symmetry

MDPI

Article

Lagrangian Zero Truncated Poisson Distribution: Properties Regression Model and Applications

Muhammed Rasheed Irshad [1], Christophe Chesneau [2,*], Damodaran Santhamani Shibu [3], Mohanan Monisha [3] and Radhakumari Maya [4]

[1] Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India
[2] Department of Mathematics, Université de Caen Basse-Normandie, LMNO, UFR de Sciences, F-14032 Caen, France
[3] Department of Statistics, University College, Thiruvananthapuram 695 034, Kerala, India
[4] Department of Statistics, Government College for Women, Thiruvananthapuram 695 014, Kerala, India
* Correspondence: christophe.chesneau@unicaen.fr

Abstract: In this paper, we construct a new Lagrangian discrete distribution, named the Lagrangian zero truncated Poisson distribution (LZTPD). It can be presented as a generalization of the zero truncated Poissson distribution (ZTPD) and an alternative to the intervened Poisson distribution (IPD), which was elaborated for modelling both over-dispersed and under-dispersed count datasets. The mathematical aspects of the LZTPD are thoroughly investigated, and its connection to other discrete distributions is crucially observed. Further, we define a finite mixture of LZTPDs and establish its identifiability condition along with some distributional aspects. Statistical work is then performed. The maximum likelihood and method of moment approaches are used to estimate the unknown parameters of the LZTPD. Simulation studies are also undertaken as an assessment of the long-term performance of the estimates. The significance of one additional parameter in the LZTPD is tested using a generalized likelihood ratio test. Moreover, we propose a new count regression model named the Lagrangian zero truncated Poisson regression model (LZTPRM) and its parameters are estimated by the maximum likelihood estimation method. Two real-world datasets are considered to demonstrate the LZTPD's real-world applicability, and healthcare data are analyzed to demonstrate the LZTPRM's superiority.

Keywords: Lagrangian zero truncated Poisson distribution; intervened Poisson distribution; index of dispersion; regression; maximum likelihood estimation; generalized likelihood ratio test; simulation

Citation: Irshad, M.R.; Chesneau, C.; Shibu, D.S.; Monisha, M.; Maya, R. Lagrangian Zero Truncated Poisson Distribution: Properties Regression Model and Applications. *Symmetry* **2022**, *14*, 1775. https://doi.org/10.3390/sym14091775

Academic Editor: Jinyu Li

Received: 3 May 2022
Accepted: 21 August 2022
Published: 25 August 2022

1. Introduction

In several cases, researchers are not capable of perceiving the unabridged distribution of counts; in particular, the zeros are not often observed, which indicates that zero truncation is found to be an important and common characteristic for various count data processes. With this in mind, ref. [1] employed the zero truncated Poisson distribution (ZTPD) to interpret a chance mechanism whose experimental device becomes active only when at least one event occurs. Ref. [2] discussed numerical examples to demonstrate the statistical applications of the ZTPD in such situations. An alternative to the ZTPD was proposed by [3], the so-called intervened Poisson distribution (IPD), to deal with the real-life situation of a manager in a supermarket who provided extra assistance to the customers at a service counter. An attraction of the IPD over the ZTPD is that it gives information on the effectiveness of intervention in the situation. Ref. [4] applied the IPD in the fields of reliability analysis, queuing problems, and epidemiological studies, etc. Ref. [5] considered a modified version of the IPD which has an advantage over the IPD in stretching the probability in all directions so that clustering of probabilities at initial values of the operating mechanism is overlooked. Ref. [6] illustrated an alternative to the IPD for prevalence reduction. However, the IPD proposed by [3] has a restriction that the variance should be less than the mean. This is referred to as 'under-dispersion'in the literature, and this phenomenon is only observed on rare occasions. To solve this limitation, we propose a new ZTPD based on a Lagrangian approach, dubbed the Lagrangian zero truncated Poisson distribution (LZTPD), that can model both under-dispersed and over-dispersed (variance greater than mean) count datasets. More information on the Lagrangian distributional approach is given below.

First, the Lagrangian family (LF) of distributions was derived from the Lagrangian expansion, which was first introduced by [7]. Later, refs. [8,9] proposed a discrete LF (DLF), which itself forms a very large and important class containing numerous families of probability distributions. For example, the Lagrangian negative binomial distribution, obtained by [10], shows its usefulness in a queuing process. The Lagranian Katz family was developed by [11]. Ref. [12] considered the applications of Lagrangian probability distributions to inferential problems in a random mapping theory. Ref. [13] derived the generalized Poisson gamma dependence model from Lagrangian probability models. Recently, ref. [14] applied the Lagrangian probability density function models for collisional turbulent fluid particle flows. Furthermore, ref. [9] proved that all discrete Lagrangian distributions converge to the normal distribution and to the inverse Gaussian distribution under certain conditions. Thus, we propose the LZTPD in this article, motivated by the adaptability of the Lagrangian distributions and the need to propose a flexible model capable of modelling versatile count datasets.

On the other hand, the regression model for count data is gaining more and more attention these days. The use of regression models to describe count data is relatively recent, as detailed in [15]. However, in some real-world situations, the system will only be engaged if at least one event occurs. Examples include the number of international conflicts, daily accidents, industrial injuries, etc. In many circumstances, counting outcomes directly with a normal linear regression model will result in inefficient, inconsistent, and biased estimation, as described in [16]. Positive count data are analyzed using the zero truncated Poisson regression model (ZTPRM), which is more accurate than the traditional Poisson regression model for this type of data. Ref. [17] has discussed the application of Poisson regression models to the analysis of truncated samples of count data. Recently, ref. [18] developed the intervened Poisson regression model (IPRM), which is an alternative to the ZTPRM. In this paper, we offer an alternative regression model to both the ZTPRM and IPRM, the so-called Lagrangian zero truncated Poisson regression model (LZTPRM). The LZTPD and the LZTPRM are motivated by their suitability for both under-dispersed and over-dispersed count datasets, as well as their applicability in situations where the modeled data excludes zero-counts.

