WJCMS

Wasit Journal of Computer and Mathematics Science

Journal Homepage: https://wjcm.uowasit.edu.iq/index.php/WJCM
e-ISSN: 2788-5879 p-ISSN: 2788-5887

Optimizing Poisson-Lindley Parameter Estimation: LQM and Reliability Analysis Applied to Guinea Pig Survival Data

Sameera Abdulsalam Othman[1,*]
Department of Mathematics, College of Basic Education, University of Dohuk, Iraq.

*Corresponding Author: Sameera Abdulsalam Othman

DOI: https://doi.org/10.31185/wjcms.261
Received 20 July 2024; Accepted 11 November 2024; Available online 30 December 2024

ABSTRACT: The main objective of this work is to estimate the scale parameter of the Poisson Lindley distribution by means of multiple approaches, such as Poisson Linear Quantile-Moment and Maximum Likelihood. Based on mean square error criteria (MSE), Akaike information criterion (AIC), and Bayesian information criterion (BIC), Linear Quantile-Moment is the most efficient estimator among these techniques. The study focuses on reliability analysis and investigates the probability functions of the distribution to create a theoretical framework for parameter estimation by Using R programming language for in-depth analysis. Through simulation and real data analysis, several estimation techniques are compared and contrasted, demonstrating the superiority of the Linear Quantile Moment approach in terms of accuracy and model fit. The Poisson Lindley Distribution parameter estimation is improved in this work, which has implications for environmental research, finance, and epidemiology. Moreover, variance estimates for the known parameters and the related Kolmogorov–Smirnov (K–S) statistics, along with their corresponding p-values for the Poisson-Lindley Distribution (PLD), are analyzed using actual data on guinea pig survival times under various tubercle bacilli dosages. An observation indicating a strong fit with the optimal estimator with (LQM=4.190217) and the lowest MSE (78.71956) is made in light of the small K–S distance and the significant p-value for the test.

Keywords: Poisson Lindley, Reliability, Linear Quantile-Moment, Stress-Strength Reliability, Maximum Likelihood

1. INTRODUCTION

In many disciplines, including epidemiology, finance, and environmental research, the Poisson–Lindley distribution (PLD) is a crucial distribution for analyzing over dispersed count data. We observe that the typical Poisson distribution, which requires that the variance and the mean be equal, cannot handle overdispersion in count data, when the variance is greater than the mean. The parameters of the Poisson–Lindley distribution must be estimated using a variety of techniques in order to increase the model's accuracy and dependability. It makes it a crucial field of research for statistical modelling. The growing need for precise and trustworthy parameter estimate techniques inside the Poisson–Lindley distribution is what spurred this investigation. To increase the prediction power of statistical models and better comprehend the behavior of the data, particularly in the presence of overdispersion, accurate parameter estimations are crucial. In applications like disease outbreak modelling, financial risk assessment, and environmental monitoring where count data is widely used, this precision is crucial. In these areas, precise estimation plays a key role in informing policy decisions, managing risks, and guiding interventions. However, despite its advantages, the Poisson-Lindley distribution presents challenges in parameter estimation due to the complexity of over-dispersed count data.

Traditional methods, such as those used for the simpler Poisson distribution, often fall short when dealing with the variability seen in real-world data. To address this, alternative distributions like the Poisson-Generalized Lindley (PGL) distribution have been proposed to handle sparse and highly variable data more effectively [1]. Additionally, recent advancements in reliability analysis have offered robust techniques for assessing the stability and accuracy of parameter estimates, contributing to more reliable statistical models [2]. Recent literature also highlights the importance of developing new methods for estimating parameters in discrete distributions. For instance, [3] introduced the Poisson quasi-XLindley distribution, a novel two-parameter discrete distribution framework that addresses limitations in existing methods and offers enhanced parameter estimation techniques. Furthermore, [4] introduced parameters and reliability

*Corresponding author: sameera.othman@uod.ac

characteristics of the failure distribution of a mixture of failure rates, which are estimated based on a complete sample using both the Markov Chain Monte Carlo (MCMC) method and Maximum Likelihood Estimation (MLE). The research landscape is further enriched by alternative mixing distributions, such as the inverse Poisson Gaussian distribution [5], and the Lindley distribution, which acts as the mixing component of the Poisson-Lindley distribution [6]. Building on these foundational works, Mahmoudi and [7] introduced the generalized Poisson-Lindley distribution using the generalized Lindley distribution as the mixing component. Similarly, [8] proposed the Poisson weighted exponential distribution, leveraging the weighted exponential distribution for modeling over-dispersed count data. In addition, recent studies have expanded the application of alternative distributions to lifetime data modeling. For example, [9] explored the use of the Weibull distribution for modeling lifetime data, showcasing its versatility in real-world scenarios. This evaluation involved using key model comparison metrics such as Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), alongside MLE for parameter estimation. Likewise, [10] focused on the transmuted Weibull distribution, emphasizing its utility in practical applications. Building on this foundation, the present study critically evaluates various parameter estimation methodologies for the Poisson-Lindley distribution, incorporating recent developments in reliability analysis and statistical techniques. By doing so, this research contributes to the ongoing efforts to improve statistical modeling for over-dispersed count data, with particular attention to its practical applications in fields where accurate estimation is essential.

This research is structured as follows. Section 2 presents the methodology of the Poisson-Lindley distribution, followed by Section 3, which introduces various estimation approaches for the distribution. Section 4 discusses the application of these techniques, and Section 5 provides a summary and conclusions.

2. METHODOLOGY

The probability density function (PDF) of Lindley distribution's is given by:

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

Where θ is the shape parameter.

The (PLD), which was proposed for modeling count data, is derived from the Poisson distribution. The parameter λ of the Poisson distribution follows the Lindley distribution, The Poisson probability density function(pdf) is given:

$$P(X=x\mid\lambda)=\frac{e^{-\lambda}\lambda^x}{x!};\lambda>0,x=0,1,2,\ldots$$

The probability density function of Lindley distribution's is:

$$f(\lambda;\theta)=\frac{\theta^2(1+\lambda)e^{-\lambda\theta}}{\theta+1};\lambda>0,\theta>0$$

Therefore, the probability mass function (pmf) of the Poisson-Lindley distribution can be written as:

$$P(X=x)=\int_0^\infty P(X=x\mid\lambda)f(\lambda)d\lambda=\int_0^\infty\frac{e^{-\lambda}\lambda^x\theta^2(1+\lambda)e^{-\lambda\theta}}{x!\,(\theta+1)}d\lambda$$

$$=\frac{\theta^2}{x!(\theta+1)}\int_0^\infty e^{-(\theta+1)\lambda}\,(\lambda^x+\lambda^{x+1})d\lambda$$

Solving the integral gives the PMF of the Poisson-Lindley distribution, which is defined by the following equation [11]

$$p(x)=\frac{\theta^2(\theta+x+2)}{(1+\theta)^{(x+3)}},\ for\ x=0,1,2,\ldots\qquad for\ \theta>0\,. \tag{1}$$

Where. X: count variable, θ: shape parameter.

