axioms

Article

A Flexible Dispersed Count Model Based on Bernoulli Poisson–Lindley Convolution and Its Regression Model

Hassan S. Bakouch [1,2], Christophe Chesneau [3], Radhakumari Maya [4], Muhammed Rasheed Irshad [5], Sreedeviamma Aswathy [5] and Najla Qarmalah [6,*]

1 Department of Mathematics, College of Science, Qassim University, Buraydah 51452, Saudi Arabia; h.bakouch@qu.edu.sa or hassan.bakouch@science.tanta.edu.eg
2 Department of Mathematics, Faculty of Science, Tanta University, Tanta 31111, Egypt
3 Department of Mathematics, University of Caen, 14032 Caen, France
4 Department of Statistics, University College, Thiruvananthapuram 695034, Kerala, India; publicationsofmaya@gmail.com
5 Department of Statistics, Cochin University of Science and Technology, Cochin 682022, Kerala, India; irshadmr@cusat.ac.in (M.R.I.)
6 Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
* Correspondence: nmbinqurmalah@pnu.edu.sa

Abstract: Count data are encountered in real-life dealings. More understanding of such data and the extraction of important information about the data require some statistical analysis or modeling. One innovative technique to increase the modeling flexibility of well-known distributions is to use the convolution of random variables. This study examines the distribution that results from adding two independent random variables, one with the Bernoulli distribution and the other with the Poisson–Lindley distribution. The considered distribution is named as the two-parameter Bernoulli–Poisson–Lindley distribution. Many of its statistical properties are investigated, such as moments, survival and hazard rate functions, mode, dispersion behavior, mean deviation about the mean, and parameter inference based on the maximum likelihood method. To evaluate the effectiveness of the bias and mean square error of the produced estimates, a simulation exercise is carried out. Then, applications to two practical data sets are given. Finally, we construct a flexible count data regression model based on the proposed distribution with two practical examples.

Keywords: discrete statistical model; dispersion index; hazard rate function; parameter estimation; simulation; regression

MSC: 62E15

Citation: Bakouch, H.S.; Chesneau, C.; Maya, R.; Irshad, M.R.; Aswathy, S.; Qarmalah, N. A Flexible Dispersed Count Model Based on Bernoulli Poisson–Lindley Convolution and Its Regression Model. *Axioms* **2023**, *12*, 813. https://doi.org/10.3390/axioms12090813

Academic Editors: Nuno Bastos, Touria Karite and Amir Khan

Received: 3 July 2023
Revised: 17 August 2023
Accepted: 22 August 2023
Published: 24 August 2023

1. Introduction

In recent decades, count data analysis has drawn interest. There are many count data sets in practical as well as theoretical domains, including medicine, sports, engineering, finance, insurance, etc. (see [1]). However, we are unable to use methodologies or typical standard probability distributions to analyze them. Building adaptable models has attracted a lot of interest from statisticians and applied scientists in order to improve the modeling of count data. Therefore, it is critical to create models that are superior to standard distributions in order to successfully investigate real-world data and its attributes.

Recently, for the purpose of modeling count data, several models have evolved. The use of conventional discrete distributions as models for dependability, hazard rates, counts, etc., is limited. The widespread parametric models for analyzing such data are the Poisson, geometric, and negative binomial (NB) models (see [2]). The Poisson regression model is the most common model for modeling count data, but an obstacle arises: there is a fact that they may exhibit over- or under-dispersion, which is when a count's conditional variance is greater or less than its conditional mean (see [3]). In these cases, the Poisson model's mean–variance relationship is a well-known drawback. This has led to the introduction of various Poisson distribution types (see [4,5]). A traditional way of overcoming over-dispersion is to allow the single parameter of the Poisson distribution to be a random variable following a given distribution. This is also known as the compounding method, and the idea was first proposed in [6]. The resultant compound distributions are also termed as mixture distributions. One such famous mixture distribution is the negative binomial distribution, obtained by mixing the Poisson distribution with a gamma distribution. In real-world count modeling applications, the negative binomial distribution with an additional dispersion parameter is widely accepted as a solution to the over-dispersion issue.

As a result, various discrete distributions based on widely used continuous distributions for reliability, hazard rates, etc., have been developed. The discrete Weibull distribution is the most well-liked of these. It was introduced in [7–9]. Since then, numerous applications have been made. There are many other recently constructed distributions with continuous analogues. The author in [10] introduced the discrete gamma distribution, which has received significant attention for applications in the areas of molecular biology and evolution. Discrete analogues of the continuous Burr and Pareto distributions were constructed in [11]. On the other hand, the authors in [12] introduced a discrete analogue of the continuous inverse Weibull distribution. The discrete Lindley distribution was proposed in [13].

There are so many models for studying over-dispersion, while only a few models are there to deal with under-dispersion, because over-dispersion exists more frequently (see [14]).

Various extensions and generalizations of the Poisson distributions were developed for both over-dispersed and under-dispersed count data in the literature over the last decade. The authors in [15] proposed the generalized Poisson (GP) regression model, whereas those of [16] introduced the Conway–Maxwell–Poisson (COM–Poisson) model. The COM–Poisson regression model was also created. The authors in [17] invented the Poisson–Tweedie regression model.

Each of the aforementioned models has some drawbacks. For instance, the GP model's range must be truncated in order to achieve under-dispersion, with the level of truncation depending on the actual model parameters. The issue is that because of the range's shortening, the probabilities no longer add up to 1. The convolutions (sum and difference) of two independent random variables are a clever way of broadening the modeling possibilities of well-known distributions.

The author in [18] proposed the discrete Poisson–Lindley distribution, a compound Poisson distribution obtained by compounding the Poisson distribution with the Lindley distribution. The authors in [19] introduced an efficient regression model for under-dispersed count data based on the Bernoulli–Poisson convolution (BerPoi) for under-dispersed count data. In it, the response variable is distributed according to the BerPoi distribution using a specific parameterization indexed by mean and dispersion parameters.

In this paper, we introduce a distribution generated from the sum of two independent random variables, one with the Bernoulli distribution and the other with the Poisson–Lindley distribution. The resulting distribution is known as the Bernoulli–Poisson–Lindley (BPL) distribution. One of its key advantages is that it is suitable for modeling both under-dispersed and over-dispersed count data, unlike the Poisson distribution. Furthermore, it has only two parameters, which reduces the complexity of the simulation study, unlike some Poisson generalizations with three parameters. Moreover, it has an increasing hazard rate, making it appropriate for modeling equipment wear and tear or ageing processes. The proposed model is appropriate for regression modeling since its moments may be retrieved in closed form.

The remaining sections of the paper are organized as follows: Section 2 presents the BPL distribution. Section 3 discusses the statistical properties of this distribution.

Section 4 introduces the parameter estimation using the maximum likelihood method, and its performance is assessed via a simulation study. The new model is shown to perform at least as well as other recently proposed two-parameter discrete models, and the conventional one-parameter discrete models using two real data sets are analyzed in Section 5. In Section 6, a regression model is developed. Finally, several key takeaways are outlined in Section 7.

2. Bernoulli-Poisson-Lindley Distribution

The BPL distribution is obtained by the distribution of the sum of two independent random variables, one with the Bernoulli distribution, and the other with the Poisson–Lindley distribution.

The result below presents a simple expression of the corresponding probability mass function (pmf).

