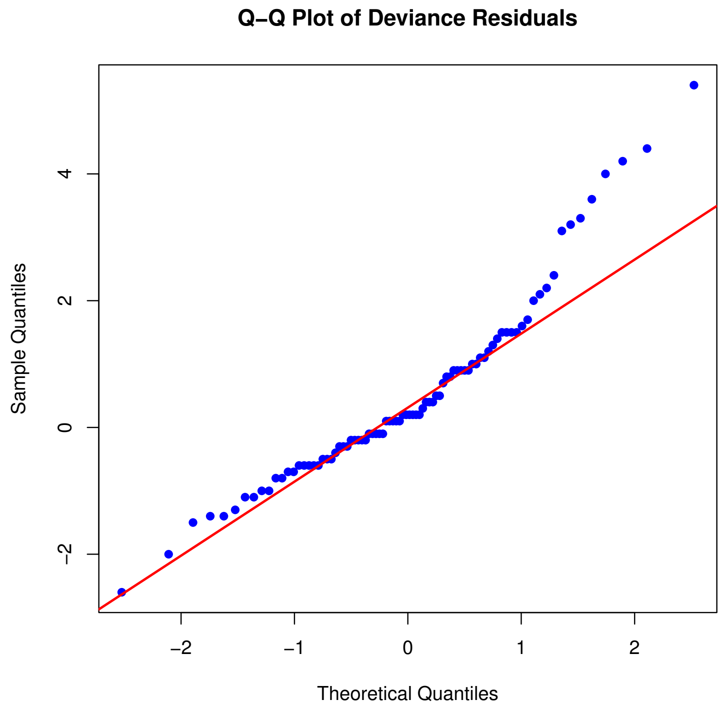

Figure 10. displays the Q-Q plot of residuals.

reliability and MRL with age. Lindley-Binomial Distribution-Based Age-Specific Evaluation of Hepatitis B Virus Positivity: Reliability, Hazard, and Reversed Hazard Functions.

5. Conclusion

This research, a new generalized discrete distribution, the $LB2$ distribution, was introduced by compounding the two-parameter Lindley distribution with the binomial distribution. Its probabilistic properties, including the shape of the probability mass function, generating functions, mean and variance, and index of dispersion, have

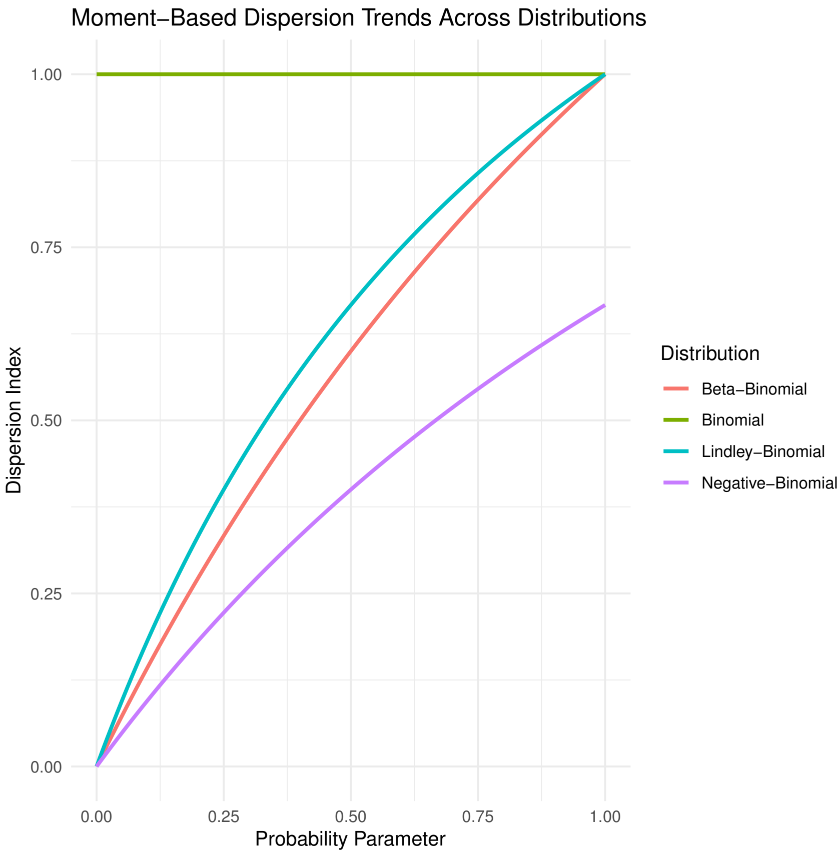

Figure 11. Index of Dispersion for various distributions Across Probability Values.