This equation expresses the probability mass function of the Poisson-Lindley distribution. We will discuss the distribution's characteristics, including the probability mass function graph, mean, variance, moment, skewness, and kurtosis. Next, we'll discuss using PLD to generate random numbers and estimate their parameters. Figure 1 shows how the pmf for the Poisson-Lindley distribution in Equation (1) evolves with θ values. It also displays the function's shape for each parameter selection.

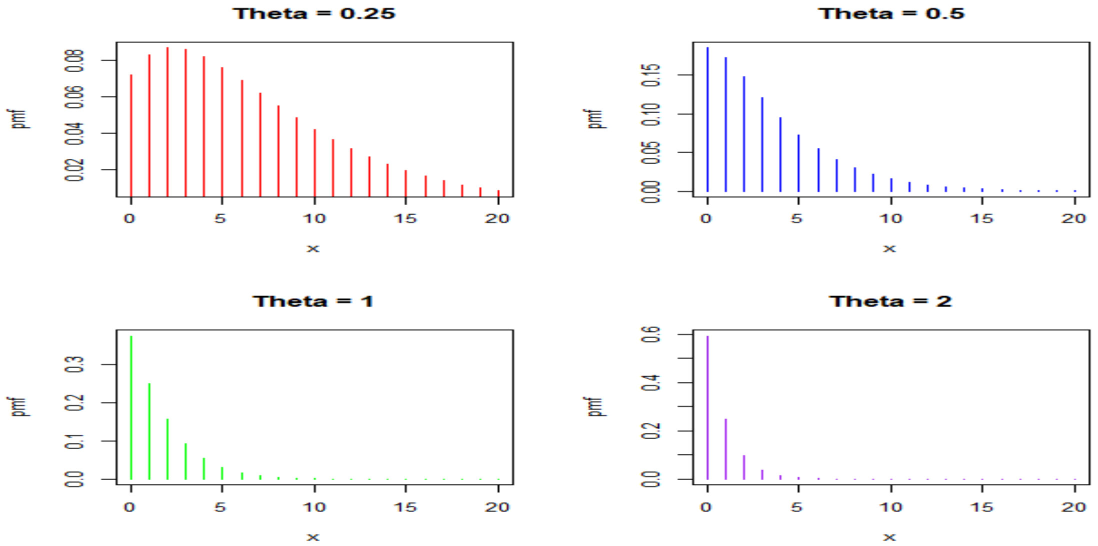

FIGURE 1: various parameter choices and the shape of the Poisson Lindley Distribution density function.
The cumulative distribution function (cdf) of the Poisson-Lindley Distribution can be expressed as follows [6]:

$$F(x) = P(X \le x) = 1 - P(X > x + 1) = 1 - \sum_{t=x+1}^{\infty} \int_0^{\infty} P(X = t|\lambda) f(\lambda) d\lambda$$

$$= 1 - \sum_{t=x+1}^{\infty} \int_0^{\infty} \frac{e^{-\lambda}\lambda^t}{t!} \frac{\theta^2}{(\theta+1)} (1 + \lambda) e^{-\theta\lambda} d\lambda$$

$$F(x) = 1 - \frac{\theta^2 + 3\theta + 1 + \theta x}{(\theta+1)^{x+3}} \quad , x = 0, 1, 2, \ldots \qquad for\ \theta > 0 \qquad (2).$$

Figure 2 illustrates the cumulative distribution function (CDF) given by Equation (2) for varying values of θ. Additionally, the quantile function (QF) of the Poisson-Lindley distribution are provided.

$$x_w = -3 - \theta - \frac{1}{\theta} - \frac{1}{\log(\theta+1)} U\left[\frac{(\theta+1)^{-\frac{\theta^2+1}{\theta}} (w-1) log(\theta+1)}{\theta}\right] \qquad (3)$$

accordingly, where $U[\cdot]$ denotes the Lambert function's negative branch.

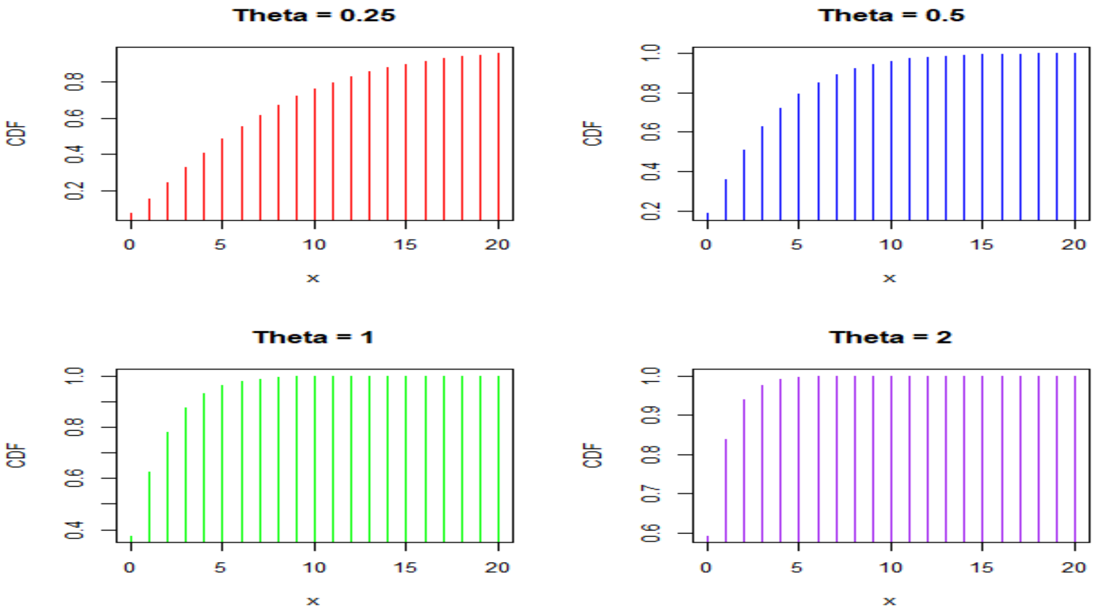

FIGURE 2: various parameter choices and the shape of the cumulative distribution function of the Poisson Lindley Distribution.