Proposition 1. *The pmf of the BPL distribution with parameters α and θ can be expressed as*

$$p(x,\alpha,\theta) = \begin{cases} \dfrac{(1-\alpha)\theta^2(\theta+2)}{(\theta+1)^3} & if \;\; x = 0 \\ \dfrac{\theta^2\left[(1+\alpha\theta)(x+\theta+1)+(1-\alpha)\right]}{(\theta+1)^{x+3}} & if \;\; x = 1,2,3,\ldots \end{cases} \tag{1}$$

Proof. Let X_1 and X_2 be two independent random variables, with X_1 following the Bernoulli distribution with parameter $0 < \alpha < 1$, i.e., $P(X_1 = 0) = 1-\alpha$ and $P(X_1 = 1) = \alpha$ and X_2 following the Poisson–Lindley distribution with parameter $\theta > 0$, i.e., $P(X_2 = x) = \dfrac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}$ with $x = 0,1,2,3,\ldots$ Then, by the definition, the BPL distribution is the distribution of $X = X_1 + X_2$. Let us now determine its pmf. For any $x = 0,1,\ldots$, we have

$$\begin{aligned} p(x,\alpha,\theta) &= P(X = x) = P(X_1 + X_2 = x) \\ &= P(X_1 = 0)P(X_2 = x) + P(X_1 = 1)P(X_2 = x-1). \end{aligned}$$

In particular, for $x = 0$, we have

$$p(x,\alpha,\theta) = P(X_1 = 0)P(X_2 = 0) = \frac{(1-\alpha)\theta^2(\theta+2)}{(\theta+1)^3}.$$

For $x = 1,2,\ldots$, we have

$$\begin{aligned} p(x,\alpha,\theta) &= P(X = x) \\ &= (1-\alpha)\frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}} + \alpha\frac{\theta^2(x-1+\theta+2)}{(\theta+1)^{x-1+3}} \\ &= \frac{\theta^2}{(\theta+1)^{x+3}}\left[\alpha(x+\theta+1)(\theta+1) + (1-\alpha)(x+2+\theta)\right] \\ &= \frac{\theta^2}{(\theta+1)^{x+3}}\left[\alpha\theta(x+\theta+1) + \alpha(x+\theta+1) + (1-\alpha)(x+\theta+1+1)\right] \\ &= \frac{\theta^2}{(\theta+1)^{x+3}}\left[\alpha\theta(x+\theta+1) + \alpha(x+\theta+1) + (1-\alpha)(x+\theta+1) + (1-\alpha)\right] \\ &= \frac{\theta^2}{(\theta+1)^{x+3}}\left[(1+\alpha\theta)(x+\theta+1) + (1-\alpha)\right]. \end{aligned}$$

This ends the proof of Proposition 1. □

Remark 1. *When $\alpha \to 0$, the Poisson–Lindley distribution is included in the BPL distribution as a special case.*

Proposition 2. *The cumulative density function (cdf) of the BPL distribution can be expressed as, for any integer x,*

$$F(x,\alpha,\theta) = 1 + \frac{[-1-\theta(3+x+\theta+x\alpha\theta+\alpha\theta(2+\theta))]}{(1+\theta)^{x+3}}, \quad x = 0,1,2,\dots \tag{2}$$

Proof. It follows from the geometric series expansions and some algebra, that

$$\begin{aligned} F(x,\alpha,\theta) &= \sum_{k=0}^{x} p(k,\alpha,\theta) \\ &= \frac{\theta^2(1-\alpha)(\theta+2)}{(\theta+1)^3} + \sum_{k=1}^{x} \frac{\theta^2\Big[[(1+\alpha\theta)(k+\theta+1)] + (1-\alpha)\Big]}{(\theta+1)^{k+3}} \\ &= 1 + \frac{[-1-\theta(3+x+\theta+x\alpha\theta+\alpha\theta(2+\theta))]}{(1+\theta)^{x+3}}. \end{aligned}$$

This ends the proof of Proposition 2. □

The corresponding survival function is given by

$$S(x,\alpha,\theta) = \frac{1+\theta[3+x+\theta+x\alpha\theta+\alpha\theta(2+\theta)]}{(1+\theta)^{x+3}}, \quad x = 0,1,2,\dots \tag{3}$$

The hazard rate function (hrf) of the BPL distribution is obtained as

$$h(x,\alpha,\theta) = \begin{cases} \dfrac{(1-\alpha)\theta^2(\theta+2)}{1+\theta[3+\theta+\alpha\theta(2+\theta)]} & if \;\; x = 0 \\ \dfrac{\theta^2[1-\alpha+(1+x+\theta)(1+\alpha\theta)]}{1+\theta[3+x+\theta+x\alpha\theta+\alpha\theta(2+\theta)]} & if \;\; x = 1,2,3,\dots \end{cases} \tag{4}$$

Figure 1 shows the different shapes of the pmf. It clearly indicates that the BPL distribution is positively skewed, unimodal and as θ goes larger, the mass concentrates more on values closer to 0 than at higher values. Figure 2 also presents different shapes of the cdf.

Figure 3 presents different shapes of the hrf, indicating that the BPL distribution exhibits increasing hazard rates with respect to both α and θ.

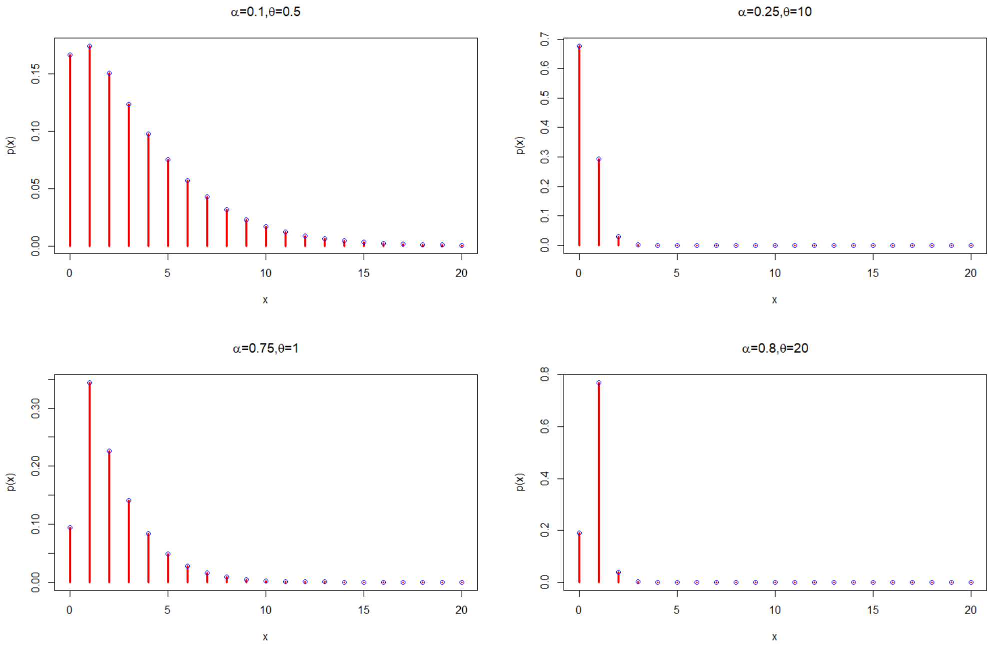

Figure 1. Pmfs of the BPL distribution for different values of the parameters.

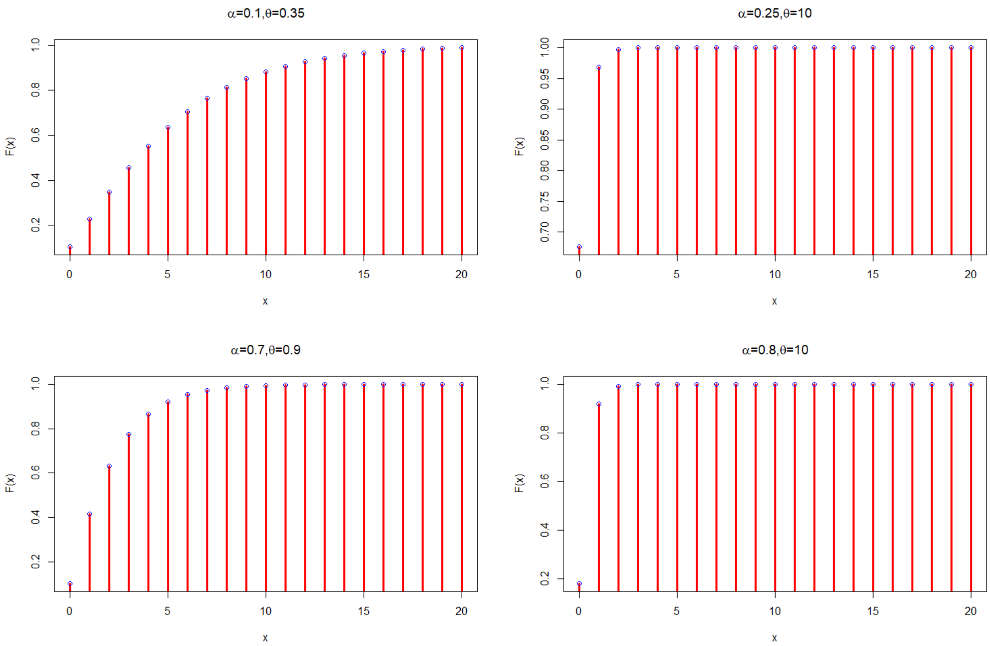

Figure 2. Cdfs of the BPL distribution for different values of the parameters.