The remaining sections of the presented study can be summarized as follows: A brief introduction to the Lagrangian expansion and DLF are given in Section 2. The LZTPD along with its statistical properties are discussed in Section 3. In Section 4, we propose a mixture LZTPD and present the identifiability conditions of finite mixtures of LZTPDs. The maximum likelihood (ML) estimation method used to investigate the parameter estimation of the LZTPD is discussed in Section 5. In Section 6, we test the significance of an additional parameter of the LZTPD using a generalized likelihood ratio test. The simulation results of the considered estimation method are presented in Section 7. The LZTPRM is elucidated in Section 8. Empirical illustrations of the proposed LZTPD and LZTPRM are given in Section 9. Discussions and conclusions are given in Section 10 and Section 11, respectively.

2. Some Preliminaries

2.1. Basics on the Discrete Lagrangian Family

In order to introduce the LZTPD, some mathematical background on the Lagrangian expansion and DLF must be recalled. Let $f_1(z)$ and $f_2(z)$ be two analytic and successively differentiable functions defined on the interval $[-1,1]$ such that $f_1(1) = f_2(1) = 1$, $f_1(0) \neq 0$, and $f_2(0) \geq 0$. The following power series expansion was obtained by inverting the Lagrange transformation $u = \frac{z}{f_1(z)}$, which provided the value of z as a power series in u:

$$\frac{f_2(z)}{1 - \frac{z\, f_1'(z)}{f_1(z)}} = \sum_{j=0}^{\infty} b_j\, u^j, \tag{1}$$

where $b_0 = f_2(0)$ and $b_j = \frac{1}{j!} D^j \left[(f_1(z))^j f_2(z) \right] \Big|_{z=0}$, with $D^j = \frac{\partial^j}{\partial z^j}$ and $f_1'(z) = \frac{\partial f_1(z)}{\partial z}$.

The details can be found in [19].

Furthermore, if

$$0 < f_1'(1) < 1 \text{ and } D^j \left[(f_1(z))^j f_2(z) \right] \Big|_{z=0} \geq 0,\ j \geq 0, \tag{2}$$

the Lagrangian expansion (1) defines the DLF. Thus, a random variable (rv) Y belonging to the DLF has the following pmf:

$$P(Y = y) = (1 - f_1'(1)) \frac{D^y[(f_1(z))^y f_2(z)]}{y!} \Big|_{z=0},\ y = 0, 1, 2, \ldots \tag{3}$$

See [19,20] for more information. The corresponding probability generating function (pgf) is given by

$$G(u) = \frac{(1 - f_1'(z)) f_2(z)}{1 - \frac{z f_1'(z)}{f_1(z)}}, \tag{4}$$

where $z = u\, f_1(z)$.

2.2. Importance of the Lagrangian Family

In the following, we list some results highlighting the choice of the following parametric exponential function:

$$f_1(z) = e^{\lambda(z-1)}, \tag{5}$$

with $0 < \lambda < 1$, into the DLF definition, which has generated several distributions of importance.

Proposition 1. *The distribution of the DLF defined with $f_1(z)$ as in (5) and $f_2(z) = \frac{(e^{z\theta}-1)(1-\lambda z)}{(e^{\theta}-1)(1-\lambda)}$ corresponds to the zero truncated generalized Poisson distribution (ZTGPD) defined by [21].*

Proof. Based on (3), the pmf of the considered distribution is obtained as

$$\begin{aligned} h_1(y) &= \frac{1-\lambda}{y!} D^y \left[e^{\lambda y(z-1)} \frac{(e^{z\theta} - 1)(1 - \lambda z)}{(e^{\theta} - 1)(1 - \lambda)} \right] \Bigg|_{z=0} \\ &= \frac{\theta(\theta + \lambda y)^{y-1} e^{-\theta - \lambda y}}{(1 - e^{-\theta}) y!}, \quad y = 1, 2, 3, \ldots, \end{aligned}$$

which corresponds to the pmf of the ZTGPD. Hence, the result. □

Proposition 2. *The distribution of the DLF defined with $f_1(z)$ as in (5) and $f_2(z) = e^{\theta(z-1)}$ corresponds to the Lagrangian type linear function Poisson distribution given by [22].*

Proof. Based on (3), the pmf of the considered distribution is obtained as

$$\begin{aligned} h_2(y) &= \frac{1-\lambda}{y!} D^y \left[e^{\lambda y(z-1)} e^{\theta(z-1)} \right] \Big|_{z=0} \\ &= \frac{1-\lambda}{y!} (\theta + \lambda y)^y e^{-\theta - \lambda y}, \quad y = 0, 1, 2, \dots, \end{aligned}$$

which is the pmf of the Lagrangian type linear function Poisson distribution. The distribution correspondence is proved. □

Proposition 3. *The distribution of the DLF defined with $f_1(z)$ as in (5) and $f_2(z) = z$ corresponds to the Sudha Lagrangian distribution given in [23].*