2.1 MOMENT

The (PLD) distribution's rth central moment has the following formula:

$$E(x^r) = \mu^r \quad (4)$$

To find the derivative of the mean of the PLD distribution, we need to first calculate the first central moment or the mean:

$$\mu = E(x)$$

After simplification and solving (4) the integral, we get:[12]

$$\mu = \frac{(\theta+2)}{\theta(1+\theta)} \quad (5)$$

$$E[X^2] = \frac{\theta^2 + 4\theta + 6}{(\theta + 1)\theta^2}$$

The formula can be used to compute the PLD distribution's variance.:

$$\sigma^2 = E[X^2] - (E[X])^2$$

$$\sigma^2 = \frac{\theta^3 + 4\theta^2 + 6\theta + 2}{\theta^2(\theta + 1)^2} \quad (6)$$

where $E[X^2]$ is the second moment and $E[X]$is the mean.

Note that $\sigma^2 = \mu\left[1 + \frac{(\theta^2+4\theta+2)}{\theta(\theta+1)(\theta+2)}\right] > \mu$, This indicates that the Poisson-Lindley distribution is overly scattered. Finally, the third and fourth central moments are:

$$E[X^3] = \frac{(\theta^2 + 6\theta + 12)(\theta + 2)}{(\theta + 1)\theta^3}$$

$$E[X^4] = \frac{(\theta^4 + 16\theta^3 + 78\theta^2 + 168\theta + 120)}{(\theta + 1)\theta^4}$$

Thus, the skewness and kurtosis of the PLD are provided by the following relations, respectively:

$$skewness = \frac{2(\theta+1)^4(\theta+2) - \theta^3(\theta+2)(\theta+3)}{[2(\theta+1)^3 - \theta^2(\theta+2)]^{3/2}} \quad (8)$$

$$\text{kurtosis} = 3 + \frac{2(\theta+1)^5[(\theta+3)^2 - 3] - \theta^4(\theta+2)[(\theta+4)^2 - 3]}{[2(\theta+1)^3 - \theta^2(\theta+2)]^2} \quad (9)$$

3. VARIOUS ESTIMATION TECHNIQUES

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

3.1 MAXIMUM LIKELIHOOD ESTIMATOR

This section explains how to derive MLEs from a PLD (θ) distribution's unknown parameter. Consider a sample of size n from a PLD (θ) distribution as $X = (X_1, X_2, \ldots, X_n)$ [12, 13]. The likelihood function can be given as follows based on the observation function (PDF) evaluated at each observation in the sample. The log-likelihood function for the PLD distribution is:

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \ln[u(x_i; \theta)]$$

where n= sample size and x_i = observed values in the sample. To maximise the log-likelihood function, we must take the partial derivatives of each parameter (θ) and set them to zero. The log-likelihood function's derivative is:

$$L(\theta) = \prod_{i=1}^{n} f(X_i, \theta) = \prod_{=1}^{n} \frac{\theta^2 * (\theta + X_i + 2)}{(1+\theta) \sum_{i=1}^{n} \ln\left(\frac{\theta^2*(\theta+X_i+2)}{(1+\theta)^{X_i+3}}\right) + 3}$$

$$l(\theta) = \ln(L(\theta)) = \sum_{i=1}^{n} \ln\left(\frac{\theta^2*(\theta+X_i+2)}{(1+\theta)^{X_i+3}}\right) \quad (10)$$

Although Equation (11) has a unique solution for all n, it lacks a closed-form solution and must be calculated numerically. [14].

$$\frac{\partial l}{\partial \theta} = \frac{2n}{\theta} - \frac{n(\bar{x} + 3)}{\theta + 1} + n \sum_{i=1}^{n} \frac{1}{x_i + \theta + 2} = 0 \quad (11)$$

3.2 LINEAR QUANTILE MOMENT METHOD

Obtaining the Quantile Function involves deriving it from the Cumulative Function (F(x)), using equation (3), [14 & 15].

The Quantile moments of a random sample of size n

$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$

As follows

$$\hat{\epsilon}_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r}{k} \hat{\tau}_{p,m}(X_{r-k:r}) \quad (12)$$

Where $X_{r-k:r}$ represents the ordered sample values from X_{r-k} to X_r.
$\hat{\tau}_{p,m}(X_{r-k:r})$ is the quantile estimator defined as a weighted combination of quantile estimates:

$$\hat{\tau}_{p,m}(X_{r-k:r}) = p\hat{Q}_{r-k:r}(m) + (1-2p)\hat{Q}_{r-k:r}\left(\frac{1}{2}\right) + p\hat{Q}_{r-k:r}(1-m) \quad (13)$$

p is the quantile level typically set to 0.5 for median estimation.
m is a parameter that determines the fraction of the quantile level on each side typically set to 0.5 for symmetric estimation.
$\hat{Q}_{r-k:r}(u)$ is the quantile estimator obtained using the sample data.

$$\hat{Q}_{r-k:r}(u) = \sum_{i=1}^{n}\left[\frac{1}{n}k_h\left(\sum_{j=1}^{i} w_{j,n} - u\right)\right]X_{i,n} \quad (14)$$

$k_{h(t)}$ is the kernel function which is defined as the standard normal density function:

$$k_h(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} \quad (15)$$

$w_{j,n}$ are the weights used in the quantile estimator defined as

$$w_{j,n} = \begin{cases} \frac{1}{2}\left(1 - \frac{n-2}{\sqrt{n(n-1)}}\right), & if\ i = 1, n \\ \frac{1}{\sqrt{n(n-1)}}, & if\ i = 1,2,\dots,n-1 \end{cases} \quad (16)$$

With the aim of deriving estimations through the LQM approach, equation 12 is employed, and the R programming language is utilized to determine the parameter

4. ANALYSIS OF RELIABILITY

The Poisson Lindley Distribution's reliability function(survival function)[16, 17], which is the survival analysis, can be represented as the complement of its cumulative distribution function. Figure 3 depicts the reliability function for various θ values., given by
R(t) = 1 - F(t)