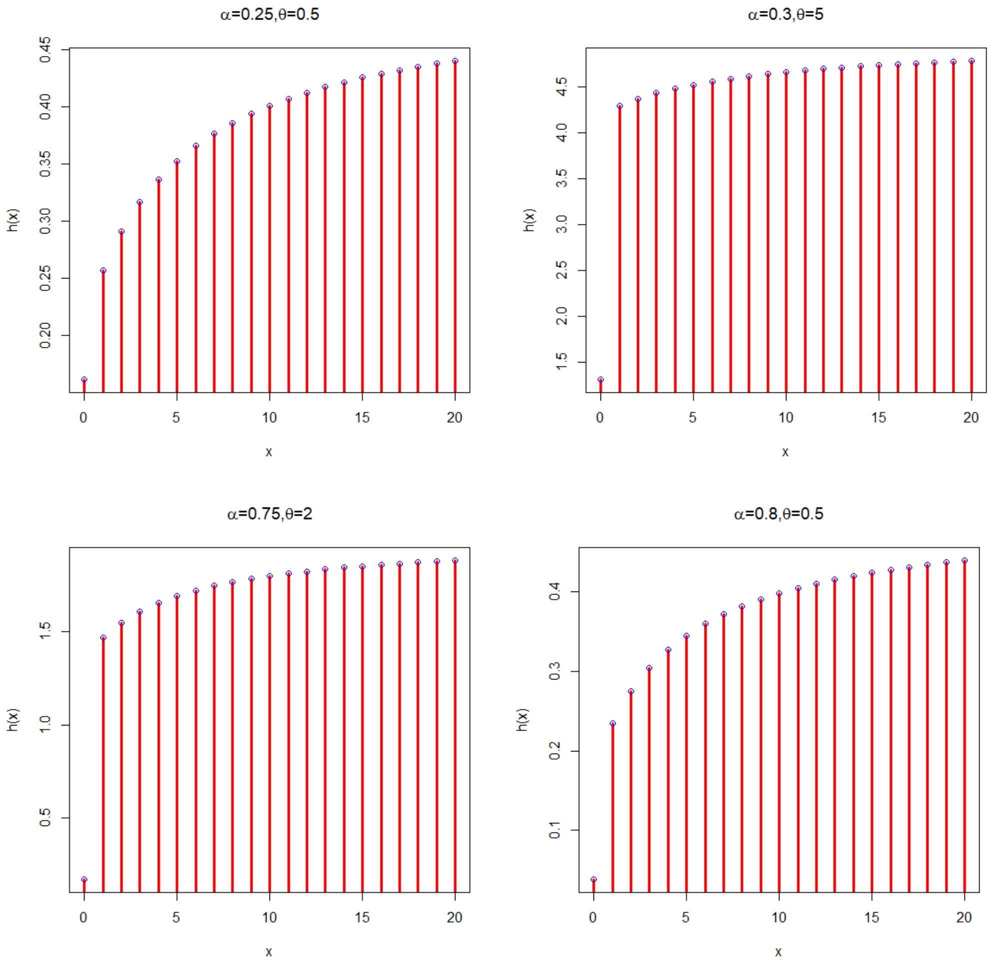

Figure 3. Hrfs of the BPL distribution for different values of the parameters.

3. Statistical Properties

3.1. Mode

We now provide some theory to the observation of the mode of the BPL distribution made in Figure 1.

Proposition 3. *Let X be a random variable following the BPL distribution. Then, the mode of X, denoted by* x_m*, exists in {0,1, 2, . . . }, and satisfies*

$$-1+\frac{1}{\theta}-\theta+\frac{2+\alpha}{1+\alpha\theta}\leq x_m\leq\frac{1}{\theta}-\theta+\frac{\alpha-1}{1+\alpha\theta}, \tag{5}$$

with $x_m = 0$ *if the upper bound is non-positive.*

Proof. By the definition of the mode, it corresponds to the integer $x = x_m$ for which $p(x, \alpha, \theta)$ has the greatest value, where we recall that

$$p(x,\alpha,\theta)=\begin{cases}(1-\alpha)\theta^2\dfrac{(\theta+2)}{(\theta+1)^3} & if\ \ x=0\\ \dfrac{\theta^2}{(\theta+1)^{x+3}}\left[(1+\alpha\theta)(x+\theta+1)+(1-\alpha)\right] & if\ \ x=1,2,3,\ldots\end{cases} \tag{6}$$

To reach our aim, we need to solve $p(x_m, \alpha, \theta) \geq p(x_m - 1, \alpha, \theta)$ and $p(x_m, \alpha, \theta) \geq p(x_m + 1, \alpha, \theta)$. Obviously, $p(x_m, \alpha, \theta) \geq p(x_m - 1, \alpha, \theta)$ implies that

$$x_m \leq \frac{1}{\theta} - \theta + \frac{\alpha - 1}{1 + \alpha\theta}. \tag{7}$$

Furthermore, $p(x_m, \alpha, \theta) \geq p(x_m + 1, \alpha, \theta)$ implies that

$$x_m \geq -1 + \frac{1}{\theta} - \theta + \frac{2 + \alpha}{1 + \alpha\theta}. \tag{8}$$

By combining Equations (7) and (8), we obtain Equation (5), hence, the proof of Proposition 3. □

3.2. Moments, Skewness, and Kurtosis

Hereafter, let X be a random variable following the BPL distribution. Then, after some algebraic developments, the probability generating function of X is given by

$$P(s) = E(s^X) = \frac{[1 + (-1 + s)\alpha]\theta^2(2 - s + \theta)}{(1 + \theta)(1 - s + \theta)^2},$$

for $s < \theta + 1$.

The moment-generating function of X can be obtained by replacing s by e^t, for $t < \log(\theta + 1)$, which gives

$$M(t) = E(e^{tX}) = \frac{[1 + (-1 + e^t)\alpha]\theta^2(2 - e^t + \theta)}{(1 + \theta)(1 - e^t + \theta)^2}.$$

Basically, the r-th moment about the origin of X is derived as

$$E(X^r) = \sum_{x=0}^{\infty} x^r p(x, \alpha, \theta) = \sum_{x=1}^{\infty} x^r \frac{\theta^2}{(\theta + 1)^{x+3}} \big[(1 + \alpha\theta)(x + \theta + 1) + (1 - \alpha)\big].$$

Thus, after an intense use of the geometric series formulas (see Appendix A), the first four moments of X are

$$E(X) = \alpha + \frac{2 + \theta}{\theta(\theta + 1)},$$
$$E(X^2) = \frac{6 + \theta[4 + \theta + \alpha(4 + \theta(3 + \theta))]}{\theta^2(1 + \theta)},$$
$$E(X^3) = \frac{24 + \theta\big[24 + \theta(8 + \theta) + \alpha(3 + \theta)(6 + \theta(4 + \theta))\big]}{\theta^3(1 + \theta)},$$

and

$$E(X^4) = \frac{120 + \theta\big[168 + \theta[78 + \theta(16 + \theta)] + \alpha[96 + \theta(132 + \theta[64 + \theta(15 + \theta)])]\big]}{\theta^4(1 + \theta)}.$$

Now, the variance of X is calculated as

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

Figure 4 presents the plots of the variance of X for different values of the parameters α and θ. We see that the variance decreases when α is fixed and θ increases.