Proof. Based on (3), the pmf of the considered distribution is obtained as

$$\begin{aligned} h_3(y) &= \frac{1-\lambda}{y!} D^y \left[e^{\lambda y(z-1)} z \right] \Big|_{z=0} \\ &= (1-\lambda) \frac{e^{-\lambda y} (\lambda y)^{y-1}}{(y-1)!}, \quad y = 1, 2, 3, \dots, \end{aligned}$$

which is the Sudha Lagrangian distribution. Hence, the result. □

Proposition 4. *The distribution of the DLF defined with $f_1(z)$ as in (5) and $f_2(z) = z^n$ corresponds to the Lagrangian type weighted delta Poisson distribution given in [23].*

Proof. Based on (3), the pmf of the considered distribution is indicated as

$$\begin{aligned} h_4(y) &= \frac{1-\lambda}{y!} D^y \left[e^{\lambda y(z-1)} z^n \right] \Big|_{z=0} \\ &= (1-\lambda) \frac{e^{-\lambda y} (\lambda y)^{y-n}}{(y-n)!}, \quad y = n, n+1, \dots, \end{aligned}$$

which is the pmf of the Lagrangian weighted delta Poisson distribution. The desired result follows. □

In view of the applications of the DLF configured with the function $f_1(z)$ in (5), it is motivating to explore a new horizon of distribution with the choice of a new function $f_2(z)$. The new distribution of the study presented below is based on this idea.

3. Lagrangian Zero Truncated Poisson Distribution (LZTPD)

In this section, based on the DLF, we explicitly define the LZTPD. We also examine its properties, such as median, mode, pgf, moment generating function (mgf), factorial moments, index of dispersion (IOD), coefficient of variation (CV) and hazard rate function (hrf), etc. Several propositions are made here to discuss the connections between the LZTPD and certain other Lagrangian distributions.

Definition 1. *The LZTPD is the special distribution of the DLF under the following original configuration: $f_1(z) = e^{\lambda(z-1)}$ and $f_2(z) = \frac{e^{z\theta}-1}{e^{\theta}-1}$. Then, a rv Y is said to follow the LZTPD, if its pmf has the following form:*

$$h(y) = \frac{(1-\lambda) e^{-\theta - \lambda y} [(\theta + \lambda y)^y - (\lambda y)^y]}{(1 - e^{-\theta}) y!}, \quad y = 1, 2, 3, \dots, \tag{6}$$

where $\theta > 0$ and $0 < \lambda < 1$.

Proof. First, note that $f_1(z)$ and $f_2(z)$ satisfy the conditions in (2). Then the pmf given in (3) can be derived as

$$\begin{aligned} P(Y=y) &= \frac{1-\lambda}{y!} D^y \left[e^{\lambda y(z-1)} \frac{e^{z\theta}-1}{e^{\theta}-1} \right] \Bigg|_{z=0} \\ &= \frac{(1-\lambda)e^{-\theta-\lambda y}[(\theta+\lambda y)^y - (\lambda y)^y]}{(1-e^{-\theta})y!}. \end{aligned}$$

Hence, the proof. □

A distribution with the pmf given in (6) will be denoted as LZTPD(θ, λ). For $\lambda \to 0$, the LZTPD(θ, λ) reduces to the ZTPD (see [24]). As a result, we can say that the LZTPD(θ, λ) is a generalization of the ZTPD. Figures 1 and 2 display the graphical representation of the pmf of the LZTPD for different parameter values of θ and λ.

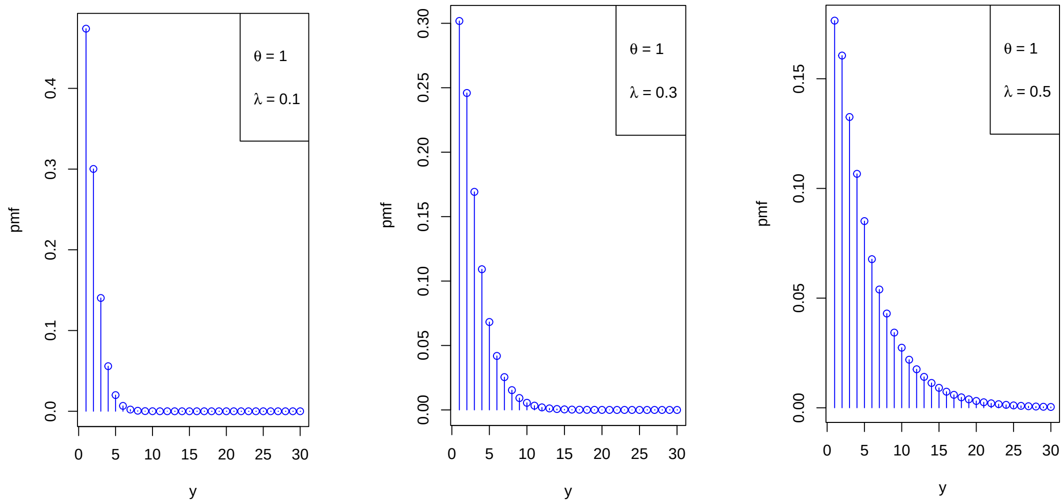

Figure 1. Various shapes of the LZTPD pmf when λ increases.