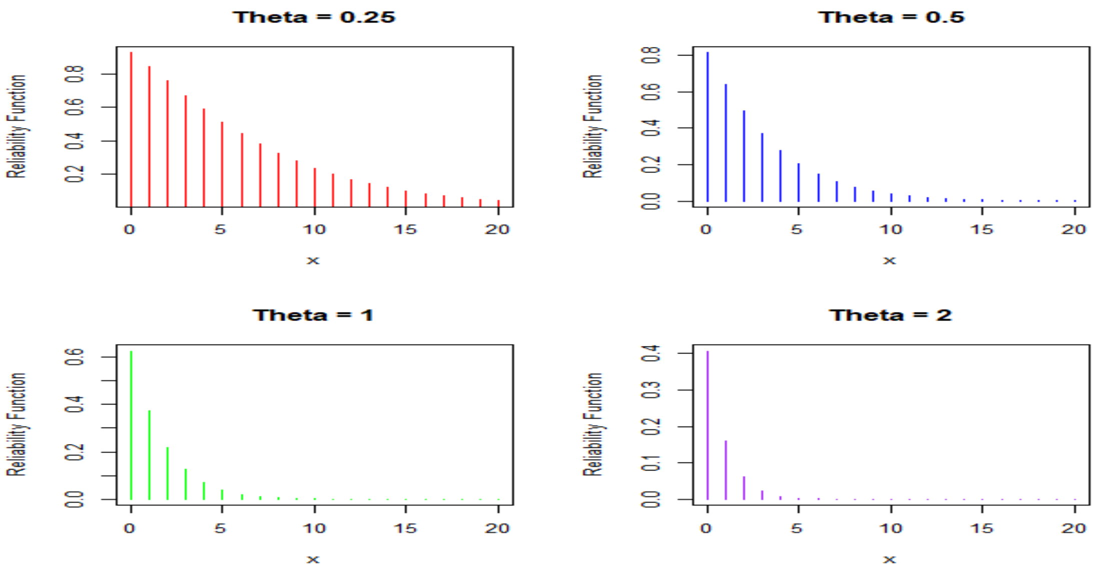

FIGURE 3: The survival analysis (Reliability Function) of the Poisson Lindley Distribution

4.1 HAZARD RATE FUNCTION

The Hazard Rate (HR) Function (failure rate) of a lifetime random variable X with the Poisson Lindley Distribution (PLD) is given by.

$$h((t)) = \frac{f(t)}{R(t)}$$

$$h(t) = \frac{\frac{\theta^2(\theta+t+2)}{(1+\theta)^{(t+3)}}}{1-\left[1-\frac{(\theta^2+3\theta+1+\theta t)}{(\theta+1)^{(t+3)}}\right]} \quad (17)$$

Figure 4 depicts the Hazard Rate Function

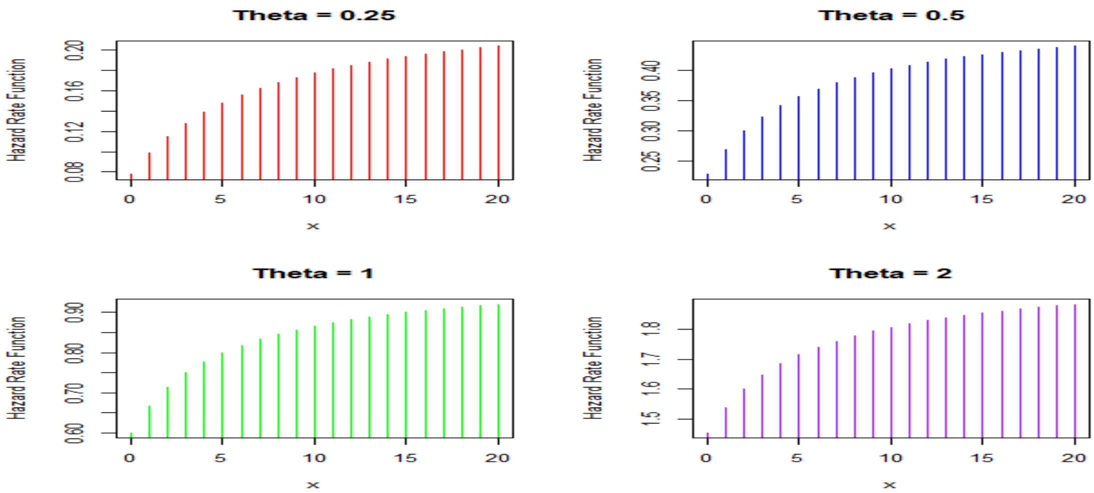

FIGURE 4: Hazard Rate Function (HR) of random variable x for the lifetime with the (PLD)

4.2 REVERSED HAZARD RATE FUNCTION

The Reversed Hazard Rate Function for a discrete random variable is defined as the ratio of the pmf f(t) to the cumulative distribution function F(t) For the Poisson-Lindley Distribution, it can be expressed as follows [18]:

$$\lambda^*(x) = \frac{f(x)}{F(x)} \tag{18}$$

The Reversed Hazard Rate Function is depicted in Figure 5.

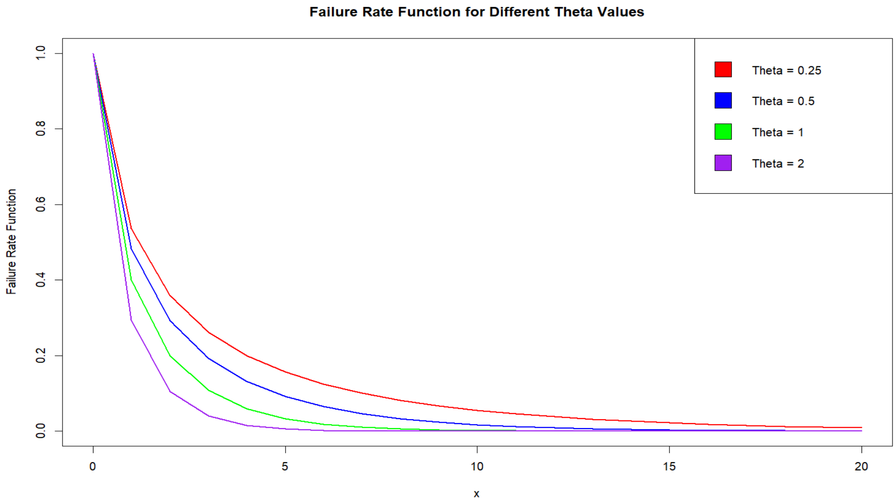

FIGURE 5: The Reversed Hazard Rate Function for PLD with different value of θ

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:

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

$$\text{MRL}(t, \theta) = \frac{1}{R(t)} \int_t^\infty x\, f(x) dx - t \quad \text{where} \quad t > 0 \tag{19}$$

Where:

- t is a specific time point for which you want to calculate the MRL.
- θ is a parameter of the 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 θ.