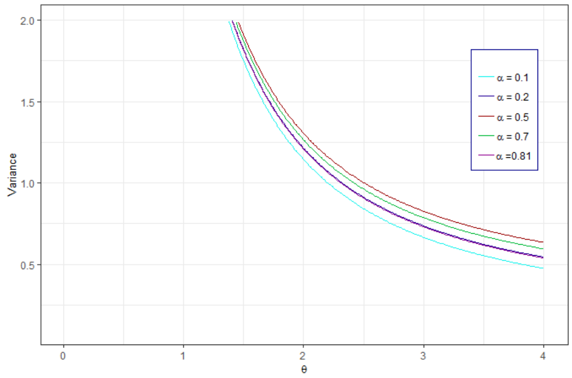

Figure 4. Variance of the BPL distribution for different values of the parameters.

On the other hand, based on the first four moments of X, the skewness of X is

$$Skewness(X) = \frac{[4+\theta(18+\theta[32+\theta(22+\alpha(1+\theta)^3-3\alpha^2(1+\theta)^3+2\alpha^3(1+\theta)^3+\theta(7+\theta))])]^2}{[2-\theta(6-\theta[4+\theta+(1-\alpha)\alpha(1+\theta^2)])]^3}.$$

Furthermore, the kurtosis of X is

$$\begin{aligned} Kurtosis(X) = {} & \frac{1}{[-2+\theta(-6+\theta[-4-\theta-(1-\alpha)\alpha(1+\theta)^2])]^2}\Big[24+\theta(144+\theta[338+ \\ & 6\alpha^3\theta^2(1+\theta)^4-3\alpha^4\theta^2(1+\theta)^4+\alpha(1+\theta)^2[12+\theta(4+\theta)(9+\theta[4+\theta])]+ \\ & \theta[406+\theta(258+\theta(87+\theta[15+\theta]))] \\ & -2\alpha^2(1+\theta)^2[6+\theta(18+\theta[14+\theta(7+2\theta)])]])\Big]. \end{aligned}$$

Figure 5 presents the plots of the skewness and kurtosis of X, respectively. From these plots, when the value of α is held constant, and θ increases, a significant effect on both the skewness and kurtosis is observed. Furthermore, when θ increases, the BPL distribution is rightly skewed and leptokurtic.

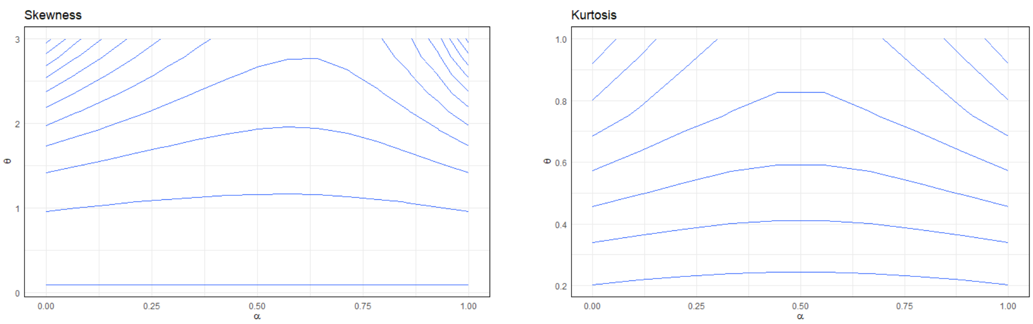

Figure 5. Skewness and kurtosis of the BPL distribution for different values of the parameters.

3.3. Dispersion Index and Coefficient of Variation

In this section, we discuss the dispersion index (DI) and coefficient of variation (CV) associated with the BPL distribution. The CV of X is obtained as

$$CV(X) = \frac{\sqrt{2+\theta[6+\theta(4+\theta+(1-\alpha)\alpha(1+\theta)^2)]}}{2+\theta+\alpha\theta(\theta+1)}.$$

The DI of X is given by

$$DI(X) = 1 + \frac{1}{\theta} + \frac{1}{1+\theta} - \left(\alpha + \frac{1-\alpha+\alpha\theta}{2+\theta(1+\alpha+\alpha\theta)}\right).$$

Clearly, $DI(X)$ is greater than 1 when θ tends to 0, and less than 1 when θ tends to ∞. Thus, the BPL distribution has under- or over-dispersed properties.

Numerical values for some moment measures, such as mean, variance, DI, skewness, and kurtosis for the BPL distribution for different sets of parameter values are given in Tables 1 and 2. It can be observed that the mean and variance decrease as θ tends to ∞ for fixed values of α.

Table 1. Numerical values for some moment measures associated with the BPL distribution for α = 0.1 and different values of θ.

Measures	θ				
	0.1	**10**	**50**	**99**	**999**
Mean	19.1909	0.2091	0.1204	0.1102	0.1010
Variance	218.3545	0.2108	0.1108	0.1003	0.0910
DI	11.3780	1.0083	0.9204	0.9102	0.9010
Skewness	2.0459	5.1086	6.4470	6.7468	7.0719
Kurtosis	6.0496	8.5888	8.2779	8.2024	8.1209

Table 2. Numerical values for some moment measures associated with the BPL distribution for α = 0.3 and different values of θ.

Measures	θ				
	0.1	**10**	**50**	**99**	**999**
Mean	19.3909	0.4091	0.3204	0.3102	0.3010
Variance	218.4745	0.3309	0.2308	0.2203	0.2110
DI	11.2669	0.8087	0.7204	0.7102	0.7010
Skewness	2.0426	1.4711	0.9079	0.8355	0.7692
Kurtosis	6.0462	4.3926	2.4144	2.1001	1.7964

3.4. Mean Deviation about the Mean

The mean deviation (MD) about the mean measures the amount of scatter in a population. Let μ be the mean of the BPL distribution, i.e., $\mu = E(X) = \alpha + \dfrac{2+\theta}{\theta(\theta+1)}$. Then the MD about the mean is defined as $MD(X) = E\big(|X-\mu|\big)$, and can be calculated as

$$\begin{aligned} MD(X) &= \sum_{x=0}^{\infty} |x-\mu| p(x,\alpha,\theta) \\ &= \mu p(0,\alpha,\theta) + \sum_{x=1}^{\lfloor\mu\rfloor} (\mu - x) p(x,\alpha,\theta) + \sum_{x=\lfloor\mu\rfloor+1}^{\infty} (x-\mu) p(x,\alpha,\theta) \\ &= \frac{(1+\theta)^{-3-\lfloor\mu\rfloor}}{\theta}\Big[2(1+\theta)^2[2+\theta(1+\alpha+\alpha\theta)] - 2\theta(1+\theta[3+\theta+\alpha\theta(2+\theta)])\mu \\ &- (1+\theta)^{2+\lfloor\mu\rfloor}[2+\theta(1+\alpha+\alpha\theta-(1+\theta)\mu)] \\ &+ 2\theta\lfloor\mu\rfloor(2+\theta[4+\alpha+\theta+\alpha\theta(3+\theta-\mu)-\mu] + \theta(1+\alpha\theta)\lfloor\mu\rfloor)\Big], \end{aligned}$$

where $\lfloor\mu\rfloor$ is the greatest integer less than or equal to μ.

Figure 6 shows the plot of the MD about the mean of X. From this plot, we observe that when θ increases, the values of the MD about the mean decrease.

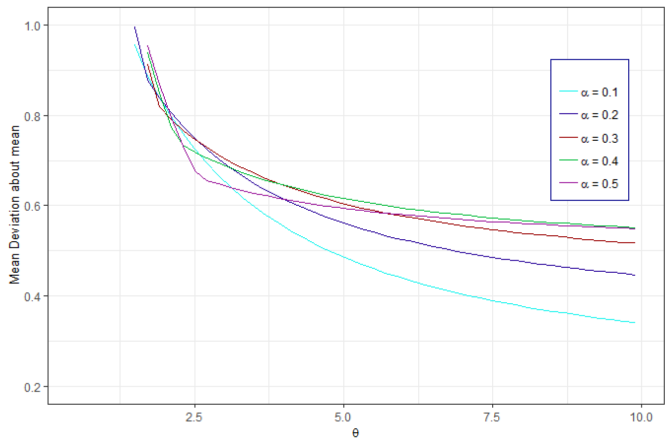

Figure 6. MD about the mean of the BPL distribution for different values of α and θ.