been discussed Furthermore, graphical representations of important functions, such as the mean residual life (MRL), hazard rate function, and reversed hazard rate function, were used to investigate reliability analysis. Three estimation methods: method of moments, maximum likelihood estimation, and the expectation-maximization algorithm—were used to estimate the distribution's parameters. The results indicate that all three methods provide good estimations; however, the expectation-maximization algorithm is the most stable and accurate compared to the method of moments and maximum likelihood estimation. The $LB2$ distribution was fitted on a dataset of the incidence of the Hepatitis B virus (HBV) and its positivity was evaluated in real world applications. To evaluate the goodness of fit of a $LB2$ distribution to those of binomial, beta binomial and negative binomial, metrics such as log likelihood, Akaike Information Criterion (AIC), chi-squared statistics and p values were used. Results indicate that the $LB2$ distribution fitted the data better than both binomial and negative binomial distributions and thus, had higher log-likelihood values and lower AIC scores. Additionally, the beta-binomial distribution and the $LB2$ distribution performed similarly. These results indicate that the two parameter Lindley Binomial distribution is particularly useful for modeling proportional data when over-dispersion is present. Because of its flexibility as well as its enhanced goodness-of-fit, the suggested distribution can be a useful tool in statistical modeling, and it may be used in epidemiology, reliability analysis, and medical research. Further extensions of the $LB2$ distribution are used in the further analysis of age specific HBV positive in this study. The findings can provide important data regarding risk patterns and thus be used to persuade public health campaigns to successfully end the spread of HBV.