- x is a variable representing time to failure.

This equation calculates the expected remaining lifetime at time t for a random variable x following a specific distribution characterized by the parameter θ. It takes into account both the shape of the distribution (through the survival function) and the time point t of interest.

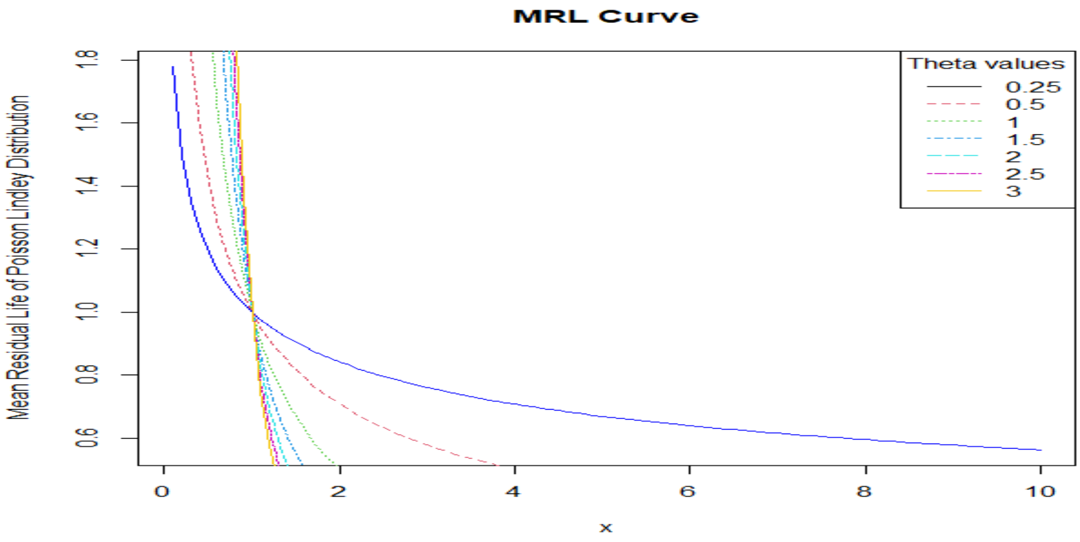

FIGURE 6: The Mean Residual Life (MRL)

4.4 MEAN INACTIVITY TIME

The Mean Inactivity Time (MIT) function is used as a reliable measure in forensic science, reliability theory, and survival analysis, to name a few fields. The MIT function for the random variable x's lifetime is shown below [5].

$$MIT(x) = \sum_{k=x+1}^{\infty} (k-x) \frac{\theta^2(k+\theta+2)}{(1+\theta)^{k+3}} \quad (20)$$

Figure (7) shows The value of the Lindley Mean Inactivity Time with parameters θ = 0.25, 0.5 , 1, 1.5, 2, 2.5, 3.

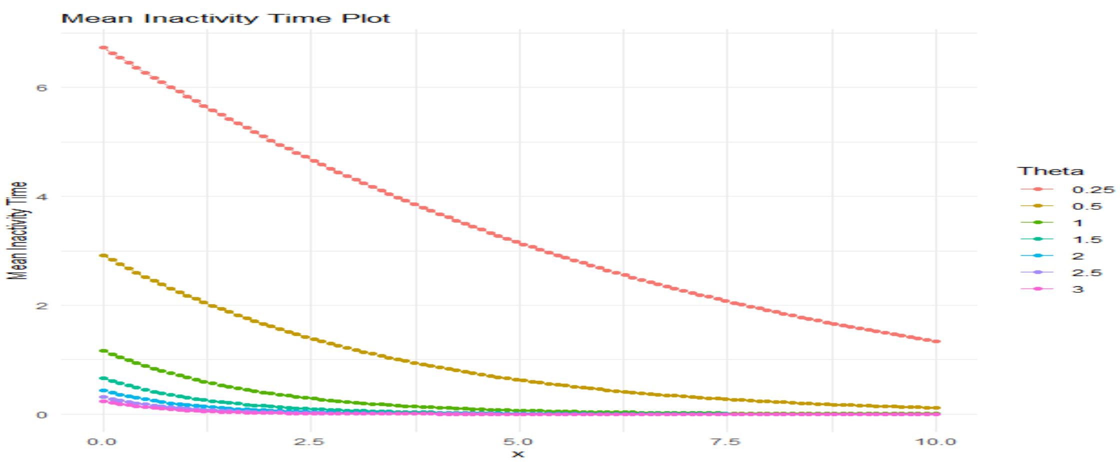

Figure 7: Mean Inactivity Time

4.5 STRESS-STRENGTH RELIABILITY

The life of a component with random strength X and random stress Z is described by the stress-strength model in dependability [19,20]. The component will stop working properly when the stress applied to it surpasses its strength, and it will fail at the moment when $X > Z$. As a result, a measure of component reliability is $R = Pr(X > Z)$. It has numerous uses in a variety of scientific and engineering fields. Now that X and Z have independent f(x, θ_1) and F(z, θ_2)) distributions, we can calculate the reliability R. One can express the PDF of X and the CDF of Z, respectively, as

$$SSR = \sum_{x=0}^{\infty} [\frac{{\theta_1}^2(\theta_1+x+2)}{(1+\theta_1)^{x+3}}] [1-(1-\frac{{\theta_2}^2+3\theta_2+1+\theta_2 x}{(\theta_2+1)^{x+3}})] \quad (21)$$

The likelihood function $L(\theta_1,\theta_2)$ can be expressed as:

$$L(\theta_1,\theta_2) = \prod_{i=1}^{n} [\frac{{\theta_1}^2(\theta_1 + t + 2)}{(1+\theta_1)^{t+3}}] [1 - \frac{{\theta_2}^2 + 3\theta_2 + 1 + \theta_2 t}{(\theta_2 + 1)^{t+3}}]$$

To maximize the likelihood, we typically take the logarithm (log-likelihood) and solve for the parameters:

$$\log\ L(\theta_1,\theta_2) = \sum_{i=1}^{n} \log [\frac{{\theta_1}^2(\theta_1 + t + 2)}{(1+\theta_1)^{t+3}}] + \log [1 - \frac{{\theta_2}^2 + 3\theta_2 + 1 + \theta_2 t}{(\theta_2 + 1)^{t+3}}]$$