4. Parameter Estimation

Parameter estimation is an important step toward a deeper understanding of the process. The classical method of estimation, the maximum likelihood (ML) method, is used here to estimate the parameters. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a BPL distribution with unknown parameters α and θ. Let $x_1, \ldots, x_n$ be the n observed values. Let y be the number of x_i taking the value 0 and $(n-y)$ of x_i's are taking the nonzero values. The log-likelihood function is given by

$$\begin{aligned}\log L(\alpha,\theta) = {}& y\log(1-\alpha) + 2y\log\theta + y\log(\theta+2) - 3y\log(\theta+1) + 2(n-y)\log\theta \\ &- 3(n-y)\log(1+\theta) \\ &+ \sum_{i=1,x_i\neq 0}^{n-y}\left\{\log\left[(1+\alpha\theta)(1+\theta+x_i)+(1-\alpha)\right] - x_i\log(\theta+1)\right\}.\end{aligned}$$

The maximum likelihood estimates (MLEs) of α and θ are the values that maximize $\log L(\alpha,\theta)$. They are denoted as $\hat{\alpha}$ and $\hat{\theta}$, respectively. The partial derivatives of $\log L(\alpha,\theta)$ with respect to each parameter are the following:

$$\frac{\partial}{\partial\alpha}\log L(\alpha,\theta) = \sum_{i=1}^{n-y}\left\{\frac{\theta(1+x_i+\theta)-1}{(1+\alpha\theta)(1+x_i+\theta)+(1-\alpha)}\right\} - \frac{y}{1-\alpha},$$

$$\frac{\partial}{\partial\theta}\log L(\alpha,\theta) = \sum_{i=1}^{n-y}\left\{\frac{(1+\alpha\theta)+(1+x_i+\theta)\alpha}{(1+\alpha\theta)(1+x_i+\theta)+(1-\alpha)}\right\} - \frac{n(3+\bar{x})}{\theta+1} + \frac{y}{\theta+2} + \frac{2n}{\theta}.$$

In order to obtain the MLEs, note that the above system of equations set to zero contains non-linear equations and does not have an explicit solution. Consequently, the system must be solved numerically, for example, using the statistical programming language **R** (see Appendix A).

Simulation Study

In this section, a brief simulation study is performed to evaluate the asymptotic behavior of the MLEs for different parametric combinations. Here the iteration is carried out for different sample sizes (50, 100, 200, 500, 1000) and $N = 1000$ replications are used for the same. The measures such as percentage relative bias (PRB) and mean square errors (MSEs) are calculated with the following formulas:

$$PRB = \frac{\sum_{i=1}^{N}(a - \hat{a}_i)}{\sum_{i=1}^{N}\hat{a}_i} \times 100,$$

where $a \in \{\alpha, \theta\}$, $\hat{a}_i$ is the MLE of a at the i-th replication, and

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(a_i - \hat{a}_i)^2.$$

It is evident from Table 3 that all the estimates are asymptotically unbiased as n increases, i.e., with the PRBs approaching zero and the MSEs decreasing to zero.

Table 3. Simulation results.

			$\boldsymbol{\alpha = 0.25, \theta = 0.6}$			
n	**MLE (α)**	**PRB (α)**	**MSE (α)**	**MLE (θ)**	**PRB (θ)**	**MSE (θ)**
50	0.24715	1.15434	0.29785	0.61781	−2.88305	0.10035
100	0.24663	1.36523	0.19457	0.60301	−0.49874	0.06998
200	0.23642	3.74246	0.15031	0.60426	−0.70581	0.05007
500	0.24617	1.55751	0.08833	0.60124	−0.20587	0.03022
1000	0.25123	−0.88448	0.06078	0.60058	−0.09602	0.02079
			$\boldsymbol{\alpha = 0.5, \theta = 1.2}$			
n	**MLE (α)**	**PRB (α)**	**MSE (α)**	**MLE (θ)**	**PRB (θ)**	**MSE (θ)**
50	0.49695	0.61431	0.15829	1.24485	−3.60276	0.24016
100	0.50188	−0.37455	0.10670	1.22124	−1.73911	0.16789
200	0.49925	0.15014	0.07770	1.21047	−0.86520	0.11252
500	0.50077	−0.15318	0.04811	1.20429	−0.35658	0.06926
1000	0.50027	−0.05312	0.03408	1.20472	−0.39213	0.04991
			$\boldsymbol{\alpha = 0.65, \theta = 3}$			
n	**MLE (α)**	**PRB (α)**	**MSE (α)**	**MLE (θ)**	**PRB (θ)**	**MSE (θ)**
50	0.64882	0.18225	0.02067	3.26433	−8.09744	1.14048
100	0.65254	−0.38979	0.06712	3.10000	−3.22588	0.60840
200	0.64524	0.73814	0.04595	3.03897	−1.28222	0.41492
500	0.65194	−0.29778	0.09402	3.03066	−1.01156	0.26135
1000	0.65068	−0.10485	0.02939	3.00499	−0.16592	0.17036

5. Empirical Studies

This section describes a comparison of the BPL model with other competing models given in Table 4, to demonstrate the BPL model's practical effectiveness. Two practical data sets are considered. The comparison of the fitted models is based on conventional metrics: the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the Kolmogorov–Smirnov test (KS) and the resulting p-value. In particular, the formulas for the AIC and BIC are

$$AIC = -2\log L + 2r$$

and

$$BIC = -2\log L + r\log n,$$

respectively, where $\log L$ is the estimation of the log-likelihood function and r is the number of parameters.

The pmfs of the competing models are given as follows:

- For the DG model:

$$p(x, \beta, \gamma) = e^{-\beta\gamma^{x+1}} - e^{-\beta\gamma^{x}},\ x = 0, 1, 2, \ldots, \beta > 0, 0 < \gamma < 1.$$

- For the DIW model:

$$p(x,\beta,\gamma) = \begin{cases} \beta & if \;\; x = 1 \\ \beta^{x^{-\gamma}} - \beta^{(x-1)^{-\gamma}} & if \;\; x = 2,3,4,\ldots, 0 < \beta < 1, \gamma > 0. \end{cases}$$

- For the PQX model:

$$p(x,\beta,\gamma) = \frac{2\beta\gamma(\gamma+1)^2 + \gamma^3(x+1)(x+2)}{2(\beta+1)(\gamma+1)^{x+3}}, \; x = 0,1,2,\ldots, \beta > 0, \gamma > 0.$$

Table 4. Discrete competitive models.

Distribution	Abbreviation	Reference
Discrete Gumbel	DG	[20]
Discrete inverse Weibull	DIW	[12]
Poisson-quasi-xgamma	PQX	[21]
Poisson	-	-
Geometric	-	-

5.1. Survival Times

The first data set consists of survival times in days for 72 guinea pigs. These data are taken from [22]. The flexibility of the BPL model is compared with other discrete flexible models, such as the DG, DIW, PQX, Poisson, and geometric models. The results of the fitted models along with their estimates together with the standard errors (SEs) are given in Table 5. This table demonstrates that the Poisson and geometric models, two of the researched models, may not be fitted to the relevant data set (based on their p-values), but we nevertheless use them for comparison since they are very common models to take into account. The BPL model, as can be observed, offers the highest p-value and the smallest AIC, BIC, and KS statistic values.

Table 5. AIC, BIC and p-values values for the survival times data.