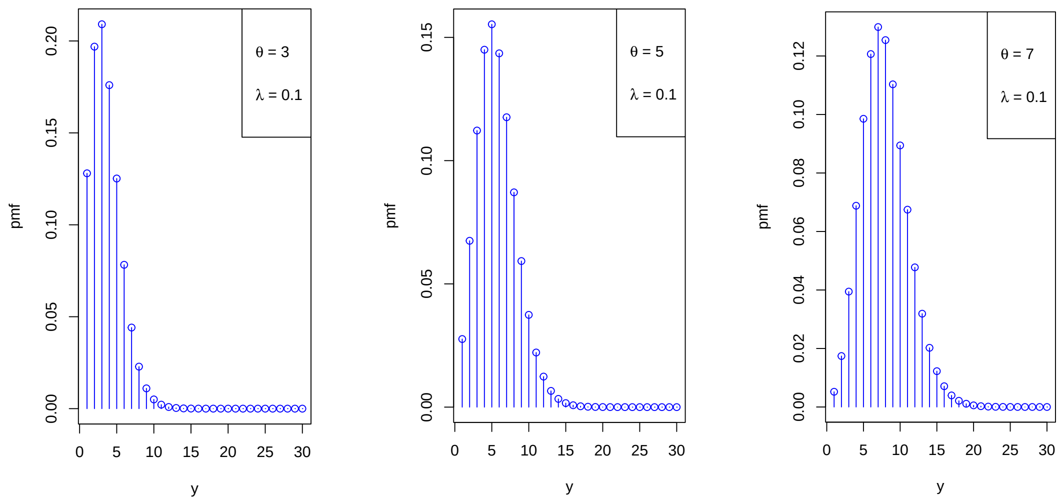

Figure 2. Various shapes of the pmf of the LZTPD when θ increases.

The hrf of the LZTPD is obtained by substituting the pmf in the following equation

$$h_y = P(Y = y | Y \geq y) = \frac{h(y)}{\sum_{j=y}^{\infty} h(j)}. \tag{7}$$

From (7), it is clear that determining the closed form expression of the hrf is more intricate. We have drawn the graph of the hrf to determine its possible shapes.

Figure 3 demonstrates that the LZTPD has an increasing hrf.

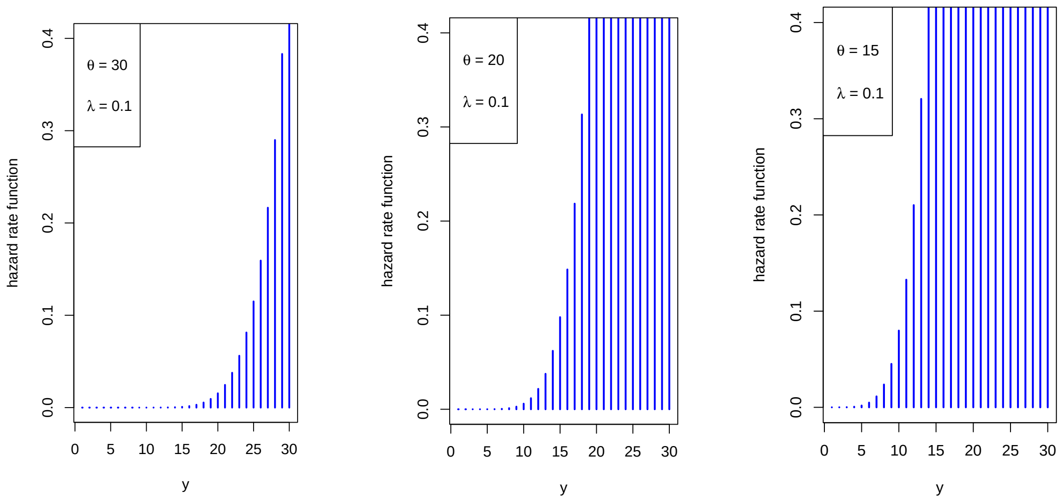

Figure 3. Plots of the hrf of the LZTPD distribution.

Proposition 5. *Let Y be a rv following the LZTPD. Then the median of Y is defined by the smaller integer m in $\{1,2,\ldots\}$ such that*

$$\sum_{y=1}^{m} \frac{e^{-\lambda y}((\theta+\lambda y)^{y}-(\lambda y)^{y})}{y!} \geq \frac{e^{\theta}-1}{2(1-\lambda)}. \tag{8}$$

Proof. By the definition, m is the smaller integer in $\{1,2,\ldots\}$ such that $P(Y \leq m) \geq \frac{1}{2}$, which is equivalent to the desired result. □

Proposition 6. *Let Y be a rv following the LZTPD. Then, the mode of Y, denoted by y_m, exists in $\{1,2,\ldots\}$, and lies in the case:*

$$\frac{\varrho(y_m+1)}{\varrho(y_m)} - e^{\lambda} \leq y_m\, e^{\lambda} \leq \frac{\varrho(y_m)}{\varrho(y_m-1)}, \tag{9}$$

where $\varrho(y_m) = (\theta+\lambda y_m)^{y_m} - (\lambda y_m)^{y_m}$.

Proof. By the definition of the mode, we must find the integer $y = y_m$ for which $h(y)$ has the greatest value. That is, we aim to solve $h(y) \geq h(y-1)$ and $h(y) \geq h(y+1)$. First, note that $h(y)$ can also be written as:

$$h(y) = \frac{1-\lambda}{e^{\theta}-1}\,\frac{e^{-\lambda y}\varrho(y)}{y!}, \tag{10}$$

where $\varrho(y) = (\theta+\lambda y)^{y} - (\lambda y)^{y}$.