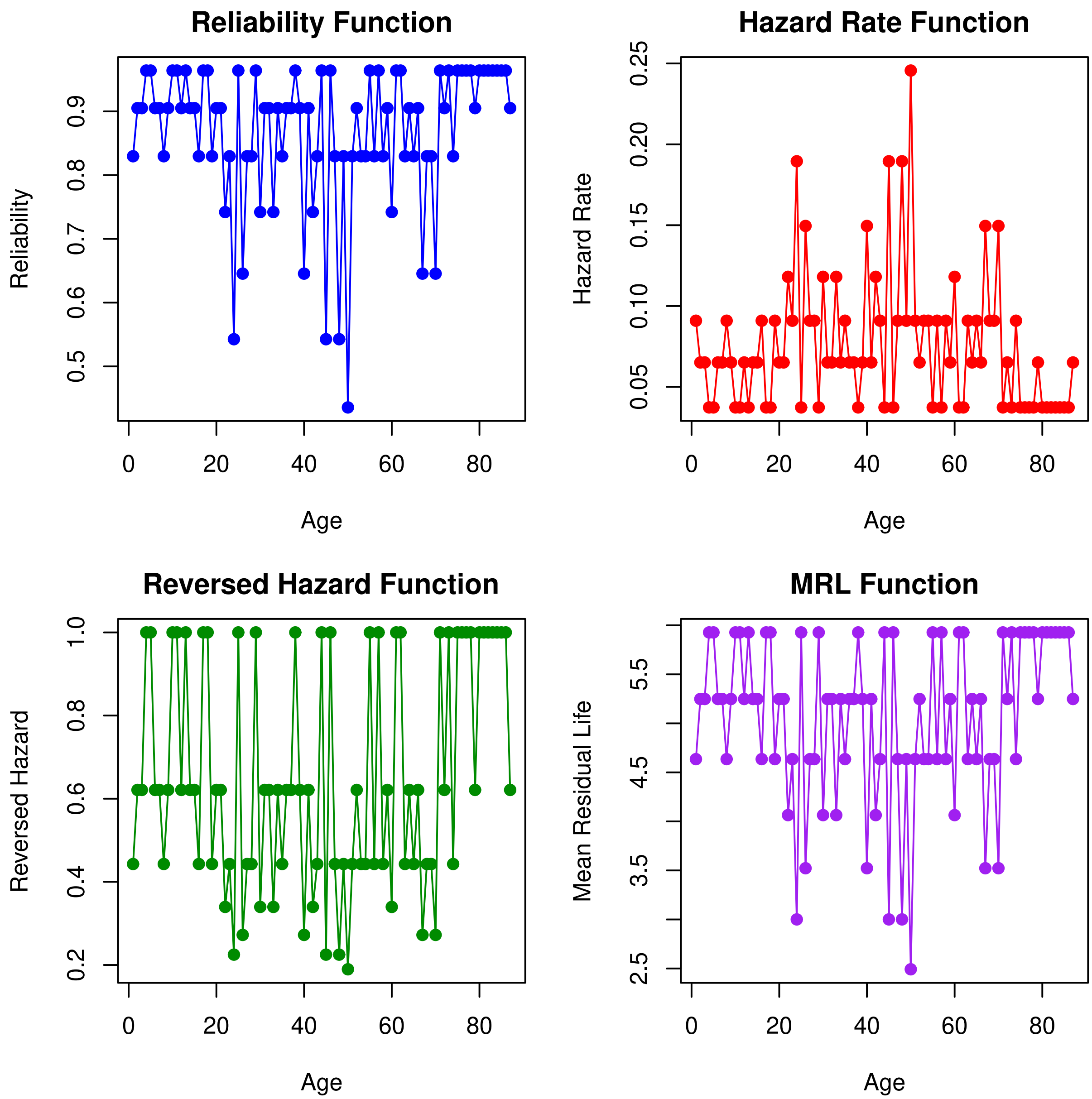

Figure 12. Lindley-Binomial Distribution-Based Age-Specific Evaluation of Hepatitis B Virus Positivity: Reliability, Hazard, , Reversed Hazard Functions and Mean Residual Life.

REFERENCES

1. Lindley, D. V. (1958). Fiducial Distributions and Bayes' Theorem. Journal of the Royal Statistical Society Series B: Statistical Methodology, 20(1), 102–107. https://doi.org/10.1111/j.2517-6161.1958.tb00278.x
2. Sankaran, M. (1970). 275. Note: The Discrete Poisson-Lindley Distribution. Biometrics, 26(1), 145–149. https://doi.org/10.2307/2529053
3. Zamani, H., & Ismail, N. (2010). Negative binomial-Lindley distribution and its application. Journal of Mathematics and Statistics, 6(1), 4–9. https://doi.org/10.3844/jmssp.2010.4.9
4. Bhati, D., Sastry, D. V. S., & Maha Qadri, P. Z. (2015). A new generalized poisson-lindley distribution: Applications and properties. Austrian Journal of Statistics, 44(4), 35–51. https://doi.org/10.17713/ajs.v44i4.54

Table 5. Hepatitis B virus positive in Duhok data set.

Age	Hepatitis B virus positive	Total	Age	Hepatitis B positive virus	Total	Age	Hepatitis B positive virus	Total	Age	Hepatitis B virus positive	Total
1	8	10	23	1	9	45	5	8	67	4	4
2	5	8	24	3	10	46	0	6	68	2	5
3	7	9	25	2	14	47	2	6	69	2	5
4	3	7	26	4	9	48	5	6	70	4	8
5	3	6	27	2	12	49	2	5	71	0	0
6	4	9	28	3	18	50	6	6	72	1	9
7	5	8	29	0	11	51	2	7	73	0	1
8	6	11	30	4	8	52	1	8	74	2	4
9	2	9	31	1	13	53	2	11	75	0	7
10	6	14	32	1	9	54	2	8	76	0	6
11	4	6	33	4	10	55	0	8	77	0	2
12	3	6	34	1	8	56	2	11	78	0	3
13	8	10	35	2	13	57	0	10	79	1	2
14	0	1	36	1	13	58	2	16	80	1	4
15	10	15	37	1	6	59	1	5	81	1	1
16	27	40	38	0	15	60	3	6	82	0	2
17	0	1	39	1	12	61	0	6	83	0	1
18	2	5	40	2	7	62	0	5	84	1	2
19	16	31	41	1	7	63	2	5	85	2	2
20	14	36	42	3	13	64	1	5	86	0	1
21	8	23	43	2	9	65	2	10	87	1	1
22	2	9	44	0	8	66	1	8			

5. S. A. Othman, "Enhancing reliability engineering through Weibull distribution in R," Al-Rafidain J. Comput. Sci. Math., vol. 18, no. 1, 2024.
6. S. A. Othman, K. M. T. Omar, and S. T. Abdulazeez, "Reliability analysis and statistical fitting for the transmuted Weibull model in R," vol. 72, no. 2, pp. 161–177, 2023.
7. S. A. Othman, 2024. Optimizing Poisson-Lindley Parameter Estimation: LQM and Reliability Analysis Applied to Guinea Pig Survival Data. Wasit Journal of Computer and Mathematics Science, 3(4), pp.83-95.
8. S. A. Othman, S. T. Qadir, and S. T. Abdulazeez, "Comparative analysis of scaling parameter estimation for the Lindley distribution in wait time analysis: Simulated and real data applications," Tuijin Jishu/Journal of Propulsion Technology, vol. 44, no. 2, 2023.
9. Shanker, R., Sharma, S., & Shanker, R. (2013). A Two-Parameter Lindley Distribution for Modeling Waiting and Survival Times Data. Applied Mathematics, 04(02), 363–368. https://doi.org/10.4236/am.2013.42056
10. Deng, D., & Zhang, X. (2024). A Lindley–binomial model for analyzing the proportions with sparseness and excessive zeros. Journal of Applied Statistics, 51(9), 1792–1817. https://doi.org/10.1080/02664763.2023.2237212