Model	Parameters	Estimates (SE)	AIC	BIC	KS Value	p-Value
BPL	α	0.9900 (2.9821)	793.0159	797.5692	0.1299	0.176
	θ	0.0200 (0.0013)				
DG	β	4.2894 (0.7061)	800.2187	804.7720	0.14825	0.08443
	γ	0.9789 (0.0021)				
DIW	β	1.517024×10^{-41} (1.1371)	801.8879	806.4412	0.14357	0.1028
	γ	1.1214 (0.4120)				
Poisson	β	99.8194 (1.1774)	795.1784	797.9551	0.5697	2.2×10^{-16}
Geometric	β	0.0100 (0.0012)	808.1606	810.4372	0.2232	0.0015
PQL	β	1.527183×10^{-7} (0.0779)	798.0983	802.6516	0.1768	0.0222
	γ	3.005888×10^{-2} (0.0025)				

5.2. Final Examination Marks

The results of 48 slow space students' final mathematics exams from the Indian Institute of Technology in Kanpur in 2003 are included in the second data set (see [23]). The results of the fitted models given in Table 6.

The BPL model has the largest p-value, the smallest KS value, and the smallest AIC and BIC values, as seen in Tables 5 and 6. We can therefore conclude that the BPL model outperforms all other competitive models for the two real-life data sets.

Table 6. AIC, BIC and p-values values for the final examination marks.

Model	Parameters	Estimates (SE)	AIC	BIC	KS Value	p-Value
BPL	α	0.9950 (4.7501)	399.4703	403.2127	0.0976	0.7507
	θ	0.0774 (0.0114)				
DG	β	4.4664 (0.8884)	402.6350	406.3774	0.0987	0.7375
	γ	0.9224 (0.0089)				
DIW	β	2.750165×10^{-15} (0.4321)	406.3307	410.0731	0.1552	0.1978
	γ	1.3479 (0.5324)				
Poisson	β	25.8958 (0.7345)	795.1784	797.0496	0.3998	4.342×10^{-7}
Geometric	β	0.0386 (0.0055)	408.5140	410.3852	0.2501	0.0049
PQX	β	1.07574×10^{-8} (0.2323)	399.9926	403.7350	0.1093	0.6149
	γ	1.158624×10^{-1} (0.0183)				

6. Bernoulli–Poisson–Lindley Regression Model

We already mentioned that the BPL distribution is capable of modeling under-dispersed as well as over-dispersed data sets. However, over-dispersed data sets are of utmost significance. In order to describe such data sets, this section introduces a count regression model based on the BPL distribution.

6.1. Model Construction

Let Y be a random variable with the BPL distribution that indicates how many times an event has been counted.

Consider the following reparametrization:

$$\theta = \frac{\alpha + 1 - \mu + \sqrt{(\mu - \alpha - 1)^2 + 8(\mu - \alpha)}}{2(\mu - \alpha)}.$$

Then the pmf of the BPL distribution can be expressed in terms of the mean $E(Y) = \mu > 0$ as

$$P(y, \alpha, \mu) = \begin{cases} (1-\alpha)\left(\frac{\alpha+1-\mu+\sqrt{(\mu-\alpha-1)^2+8(\mu-\alpha)}}{2(\mu-\alpha)}\right)^2 \dfrac{\left(\frac{\alpha+1-\mu+\sqrt{(\mu-\alpha-1)^2+8(\mu-\alpha)}}{2(\mu-\alpha)}+2\right)}{\left(\frac{\alpha+1-\mu+\sqrt{(\mu-\alpha-1)^2+8(\mu-\alpha)}}{2(\mu-\alpha)}+1\right)^3}, & if \quad y = 0 \\ \dfrac{\left(\frac{\alpha+1-\mu+\sqrt{(\mu-\alpha-1)^2+8(\mu-\alpha)}}{2(\mu-\alpha)}\right)^2}{\left(\frac{\alpha+1-\mu+\sqrt{(\mu-\alpha-1)^2+8(\mu-\alpha)}}{2(\mu-\alpha)}+1\right)^{y+3}} \left(\left[(1+\alpha\frac{\alpha+1-\mu+\sqrt{(\mu-\alpha-1)^2+8(\mu-\alpha)}}{2(\mu-\alpha)}) [y+\frac{\alpha+1-\mu+\sqrt{(\mu-\alpha-1)^2+8(\mu-\alpha)}}{2(\mu-\alpha)}+1]\right] + (1-\alpha)\right), & if \quad y = 1,2,3,\ldots \end{cases} \tag{9}$$

with $0 < \alpha < 1$, $\mu > 0$ and $\mu - \alpha > 0$.

Assume that we have n observations of the response variable Y, which is also the response variable, with the i-th observation being a realization of a random variable Y_i for $i = 1, 2, \ldots, n$. In addition, assume that the mean of the response variable Y_i is linked to the covariates with a log link function given by

$$\mu_i = e^{x_i^T \gamma}, \quad i = 1, 2, \ldots, n \tag{10}$$

where $x_i^T = (1, x_{i1}, x_{i2}, x_{i3}, \ldots, x_{ik})^T$ is the covariate vector and $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_k)$ is the unknown regression coefficient vector. Substituting Equation (10) in Equation (9), a linear form for the pmf of Y_i provided that $\{X_i^T = x_i^T\}$ is realized and the BPL distribution with parameters α and μ_i, is obtained as

$$P(y_i, \alpha, e^{x_i^T\gamma}) = \begin{cases} (1-\alpha)\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)^2 \frac{\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}+2\right)}{\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}+1\right)^3}, & if \;\; y_i = 0 \\ \frac{\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)^2}{\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}+1\right)^{y_i+3}} \left((1+\alpha\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}) (y_i+\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}+1)+(1-\alpha)\right), & if \;\; y_i = 1,2,3,\ldots \end{cases}$$

6.2. Estimation of the Model Parameters

The ML method is used to estimate the parameter α and the regression coefficient vector γ of the model. The logarithm of the likelihood function L of the BPL count regression model is given by

$$\begin{aligned} \log L = \sum_{i=1}^{y} \Bigg\{ & \log(1-\alpha) + 2\log\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)^2 + \\ & \log\left(\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)^2+2\right) - \\ & 3\log\left(\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)^2+1\right)\Bigg\} + \\ & \sum_{i=1,x_i\neq 0}^{n-y} \Bigg\{ 2\log\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)^2 + \\ & \log\left(\left(1+\alpha\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)\right)\right. \\ & \left(y_i+\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)+1\right) + \\ & \left.(1-\alpha)\right) - (y_i+3)\log\left(\left(\frac{\alpha+1-e^{x_i^T\gamma}+\sqrt{(e^{x_i^T\gamma}-\alpha-1)^2+8(e^{x_i^T\gamma}-\alpha)}}{2(e^{x_i^T\gamma}-\alpha)}\right)+1\right)\Bigg\}. \end{aligned} \tag{11}$$

Now the unknown parameters α and γ are obtained by maximizing Equation (11).

6.3. Residual Analysis

This part introduces a residual to test the goodness-of-fit of the BPL model defined in Section 6.1 based on randomized quantile (RQ) residuals. Let $F(y, \mu)$ be the cdf of the BPL model in which the regression structures are assumed in the parameter as in Equation (10). The i-th RQ residual of the BPL regression model is

$$r_i^q = \Phi^{-1}(F(U_i, \hat{\mu}_i)), \ \ i = 1, 2, \ldots, n,$$

where $\hat{\mu}_i = e^{x_i^T \hat{\gamma}}$, and $\Phi^{-1}(\cdot)$ represents the quantile function of the standard normal distribution. Furthermore, U_i is a random variable that follows the uniform $U\left(F(y_i - 1, \hat{\mu}_i), F(y_i, \hat{\mu}_i)\right)$ distribution. When the fitted model is correct, the RQ residuals are normally distributed with zero mean and unit variance.

6.4. Simulation of the Bernoulli–Poisson–Lindley Regression Model

This section provides a simulation exercise to assess how well the MLEs of the BPL regression model's parameters performed. We generate $N = 1000$ samples of sizes n = 100, 200, 300, and 500 for the parametric combinations (α = 0.25, γ_0 = 0.5, γ_1 = 0.4, γ_2 = 0.6) and (α = 0.5, γ_0 = 0.3, γ_1 = 1.2, γ_2 = 2) by using $\mu_i = \exp(\gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2})$. The independent variables x_{i1} and x_{i2} are generated from the standard uniform distribution, i.e., $U(0, 1)$. On the basis of the estimates, biases, and MSEs, the simulation findings are discussed. The simulation results are listed in Table 7.