Obviously, $h(y) \geq h(y-1)$ implies that

$$\frac{\varrho(y)}{\varrho(y-1)} \geq y\, e^{\lambda}. \tag{11}$$

Also, $h(y) \geq h(y+1)$ implies that

$$\frac{\varrho(y+1)}{\varrho(y)} \leq (y+1)\, e^{\lambda}. \tag{12}$$

By combining (11) and (12), we get (9), hence, the proof. □

Proposition 7. *The LZTPD(θ, λ) is a member of the modified power series family of distributions defined by [25].*

Proof. According to [25], the pmf of the modified power series distribution (MPSD) is given by

$$h_5(y) = \frac{r_y[\Psi(\theta)]^y}{\kappa(\theta)}, \quad y \in G \subseteq N,$$

where N is the set of non-negative integers, G is a subset of N, $r_y \geq 0$ for all $y \in N$, and $\kappa(\theta) = \sum_{y \in G} r_y[\Psi(\theta)]^y$ can be viewed as a normalization constant. By its basic definition, the pmf of the LZTPD in (6) satisfies

$$\sum_{y=1}^{\infty} h(y) = 1,$$

which implies that

$$\sum_{y=1}^{\infty} \frac{e^{-\lambda y}[(\theta + \lambda y)^y - (\lambda y)^y]}{y!} = \frac{e^{\theta} - 1}{1 - \lambda}.$$

Also, we have

$$\frac{e^{\theta} - 1}{1 - \lambda} = \sum_{y=1}^{\infty} \frac{(1 + y)^{y-1}(\theta e^{-\theta})^y}{(1 - \lambda)y!} = \sum_{y=1}^{\infty} r_y[\Psi(\theta)]^y,$$

where

$$r_y = \frac{(1 + y)^{y-1}}{(1 - \lambda)y!}, \quad \Psi(\theta) = \theta e^{-\theta},$$

and $\kappa(\theta) = \sum_{y=1}^{\infty} r_y[\Psi(\theta)]^y$.

Hence, the pmf of the LZTPD given in (6) can be expressed under the following form:

$$h(y) = \frac{r_y[\Psi(\theta)]^y}{\kappa(\theta)}.$$

This completes the proof. □

Proposition 8. *The pgf of a rv Y following the LZTPD(θ, λ) is expressed as*

$$G(u) = \mathrm{E}(u^Y) = \frac{(1 - \lambda)(e^{z\theta} - 1)}{(e^{\theta} - 1)(1 - \lambda z)}, \tag{13}$$

where we recall that z and u are related by the following equation: $z = u\, e^{\lambda(z-1)}$.

Proof. Based on (4), we directly obtain

$$G(u) = \frac{(1 - f_1'(1))f_2(z)}{1 - \frac{z f_1'(z)}{f_2(z)}} = \frac{(1 - \lambda)(e^{z\theta} - 1)}{(e^{\theta} - 1)(1 - \lambda z)}.$$

Thus, the proof is obtained. □

Corollary 1. *The mgf of a rv Y following the LZTPD(θ, λ) is obtained by putting $z = e^s$ and $u = e^k$ in (13), and we get*

$$M(k) = \mathrm{E}(e^{kY}) = \frac{(1-\lambda)(e^{\theta\, e^s} - 1)}{(e^\theta - 1)(1 - \lambda e^s)},$$

where $s = k + \lambda(e^s - 1)$.

Corollary 2. *The cumulant generating function (cgf) of a rv Y following the LZTPD(θ, λ) given in (6) becomes*

$$C(k) = \log\left[M_Y(k)\right] = \log\left[\frac{(1-\lambda)(e^{\theta\, e^s} - 1)}{(e^\theta - 1)(1 - \lambda e^s)}\right],$$

where $s = k + \lambda(e^s - 1)$.

Proposition 9. *Let $Y_1, Y_2, \ldots, Y_n$ be n independently and identically distributed (iid) rvs following the LZTPD(θ, λ). Then the distribution of the sample sum $V = \sum_{k=1}^{n} Y_i$ has the following pgf:*

$$\Psi(u) = \frac{(1-\lambda)^n (e^{z\theta} - 1)^n}{(e^\theta - 1)^n (1 - \lambda z)^n},$$

where $z = u\, e^{\lambda(z-1)}$.

Proof. Based on the pgf of the LZTPD given in (13), the pgf of the rv V becomes

$$\begin{aligned}\Psi(u) &= \mathrm{E}(u^V) = \mathrm{E}(u^{Y_1+Y_2+\ldots+Y_n}) = \prod_{k=1}^{n} \mathrm{E}(u^{Y_k}) = \prod_{k=1}^{n} G(u) = [G(u)]^n \\ &= \frac{(1-\lambda)^n (e^{z\theta} - 1)^n}{(e^\theta - 1)^n (1 - \lambda z)^n}.\end{aligned}$$

This completes the proof. □

Proposition 10. *For any integer $r \geq 1$, the rth factorial moment of a rv Y following the LZTPD(θ, λ) is given by*

$$\begin{aligned}\mu_{[r]} &= \mathrm{E}[Y(Y-1)\ldots(Y-r+1)] \\ &= \left\{(e^\theta - 1)^{-1} D_r\left(e^{z\theta}\right) + \lambda \frac{\sum_{i=1}^{r}(r-i+1)\mu_{[r-i]} D_i\left(u\, e^{\lambda(z-1)}\right)}{1-\lambda}\right\}\Bigg|_{u=z=1},\end{aligned} \tag{14}$$

where $z = u\, e^{\lambda(z-1)}$.