Next, compute the partial derivatives of the log-likelihood function with respect to θ_1 and θ_2 .

$$\frac{\partial \log\ L}{\partial \theta_1} = \sum_{i=1}^{n} [\frac{\partial}{\partial \theta_1} \log (\frac{{\theta_1}^2(\theta_1 + t + 2)}{(1+\theta_1)^{t+3}})] = 0$$

$$\frac{\partial \log\ L}{\partial \theta_2} = \sum_{i=1}^{n} [\frac{\partial}{\partial \theta_2} \log (1 - \frac{{\theta_2}^2 + 3\theta_2 + 1 + \theta_2 t}{(\theta_2 + 1)^{t+3}})] = 0$$

After obtaining the log-likelihood function, numerical optimization methods (e.g., using the optim function in R) can be applied to find the values of $\hat{\theta}_1$ and $\hat{\theta}_2$ that maximize the log-likelihood, thereby providing the MLE estimates. The results are presented in Table 1.

Table 1: Estimates of the stress- strength Reliability with parameters, at t=1

parameter	Estimate	R(t)	HR	RH	MRL	AIC	BIC
$\theta_1 = 0.9$ $\theta_2 = 0.8$	38.37324 57.38643	2.011730e-01	37.409	0.02673	4.940523e-02	77.22048	82.43082
$\theta_1 = 0.7$ $\theta_2 = 0.8$	64.76526 57.38639	1.598804e-01	13.233	0.01568	4.028622e-02	81.90263	87.11297

5. APPLICATION

5.1 SIMULATION STUDY

Simulation serves as the method for representing real-world phenomena through specific models. In the realm of complex operations, which are often challenging to comprehend and analyze, models resembling real-world scenarios are instrumental. These models aid in grasping and scrutinizing intricate processes. Simulation, as a tool, enhances our understanding of original processes and real-world dynamics. This research stage focuses on a simulation study involving the generation of data using the inverse transformation method based on the cumulative distribution function equation (3). The primary objective is to compare the performance of two distinct estimators: Maximum Likelihood Estimators (MLEs) and the combination of Least Square and LMQ estimators. This comparison will be grounded in their respective estimates and mean squared errors (MSEs). The study incorporates various sample sizes (25, 50, 75, 100, 150), employs the R program, and explores diverse values for the theta parameters. To encompass all conceivable combinations of sample size and shape parameter values, the experiment will be iterated 1000 times. The outcomes, encompassing the estimated parameters and MSEs for both estimators, will be systematically presented in Tables 2 .

Table 2: MSE of the parameter estimations and a comparison of the two methods of estimation at the sample sizes (25,50,100,150) For the initial value set.

methods	Sample size	value	estimate	Statistics			Reliability
		θ	$\hat{\theta}$	MSE($\hat{\theta}$)	AIC	BIC	

MLE	25	0.25	0.316	0.004356	25.2783	25.668	0.47116
LQM			0.2818	0.00101124	17.4767	17.866	0.7292351
MLE	50	0.25	0.2123	0.00142129	18.8638	18.9791	0.441222
LQM			0.2715	0.00046225	14.8947	14.9100	0.79917
MLE	100	0.25	0.234487	0.0002406532	18.6254	18.7666	0.42775
LQM			0.2382	0.00013924	14.14601	14.56310	0.8128402
MLE	150	0.25	0.24365	4.03225e-05	18.58464	18.6915	0.423222
LQM			0.2441	3.481e-05	14.05621	14.2872	0.820173
MLE	25	0.5	0.35083	0.02225169	24.148545	24.5386	0.45901
LQM			0.5554304	0.00307252	20.337444	20.72752	0.729602
MLE	50	0.5	0.52720473	0.000740097	22.68327	22.89854	0.435071
LQM			0.48514327	0.000220722	16.12748	16.34275	0.5130255
MLE	100	0.5	1.1470	0.4186	349.5226	56.4329	0.4245096
LQM			0.7312	0.0534	-15.5903	-8.38004	0.4964696
MLE	150	0.5	1.14244	0.4127	509.1133	517.1346	0.42103
LQM			0.7610	0.0681	-27.2034	-19.3822	0.4807262
MLE	25	1	1.0738264	0.005450345	19.31630	19.45751	0.437082
LQM			1.028069	0.000787868	11.145448	11.53552	0.8586
MLE	50	1	1.06510	0.00423801	19.25684	19.39238	0.4256
LQM			1.01817774	0.000330430	10.095828	10.2370	0.8101
MLE	100	1	0.9854590	0.000211440	19.177117	19.36374	0.4203
LQM			0.9887855	0.000125763	10.095733	10.20262	0.842517
MLE	150	1	1.0007744	5.997263e-07	18.51142	18.90150	0.41842
LQM			1.0000875	7.668505e-09	9.3375461	9.552818	0.89574
MLE	25	1.5	2.1289	0.3955	20.7402125	20.9554848	0.615369
LQM			0.9031	0.3563	13.2556176	13.645697	0.60930
MLE	50	1.5	2.0579	0.3112	20.3481738	20.4893795	0.42187
LQM			0.9455	0.3074	11.5160797	11.6572854	0.6968512
MLE	100	1.5	2.0215	0.2719	20.227250	20.6173304	0.41865
LQM			0.9861	0.2640	11.480190	11.4801902	0.8265945
MLE	150	1.5	2.0096	0.2596	20.1585418	20.2654372	0.4173
LQM			0.9947	0.2553	11.3844624	11.5997348	0.8263621
MLE	25	2	2.99975	0.9995	19.4512682	19.592473	0.568742
LQM			2.35840	0.1284506	11.0760328	11.466112	0.46353