Table 7. Simulation results for the BPL regression model.

	$\alpha = 0.25, \gamma_0 = 0.5, \gamma_1 = 0.4, \gamma_2 = 0.6$					$\alpha = 0.5, \gamma_0 = 0.3, \gamma_1 = 1.2, \gamma_2 = 2$			
n	**Parameters**	**Estimates**	**Bias**	**MSE**	**n**	**Parameters**	**Estimates**	**Bias**	**MSE**
100	α	0.25781	0.00781	0.01867	100	α	0.51368	0.01368	0.01360
	γ_0	0.53025	0.03025	0.49531		γ_0	0.37353	0.07353	0.16408
	γ_1	0.49863	0.09863	0.26276		γ_1	1.19985	0.00015	0.37260
	γ_2	0.65218	0.05218	0.31935		γ_2	1.80780	0.19220	1.21552
200	α	0.25420	0.00420	0.00987	200	α	0.50673	0.00673	0.00525
	γ_0	0.53000	0.03000	0.55058		γ_0	0.35115	0.05115	0.11311
	γ_1	0.47112	0.07112	0.20901		γ_1	1.18296	0.01705	0.74723
	γ_2	0.63384	0.03384	0.24494		γ_2	1.93278	0.06722	1.10588
300	α	0.25214	0.00214	0.00223	300	α	0.50106	0.00106	0.00370
	γ_0	0.50183	0.00183	0.38789		γ_0	0.31464	0.01464	0.08764
	γ_1	0.44939	0.04939	0.16479		γ_1	1.20512	0.00512	0.52853
	γ_2	0.61069	0.01069	0.17588		γ_2	1.93557	0.06443	0.53403
500	α	0.25051	0.00051	0.00430	500	α	0.50121	0.00121	0.00215
	γ_0	0.50031	0.00031	0.00031		γ_0	0.30628	0.00628	0.07150
	γ_1	0.40352	0.01352	0.00141		γ_1	1.20053	0.00052	0.35168
	γ_2	0.60321	0.00321	0.16040		γ_2	1.96866	0.03134	0.36140

Table 7 shows that the bias and MSEs reduce as sample size rises, indicating the consistency property of the MLEs for estimating the regression parameters.

6.5. Applications

Two data sets are used here to assess the performance of the BPL regression model. Only the Poisson distribution is considered in both scenarios for comparison.

6.5.1. Titanic Survivors Data

The first data set used is the Titanic survivors data. These data, which come from the Titanic's survival record, show the proportion of survivors among all the passengers, broken down by age, sex, and class. They are available in the **CountsEPPM** package of the statistical programming language **R**. The aim of the study is to investigate the effects of *age (adult)* (x_{1i}), *sex (male)* (x_{2i}), and *classes* (2-nd *class* and 3-rd *class*) (x_{3i} and x_{4i}) on the number of survivors (y_i).

The summary statistics for the Titanic survivors data are shown in Table 8.

Table 8. Summary statistics for the Titanic survivors data set.

Variables	Min	Max	Median
survive	1	140	14
age adult	0	0.5	1
sex male	0	0.5	1
2-nd *class*	0	0	1
3-rd *class*	0	0	1

The results of the regression analysis applied to the Titanic survivors data are given in Table 9.

Table 9. Modeling results for the Titanic survivors data set.

Covariates	Poisson		BPL	
	Estimates	*p*-Values	Estimates	*p*-Values
γ_0	2.71128	<0.001	2.25802	<0.001
γ_1	2.04421	<0.001	2.03979	<0.001
γ_2	−0.59605	<0.001	−0.37823	0.01094
γ_3	−0.52602	<0.001	0.07812	0.03181
γ_4	−0.12805	0.02179	0.39305	<0.001
AIC	145.83530		111.45620	
BIC	148.74480		114.85050	

From this table, it is clear that the BPL regression model has a better fit than the Poisson regression model with the smallest AIC and BIC. In conclusion, all the covariates can explain the number of survivors.

The corresponding quantile–quantile (Q–Q) plots are shown in Figure 7. These graphs demonstrate that the BPL regression model is better than the Poisson regression model.

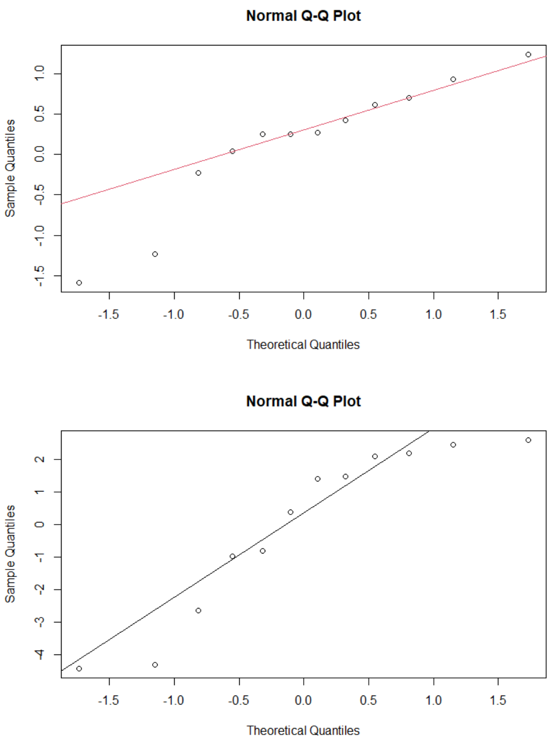

Figure 7. The Q–Q plots of the BPL and Poisson regression models, respectively.

6.5.2. Low Birth Weight Data

The second data set used here is the low birth weight data. It is taken from the **COUNT** package in the statistical programming language **R**. The BPL regression model is used to model the number of low-weight babies (*lowbw*) (y_i) by using the covariates, *cases* (x_{1i}), *race1* (x_{2i}) and *race2* (x_{3i}). The summary statistics for the low birth weight data are shown in Table 10.

Table 10. Summary statistics for the low birth weight data set.

Variables	Min	Max	Median
lowbw	12	60	16.5
cases	30	90	165
race1	0	0.5	1
race2	0	0	1

The results of the regression analysis applied to the low birth weight data are given in Table 11.

Table 11. Modeling results for the low birth weight data set.

Covariates	Poisson		BPL	
	Estimates	*p*-Values	Estimates	*p*-Values
γ_0	2.0679	<0.001	2.2041	0.0194
γ_1	0.0124	<0.001	0.0119	0.2390
γ_2	−0.3287	0.0690	−0.4641	0.8689
γ_3	0.2192	0.0505	0.1506	0.8273
AIC	61.9544		59.31121	
BIC	60.9132		58.06177	

According to this table, the BPL regression model offers a better fit than the Poisson regression model since it has lower AIC and BIC values. Additionally, the covariates have no statistically significant effect on the number of low-weight babies.

Figure 8 presents the Q–Q plots corresponding with the low birth weight data. Here also, these graphs demonstrate that the BPL regression model is better than the Poisson regression model.

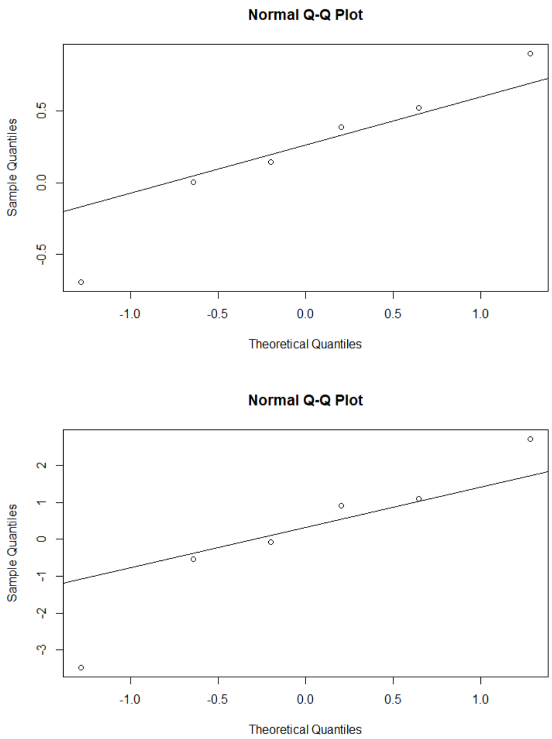

Figure 8. The Q–Q plots of the BPL and Poisson regression models, respectively.