Proof. By its definition, it is obtained by successively differentiating $G(u)$ given in (4) in r times with respect to u and by putting $u = z = 1$. Thus, it is given by

$$G(u) = \frac{(1 - f_1'(1)) f_2(z)}{1 - u\, f_1'(z)},$$

implying that

$$(1 - u f_1'(z)) G(u) = (1 - f_1'(1)) f_2(z).$$

Taking the first derivative with respect to u on both sides, we get

$$G(u) D_1 (1 - u f_1'(z)) + G'(u)(1 - u f_1'(z)) = (1 - f_1'(1)) D_1 f_2(z). \tag{15}$$

Again, by taking the derivative of (15) with respect to u on both sides, we get

$$G(u)D_2(1-uf_1'(z))+2D_1(1-uf_1'(z))G'(u)+(1-uf_1'(z))G''(u)$$
$$=(1-f_1'(1))D_2f_2(z).$$

Proceeding like this, we get the rth derivative is of the following form:

$$G^r(u)=\frac{(1-f_1'(1))D_rf_2(z)-\sum_{i=1}^{r}(r-i+1)\,D_i(1-uf_1'(z))\,G^{r-i}(u)}{1-uf_1'(z)} \tag{16}$$

Substitute $f_1(z)=e^{\lambda(z-1)}$, $f_2(z)=\frac{e^{z\theta}-1}{e^{\theta}-1}$ and $z=u=1$ in (16), we get (14). Thus the proof is obtained. □

Proposition 11. *The mean and variance of a rv Y following the LZTPD(θ,λ) are*

$$\mu=\mathrm{E}(Y)=\frac{\lambda}{(1-\lambda)^2}+\frac{\theta}{(1-e^{-\theta})(1-\lambda)}$$

and

$$\sigma^2=\mathrm{Var}(Y)=\frac{\lambda+\lambda^2}{(1-\lambda)^4}+\frac{\theta^2(1-\lambda)+\theta}{(1-e^{-\theta})(1-\lambda)^3}-\frac{\theta^2}{(1-e^{-\theta})^2(1-\lambda)^2},$$

respectively.

Proof. The first two factorial moments can be obtained by using (14) as follows:

$$\begin{aligned}\mathrm{E}(Y)=\mu&=\frac{f_2'(1)}{1-f_1'(1)}+\frac{f_1''(1)+f_1'(1)-(f_1'(1))^2}{(1-f_1'(1))^2}\\&=\frac{\lambda}{(1-\lambda)^2}+\frac{\theta}{(1-e^{-\theta})(1-\lambda)}\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}[Y(Y-1)]=&\frac{f_2'(1)+f_1''(1)+4f_2'(1)f_1'(1)+2(f_1'(1))^2}{(1-f_1'(1))^2}\\&+\frac{f_1'''(1)+f_1''(1)+3f_2'(1)f_1''(1)+5f_1'(1)f_1''(1)}{(1-f_1'(1))^3}\\&+\frac{3(f_1''(1))^2}{(1-f_1'(1))^4}.\end{aligned}$$

Furthermore, we have

$$\begin{aligned}\mathrm{Var}(Y)=\sigma^2&=\mathrm{E}[Y(Y-1)]+\mathrm{E}(Y)-[\mathrm{E}(Y)]^2\\&=\frac{f_2''(1)+f_2'(1)-(f_2'(1))^2}{(1-f_1'(1))^2}+\frac{(1+f_2'(1))(f_1''(1)+f_1'(1)-(f_1'(1))^2)}{(1-f_1'(1))^3}\\&+\frac{f_1'''(1)+f_1'(1)f_1''(1)+2f_1''(1)}{(1-f_1'(1))^3}+\frac{2(f_1''(1))^2}{(1-f_1'(1))^4}\\&=\frac{\lambda+\lambda^2}{(1-\lambda)^4}+\frac{\theta^2(1-\lambda)+\theta}{(1-e^{-\theta})(1-\lambda)^3}-\frac{\theta^2}{(1-e^{-\theta})^2(1-\lambda)^2},\end{aligned}$$

where $f_1'(1), f_1''(1), f_1'''(1), f_2'(1), f_2''(1)$ denote the values of the successive derivatives of $f_1(z)$ and $f_2(z)$, respectively, evaluated at the special value $z=1$.
Hence, the proof. □

Proposition 12. *The index of dispersion and coefficient of variation of a rv Y following the LZTPD(θ, λ) are*

$$IOD = \frac{(\lambda+\lambda^2)(1-e^{-\theta})^2 + (\theta^2 - \lambda\theta^2 + \theta)(1-\lambda)(1-e^{-\theta}) - \theta^2(1-\lambda)^2}{\lambda(1-e^{-\theta})^2(1-\lambda)^2 + \theta(1-e^{-\theta})(1-\lambda)^3}$$

and

$$CV = \frac{\sqrt{(\lambda+\lambda^2)(1-e^{-\theta})^2 + (\theta^2 - \lambda\theta^2 + \theta)(1-\lambda)(1-e^{-\theta}) - \theta^2(1-\lambda)^2}}{\lambda(1-e^{-\theta}) + \theta(1-\lambda)},$$

respectively.

Proof. A normalized measure of dispersion can be obtained by using the variance to mean relationship. This measure is the well-known IOD, and it is given by

$$\begin{aligned} IOD &= \frac{\sigma^2}{\mu} \\ &= \frac{(\lambda+\lambda^2)(1-e^{-\theta})^2 + (\theta^2 - \lambda\theta^2 + \theta)(1-\lambda)(1-e^{-\theta}) - \theta^2(1-\lambda)^2}{\lambda(1-e^{-\theta})^2(1-\lambda)^2 + \theta(1-e^{-\theta})(1-\lambda)^3}. \end{aligned}$$

Analogously, the CV is given by

$$\begin{aligned} CV &= \frac{\sqrt{\sigma^2}}{\mu} \\ &= \frac{\sqrt{(\lambda+\lambda^2)(1-e^{-\theta})^2 + (\theta^2 - \lambda\theta^2 + \theta)(1-\lambda)(1-e^{-\theta}) - \theta^2(1-\lambda)^2}}{\lambda(1-e^{-\theta}) + \theta(1-\lambda)}. \end{aligned}$$

Hence, the proof. □

The coefficients of skewness and kurtosis, respectively, are used to calculate the asymmetry degree and flatness of a distribution. The first is derived by dividing the third central moment by the variance raised to the power of 3/2, and the second is acquired by dividing the fourth central moment by the square of the variance. These coefficients are required to determine the shape of any distribution. The mean, variance, median, mode, CV, IOD, skewness and kurtosis for selected values of parameters of the LZTPD(θ, λ) are summarized in Table 1.