MLE	50	2	2.461513	0.2129942	19.3991800	19.529776	0.5391
LQM			2.344135	0.1184	10.344744	10.48594	0.460023
MLE	100	2	2.295922	0.0875698	19.3145039	19.506075	0.4177
LQM			2.258038	0.0665836	10.3378268	10.444722	0.458388
MLE	150	2	2.286174	0.0818955	17.8485466	18.238626	0.55919
MLE	25	2.5	2.21659	0.080321	18.6930806	18.83428	0.5767
LQM			2.25232	0.061345	10.9456870	11.33576	0.3637
MLE	50	2.5	2.76611	0.070818	18.6209113	18.72780	0.51996
LQM			2.71842	0.047707	10.3966089	10.53781	0.36289
MLE	100	2.5	2.27823	0.049181	18.2968551	18.51212	0.55242
LQM			2.30279	0.038891	10.3799173	10.48681	0.36328
MLE	150	2.5	2.304577	0.038190	16.9933777	17.38345	0.558450
LQM			2.34514	0.023981	10.233130	10.44840	0.36337
MLE	25	3	2.18756	0.660058	18.29747	18.438684	0.57564
LQM			2.70125	0.089251	10.89592	11.28600306	0.30306
MLE	50	3	2.3806	0.560536	18.17272	18.2796201	0.51883
LQM			2.851310	0.022108	10.31502	10.42192432	0.30297
MLE	100	3	2.380371	0.383940	17.82421	18.03948788	0.54345
LQM			2.88747	0.012663	10.24445	10.38566187	0.30372
MLE	150	3	2.81755	0.033288	16.79956	17.189648	0.49514837
LQM			2.9344	0.004303	10.07834	10.293615	0.3035

5.2 DISCUSSION

Table 2's simulation experiments show that the LQM method consistently produces lower MSE values and has a better model fit, as indicated by the AIC and BIC, than MLE Estimation. This constant advantage motivates additional investigation into why LQM beats MLE, particularly given the parameters of this study, which include small to medium sample sizes, a wide range of parameter values, and unique distribution features.

Theoretical Foundations of MLE and LQM: MLE has long been the standard approach for parameter estimation due to its strong theoretical foundations. MLE determines parameter values that maximise the likelihood function and make the observed data most likely. As sample sizes increase, MLE estimators become more efficient, producing unbiased estimates with low variation. However, these asymptotic gains might be lost in smaller samples, when MLE may become unreliable, particularly in the presence of outliers or when important distributional assumptions are violated. LQM, on the other hand, focusses on quantiles (such as medians or percentiles) rather than MLE's moments (mean, variance). This focus allows LQM to be more resilient to departures from normality and the assumptions that MLE is based on. Because LQM is less impacted by extreme values and outliers, it produces more stable estimates in small to medium sample sizes, making it a viable option when MLE assumptions are not met.

Flexibility with Model Misspecification: One key advantage of LQM is its adaptability in the situation of model misspecification. MLE relies largely on the chosen model accurately representing the underlying data-generation process. Deviations from the anticipated distribution can produce biassed estimates, raising MSE and reducing model fit, as evidenced by AIC and BIC. In contrast, LQM's lower reliance on the specifics of the likelihood function increases its adaptability in such scenarios, allowing it to perform effectively even when the model does not completely correlate with the data.

This work demonstrates that LQM regularly delivers correct parameter estimates under varied settings, retaining resilience across varying parameter values and distributions. For lesser parameter values (e.g., θ=0.25), LQM surpasses MLE in terms of MSE. Although the disparity narrows as θ grows, LQM still has an edge in MSE.

Specific Distributions and Computational Complexity: It is important to highlight that these conclusions only apply to the specific distributions and parameter values studied. MLE may outperform LQM under different conditions, necessitating more research into a larger range of distributions. Furthermore, because LQM is a relatively new approach, it may have greater computational costs, particularly with large datasets, which could provide issues in large-scale applications in terms of processing time and resource allocation.

5.3 REAL DATA

This section shows how the Poisson Lindley Distribution can be used in practice by fitting it to a real-world dataset. Bjerkedal initially supplied the dataset for this analysis in 1960. The dataset includes the survival durations of guinea pigs that were subjected to different tubercle bacilli dosages. It is known that guinea pigs are more vulnerable to tuberculosis than humans. Interestingly, a minimal infection with a few virulent tubercle bacilli can cause the disease to progress and eventually kill the patient. There are a total of 72 observations in Table 3, and the individual data points are listed below:

Table 3: Seventy-two Guinea Pigs under Regime 6.6: Survival Times

0.12	**0.15**	**0.22**	**0.24**	**0.24**	**0.32**	**0.32**	**0.33**
0.34	0.38	0.38	0.43	0.44	0.48	0.52	0.53
0.54	0.54	0.55	0.56	0.57	0.58	0.58	0.59
0.60	0.60	0.60	0.60	0.61	0.62	0.63	0.65
0.65	0.67	0.68	0.70	0.70	0.72	0.73	0.75
0.76	0.76	0.81	0.83	0.84	0.85	0.87	0.91
0.95	0.96	0.98	0.99	1.09	1.10	1.21	1.27
1.29	1.31	1.43	1.46	1.46	1.75	1.75	2.11
2.33	2.58	2.58	2.63	2.97	3.41	3.41	3.76

A descriptive summary of these data can be found in the following table 4.

Table 4: A few properties of the data collection.

mean	variance	**Skewness**	**Kurtosis**
0.9981944	0.6580122	1.758953	2.459565

Table 5 presents variance estimates for the unknown parameters and the associated Kolmogorov–Smirnov (K–S) statistics, along with their corresponding p-values for the Poisson-Lindley Distribution (PLD). An observation from Table 5 reveals that the small K–S distance and the substantial p-value for the test collectively suggest a strong fit of the LQM estimates to the Poisson-Lindley Distribution.

Table 5: Outcomes and Parameter Estimates with Goodness-of-Fit Test P-Values (P-values shown in parentheses) for Real Data

methods	**estimate**	**Statistics**				
	$\hat{\theta}$	MSE	AIC	BIC	Kolmogorov Smirnov	χ^2
MLE	0.7806483	78.82717	297.4119	303.9652	0.20167 (0.005723)	398.28 (2.2e-16)
LQM	4.190217	78.71956	105.223	111.7763	0.076257 (0.7965)	2.0845 (1)

Based on a combined evaluation of Mean Squared Error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), LQM is suggested as one of these approaches for this dataset. When it comes to parameter estimation for the Poisson Lindley distribution, LQM performs better. The data presented in Table 5 and Figure 8 provide substantial support for this inference, demonstrating that the LQM approach provides the most accurate fit to the dataset.