7. Conclusions

This paper focused on a two-parameter discrete distribution generated from the sum of two independent random variables, one with the Bernoulli distribution and the other with the Poisson–Lindley distribution. We have naturally called it the Bernoulli–Poisson–Lindley distribution. This distribution has a number of advantages, including the absence of special functions in its pmf and cdf, as well as its utilization of only two parameters. Furthermore, the model's ability to exhibit under- or over-dispersion makes it well-suited for modeling purposes. With the aim of estimating the unknown parameter, the ML method was used, and a simulation exercise was conducted. Furthermore, its associated count regression model was developed and discussed from an inferential viewpoint. The regression model is applied to two real-life data sets, and it is observed that our model is competitive in modeling practical data. To assess the viability of the suggested paradigm, two real-world data sets are examined. Favorable results were obtained for the proposed modeling strategy in all cases. Thus, the BPL distribution will be productive in modeling count data, beyond the scope of this paper.

Author Contributions: Conceptualization, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; methodology, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; software, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; validation, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; formal analysis, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; investigation, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; writing—original draft preparation, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; writing—review and editing, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; visualization, H.S.B., C.C., R.M., M.R.I., S.A. and N.Q.; funding acquisition, N.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used in this paper are well referenced.

Acknowledgments: The authors gratefully acknowledge Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R376), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia for the financial support for this project.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

- The formula for a finite geometric series is as follows:

$$\sum_{i=0}^{n} r^i = \frac{1 - r^{n+1}}{1 - r},$$

where $r \in \mathbb{R}$ and n is a positive integer. When $|r| < 1$, by applying $n \to \infty$, we obtain the standard infinite geometric formula, which can be generalized for any non-negative integer k as follows:

$$\sum_{i=0}^{\infty} i(i-1)\dots(i-k+1)r^{i-k} = \frac{k!}{(1-r)^{k+1}}.$$

- The R-code for the empirical study of BPL distribution is given below.

```
library(AdequacyModel)
data<-NULL

n<-length(data)
n
x<-mean(y)
x
TTT(y)
dbpl <- function(x,alpha,theta) {
```

```
ifelse (x==0,(((1-alpha)*(theta^2)*(theta+2))/((theta+1)^3)),  \\
(((theta^2)*((1+alpha*theta)*(x+theta+1)+(1-alpha))/((theta+1)^(x+3)))))
}
dbpl(1,0.25,0.66)
pbpl <- function(q,alpha,theta){
(1-(1+theta*(3+q+theta+(q*alpha*theta)+ \\
 (alpha*theta*(2+theta))))/((1+theta)^(q+3)))
}

z<-sort(y)
c1=c(0,-1)
a1=matrix(c(1,0,-1,0),byrow = TRUE,2)
a1

L<-function(par)
{alpha=par[1];theta=par[2]
res= - sum(log(dbpl(y,alpha,theta)));
return(res);
}
initial<-c()
est=constrOptim(initial,L,ci=c1,ui=a1,grad = NULL)
est
ks.test(y,"pbpl",initial)
```

References

1. Plan, E. Modeling and simulation of count data. In *CPT: Pharmacometrics & Systems Pharmacology*; ASCPT: Alexandria, VA, USA , 2014; Volume 3, pp. 1–12.
2. Miaou, S.P. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accid. Anal. Prev.* **1994**, *26*, 471–482. [CrossRef] [PubMed]
3. Sáez-Castillo, A.; Conde-Sánchez, A. A hyper-Poisson regression model for overdispersed and under-dispersed count data. *Comput. Stat. Data Anal.* **2013**, *61*, 148–157. [CrossRef]
4. Del Castillo, J.; Pérez-Casany, M. Weighted Poisson distributions for overdispersion and underdispersion situations. *Ann. Inst. Stat. Math.* **1998**, *50*, 567–585. [CrossRef]
5. del Castillo, J.; Pérez-Casany, M. Overdispersed and under-dispersed Poisson generalizations. *J. Stat. Plan. Inference* **2005**, *134*, 486–500. [CrossRef]
6. Greenwood, M.; Yule, G.U. An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *J. R. Stat. Soc.* **1920**, *83*, 255–279. [CrossRef]
7. Nakagawa, T.; Osaki, S. The discrete Weibull distribution. *IEEE Trans. Reliab.* **1975**, *24*, 300–301. [CrossRef]
8. Stein, W.E.; Dattero, R. A new discrete Weibull distribution. *IEEE Trans. Reliab.* **1984**, *33*, 196–197. [CrossRef]
9. Khan, M.A.; Khalique, A.; Abouammoh, A. On estimating parameters in a discrete Weibull distribution. *IEEE Trans. Reliab.* **1989**, *38*, 348–350. [CrossRef]
10. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* **1994**, *39*, 306–314. [CrossRef]
11. Krishna, H.; Pundir, P.S. Discrete Burr and discrete Pareto distributions. *Stat. Methodol.* **2009**, *6*, 177–188. [CrossRef]
12. Jazi, M.A.; Lai, C.D.; Alamatsaz, M.H. A discrete inverse Weibull distribution and estimation of its parameters. *Stat. Methodol.* **2010**, *7*, 121–132. [CrossRef]
13. Bakouch, H.S.; Jazi, M.A.; Nadarajah, S. A new discrete distribution. *Statistics* **2014**, *48*, 200–240. [CrossRef]
14. Dean, C.B.; Lundy, E.R. Overdispersion. In *Wiley StatsRef: Statistics Reference Online*; Wiley: Hoboken, NJ, USA, 2014; pp. 1–9.
15. Consul, P.; Famoye, F. Generalized Poisson regression model. *Commun. Stat.-Theory Methods* **1992**, *21*, 89–109. [CrossRef]
16. Sellers, K.F.; Shmueli, G. A flexible regression model for count data. *Ann. Appl. Stat.* **2010**, *4*, 943–961. [CrossRef]
17. Bonat, W.H.; Jørgensen, B.; Kokonendji, C.C.; Hinde, J.; Demétrio, C.G. Extended Poisson-Tweedie: Properties and regression models for count data. *Stat. Model.* **2018**, *18*, 24–49. [CrossRef]
18. Sankaran, M. 275. note: The discrete Poisson–Lindley distribution. *Biometrics* **1970**, *26*, 145–149. [CrossRef]
19. Bourguignon, M.; Gallardo, D.I.; Medeiros, R.M. A simple and useful regression model for under-dispersed count data based on Bernoulli–Poisson convolution. *Stat. Pap.* **2022**, *63*, 821–848. [CrossRef]

20. Chakraborty, S.; Chakravarty, D.; Mazucheli, J.; Bertoli, W. A discrete analog of Gumbel distribution: Properties, parameter estimation and applications. *J. Appl. Stat.* **2021**, *48*, 712–737. [CrossRef]
21. Altun, E.; Bhati, D.; Khan, N.M. A new approach to model the counts of earthquakes: INARPQX (1) process. *SN Appl. Sci.* **2021**, *3*, 274. [CrossRef]
22. Bjerkedal, T. Acquisition of Resistance in Guinea Pies infected with Different Doses of Virulent Tubercle Bacilli. *Am. J. Hyg.* **1960**, *72*, 130–148.
23. Gupta, R.D.; Kundu, D. A new class of weighted exponential distributions. *Statistics* **2009**, *43*, 621–634. [CrossRef]