From Table 1, it can be observed that the LZTPD(θ, λ) is both under-dispersed, i.e., $IOD < 1$, and over-dispersed, i.e., $IOD > 1$. This makes a strong difference with the ZTPD and IPD, defined on a similar mathematical basis.

Table 1. Values for some moment measures of the LZTPD for different values of θ and λ.

θ	λ	Mean	Variance	Median	Mode	IOD	CV	Skewness	Kurtosis
0.5	0.05	1.4045	1.0797	1	1	0.7687	0.7398	1.4039	5.1257
	0.1	1.5531	1.3000	1	1	0.8370	0.7341	1.9099	7.8036
	0.2	1.9061	1.9476	1	1	1.0217	0.7321	2.0265	10.0765
	0.3	2.3599	3.0631	2	1	1.2979	0.7416	1.3564	14.9253
	0.4	2.9650	5.1144	2	1	1.7248	0.7627	2.8461	15.4684
	0.5	3.8122	9.2482	3	1	2.4259	0.7977	3.5461	17.4684
1	0.05	1.6652	2.4915	1	1	1.4962	0.9478	1.9810	5.1030
	0.1	1.7577	2.8956	2	1	1.6473	0.9680	2.0319	9.7862
	0.2	1.9774	4.0345	2	1	2.0402	1.0157	2.4052	10.8162
	0.3	2.2599	5.9023	2	1	2.6117	1.0750	2.9737	12.2294
	0.4	2.6366	9.1847	3	1	3.4835	1.1494	3.1804	15.0489
	0.5	3.1639	15.5245	4	1	4.9066	1.2453	3.6804	17.0489

Table 1. *Cont.*

θ	λ	Mean	Variance	Median	Mode	*IOD*	*CV*	Skewness	Kurtosis
2	0.05	2.3739	6.7172	2	2	2.8296	1.0917	2.0589	6.1850
	0.1	2.4415	7.6566	2	2	3.1359	1.1333	2.8669	9.0253
	0.2	2.6021	10.2187	3	2	3.9270	1.2284	3.1943	11.4901
	0.3	2.8086	14.2461	3	2	5.0721	1.3438	3.5848	15.2456
	0.4	3.0840	21.0251	4	2	6.8173	1.4867	3.8983	16.0534
	0.5	3.4695	33.5493	5	2	9.6696	1.6694	3.9983	17.0534
3	0.05	3.2125	13.0707	3	3	4.0686	1.1253	1.4589	6.9185
	0.1	3.2741	14.7966	3	3	4.5192	1.1748	1.9866	10.0253
	0.2	3.4202	19.4386	4	3	5.6833	1.2890	2.6943	11.4901
	0.3	3.6082	26.5960	5	3	7.3709	1.4292	2.9848	13.2456
	0.4	3.8587	38.3929	5	4	9.9494	1.6057	3.1983	16.7534
	0.5	4.2095	59.6845	7	4	14.1782	1.8352	3.8983	18.0534

4. Finite Mixtures of Lagrangian Zero Truncated Poisson Distribution

In recent years, finite mixture models have been given much attention in practical situations. Mixture models are widely used in astronomy, biology, genetics, medicine, psychiatry, marketing, etc. For the details, see [26]. The properties of finite mixtures of the IPD and the modified IPD are discussed by [27]. In this section, we derive finite mixtures based on the LZTPD(θ, λ). This mixture model may be consistent with the situation of further interventions.

Let Z be a discrete rv with pmf $h(z) = \sum_{i=1}^{g} l_i h_i(z)$, where $i = 1, 2, \dots g$, $l_i > 0$ such that $\sum_{i=1}^{g} l_i = 1$, $h_i(z) \geq 0$ and $\sum_z h_i(z) = 1$. Then, we state Z has a mixture distribution and $h(z)$ is a finite mixture of distributions. The parameters $l_1, l_2, \dots, l_g$ are known as the mixing weights and $h_1, h_2, \dots, h_g$ as the components of the mixture. The collection of all parameters occurring in the components is represented as Θ and the complete collection of all parameters in the mixture model is represented as Ψ.

Suppose that $\Delta = \{U(z; \theta_i) : \theta_i \in \Theta, z \in R\}$ is the class of pmf's from which mixtures are to be formed. Then the class of finite mixtures of Δ with the appropriate class of pmf's is $\hat{H} = \{H(z) : H(z) = \sum_{i=1}^{g} l_i U(z; \theta_i), l_i > 0, U(z; \theta_i) \in \Delta, i = 1, 2, \dots g\}$. In this setting, $\hat{H}$ is the convex hull of Δ.

Definition 2. *A rv Z is said to have a g component mixture of LZTPDs if it has the pmf $h(z) = P(Z = z)$ of the following form:*

$$h(z) = \sum_{i=1}^{g} l_i h_i(z), \quad z = 1, 2, \dots, \tag{17}$$

where $0 \leq l_i \leq 1$, $\sum_{i=1}^{g} l_i = 1$, and for each $i = 1, 2, 3, \dots g$,

$$h_i(z) = \frac{(1 - \lambda_i) e^{-\theta_i - \lambda_i z} [(\theta_i + \lambda_i z)^x - (\lambda_i z)^z]}{(1 - e^{-\theta_i}) z!},$$

with $0 < \lambda_i < 1$ and $\theta_i > 0$.