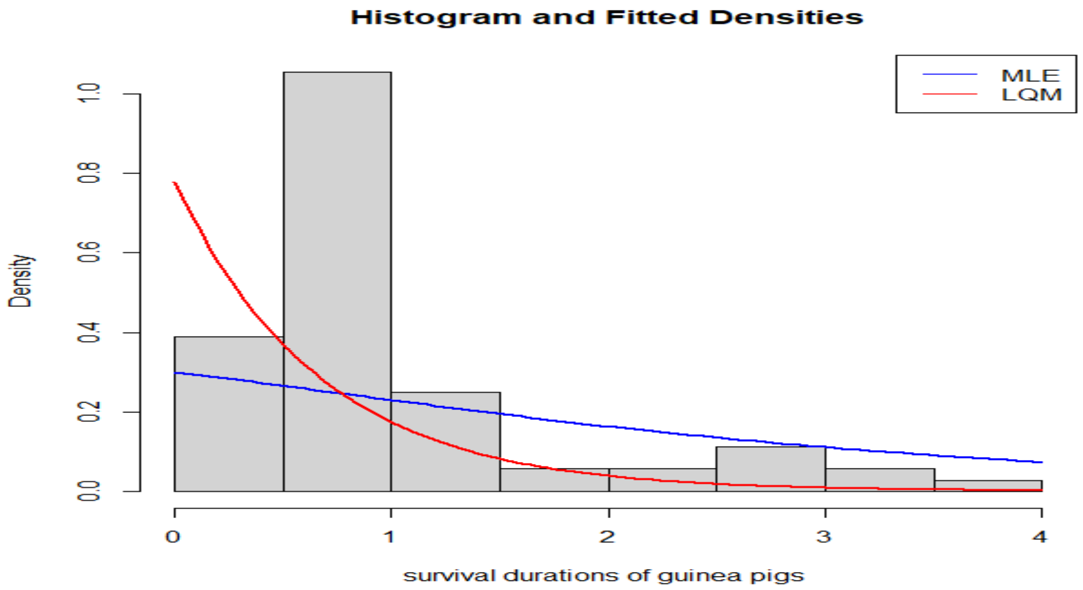

FIGURE 8: LQM estimate approach offers the best fit to the dataset

6. CONCLUSION

In conclusion, this study on Poisson Lindley Distribution parameter estimation techniques emphasises the superiority of the Linear Quantile Moment (LQM) approach and the significance of reliability analysis in assessing their dependability. In terms of accuracy and model fit, LQM consistently outperformed Maximum Likelihood Estimation (MLE) and a combined method, as demonstrated by extensive simulations and real guinea pig survival data analysis. These results highlight the effectiveness of LQM for this distribution and provide insightful information for applications in epidemiology, finance, and environmental studies. The work advances statistical methods for intricate processes and highlights the usefulness of using LQM for more accurate parameter estimation.

Funding

There was no outside funding for this study

ACKNOWLEDGEMENT

The authors have no acknowledgements to declare

CONFLICTS OF INTEREST

According to the writers, they have no conflicts of interest.

REFERENCES

[1] I. Elbatal, M. S. Rahman, and F. M. Al-Marzouki, "The generalized Lindley distribution: Properties and applications," Commun. Stat. - Theory Methods, vol. 42, no. 12, pp. 2246–2262, 2013.
[2] A. A. Al-Babtain, A. M. Gemeay, and A. Z. Afify, "Estimation methods for the discrete Poisson-Lindley and discrete Lindley distributions with actuarial measures and applications in medicine," J. King Saud Univ.-Sci., vol. 33, no. 2, p. 101224, 2021.
[3] F. M. Alghamdi, M. Ahsan-ul-Haq, M. N. S. Hussain, E. Hussam, E. M. Almetwally, H. M. Aljohani, M. S. Mustafa, E. Alshawarbeh, and M. Yusuf, "Discrete Poisson Quasi-X-Lindley distribution with mathematical properties, regression model, and data analysis," J. Radiat. Res. Appl. Sci., vol. 17, no. 2, p. 100874, 2024.
[4] T. T. Thach and R. Briš, "MLE versus MCMC estimators of the mixture of failure rate model," in Safety and Reliability – Safe Societies in a Changing World, CRC Press, pp. 937–944, 2018.
[5] M. Holla, "On a Poisson-inverse Gaussian distribution," Metrika, vol. 11, no. 1, pp. 115–121, 1967.
[6] M. Sankaran, "The discrete Poisson-Lindley distribution," Int. Biometric Soc., vol. 26, no. 1, pp. 145–149, 1970.
[7] E. Mahmoudi and H. Zakerzadeh, "Generalized Poisson-Lindley distribution," Commun. Stat. - Theory Methods, vol. 39, no. 10, pp. 1785–1798, 2010.
[8] H. Zamani, N. Ismail, and P. Faroughi, "Poisson-weighted exponential univariate version and regression model with applications," J. Math. Stat., vol. 10, no. 2, pp. 148–154, 2014.
[9] S. A. Othman, "Enhancing reliability engineering through Weibull distribution in R," Al-Rafidain J. Comput. Sci. Math., vol. 18, no. 1, 2024.

[10] 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.
[11] C. Chesneau, L. Tomy, and G. Veena, "The Poisson-Modified Lindley Distribution," Appl. Math. E-Notes, vol. 22, pp. 18–31, 2022.
[12] W. A. H. Al-Nuaami, A. A. Heydari, and H. J. Khamnei, "The Poisson–Lindley distribution: Some characteristics, with its application to SPC," Mathematics, vol. 11, no. 11, p. 2428, 2023.
[13] D. V. Lindley, "Fiducial distributions and Bayes' theorem," J. R. Stat. Soc., Ser. B (Methodological), pp. 102–107, 1958.
[14] M. Ghitany and D. Al-Mutairi, "Estimation methods for the discrete Poisson–Lindley distribution," J. Stat. Comput. Simul., vol. 79, pp. 1–9, 2009.
[15] E. Altun, "A new model for over-dispersed count data: Poisson quasi-Lindley regression model," Math. Sci., vol. 13, pp. 241–247, 2019.
[16] D. Karlis and E. Xekalaki, "Mixed Poisson distributions," Int. Stat. Rev., vol. 73, no. 1, pp. 35–58, 2005.
[17] 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.
[18] S. A. Dar, A. Hassan, and P. B. Ahmad, "Poisson size-biased Lindley distribution and its applications," Int. J. Model. Simul. Sci. Comput., vol. 13, no. 4, p. 2250031, 2022.
[19] M. Abdi, A. Asgharzadeh, H. S. Bakouch, and Z. Alipour, "A new compound gamma and Lindley distribution with application to failure data," Austrian J. Stat., vol. 48, no. 3, pp. 54–75, 2019.
[20] B. O. Oluyede, B. Mashabe, A. Fagbamigbe, B. Makubate, and D. Wanduku, "The exponentiated generalized power series: Family of distributions: Theory, properties and applications," Heliyon, vol. 6, no. 8, 2020.