A distribution with the pmf given in (17) is called the Lagrangian zero truncated Poisson mixture distribution with g components, or in short, LZTPMD$_g$.

The following theorem from [28] is adopted to construct the identifiability condition of the finite mixture model:

Theorem 1. *A necessary and sufficient condition for $\hat{H}$ to be identifiable is that Δ is linearly independent over the field of real numbers.*

Proof. Proof is given in [28] and hence omitted. □

We are now able to present the identifiability conditions of the LZTPMD$_g$.

Theorem 2. *The identifiability conditions for the LZTPMD$_g$ with the pmf $h(z)$ as given in (17) are $\theta_i \neq \theta_j$, $\lambda_i \neq \lambda_j$ for $i, j \in \{1, 2, \ldots, g\}$ such that $i \neq j$.*

Proof. Take $g = 2$ and consider the equation

$$d_1 H_1(z) + d_2 H_2(z) = 0, \tag{18}$$

where d_1 and d_2 are any two arbitrary real numbers, $H_1(z) = \sum_{j=1}^{z} h(j)$ and $H_2(z) = \sum_{j=1}^{z} \phi(j)$ for $z = 1, 2, \ldots,$ in which $\phi(j)$ is obtained from $h(j)$ by replacing θ_j by δ_j and λ_j by γ_j.

Assume that, $\theta_i \neq \theta_j$, $\lambda_i \neq \lambda_j$ for $i, j \in (1, 2)$ such that $i \neq j$, $\theta_i \neq \theta_j$ and $\lambda_i \neq \lambda_j$. Thus, for $l_1 = l$, we have

$$\begin{aligned} H_1(z) = \sum_{j=1}^{z} \Bigg\{ & l \frac{(1-\lambda_1)e^{-\theta_1-\lambda_1 j}[(\theta_1+\lambda_1 j)^j - (\lambda_1 j)^j]}{(1-e^{-\theta_1})j!} \\ & + (1-l)\frac{(1-\lambda_2)e^{-\theta_2-\lambda_2 j}[(\theta_2+\lambda_2 j)^j - (\lambda_2 j)^j]}{(1-e^{-\theta_2})j!} \Bigg\} \end{aligned} \tag{19}$$

and

$$\begin{aligned} H_2(z) = \sum_{j=1}^{z} \Bigg\{ & l \frac{(1-\gamma_1)e^{-\delta_1-\gamma_1 j}[(\delta_1+\gamma_1 j)^j - (\gamma_1 j)^j]}{(1-e^{-\delta_1})j!} \\ & + (1-l)\frac{(1-\gamma_2)e^{-\delta_2-\gamma_2 j}[(\delta_2+\gamma_2 j)^j - (\gamma_2 j)^j]}{(1-e^{-\delta_2})j!} \Bigg\}. \end{aligned} \tag{20}$$

Now, from (18)–(20), we obtain the following equations:

$$\begin{aligned} & d_1 \sum_{j=1}^{z} \frac{(1-\lambda_1)e^{-\theta_1-\lambda_1 j}[(\theta_1+\lambda_1 j)^j - (\lambda_1 j)^j]}{(1-e^{-\theta_1})j!} \\ & + d_2 \sum_{j=1}^{z} \frac{(1-\gamma_1)e^{-\delta_1-\gamma_1 j}[(\delta_1+\gamma_1 j)^j - (\gamma_1 j)^j]}{(1-e^{-\delta_1})j!} = 0 \end{aligned} \tag{21}$$

and

$$\begin{aligned} & d_1 \sum_{j=1}^{z} \frac{(1-\lambda_2)e^{-\theta_2-\lambda_2 j}[(\theta_2+\lambda_2 j)^j - (\lambda_2 j)^j]}{(1-e^{-\theta_2})j!} \\ & + d_2 \sum_{j=1}^{x} \frac{(1-\gamma_2)e^{-\delta_2-\gamma_2 j}[(\delta_2+\gamma_2 j)^j - (\gamma_2 j)^j]}{(1-e^{-\delta_2})j!} = 0. \end{aligned} \tag{22}$$

Solving (21) and (22), we get

$$\begin{aligned} & d_1 \sum_{j=1}^{z} \frac{(1-\lambda_1)(1-\gamma_2)e^{-\theta_1-\lambda_1 j-\delta_2-\gamma_2 j}[(\theta_1+\lambda_1 j)^j - (\lambda_1 j)^j][(\delta_2+\gamma_2 j)^j - (\gamma_2 j)^j]}{(1-e^{-\theta_1})(1-e^{-\delta_2})j!} \\ & = d_1 \sum_{j=1}^{z} \frac{(1-\lambda_2)(1-\gamma_1)e^{-\theta_2-\lambda_2 j-\delta_1-\gamma_1 j}[(\theta_2+\lambda_2 j)^j - (\lambda_2 j)^j][(\delta_1+\gamma_1 j)^j - (\gamma_1 j)^j]}{(1-e^{-\theta_2})(1-e^{-\delta_1})j!}, \end{aligned} \tag{23}$$

which implies $d_1 = 0$ and, logically, $d_2 = 0$ by (18). Hence, by Theorem 1, it can be concluded that H_1 and H_2 are linearly independent. Now, the argument can be extended to the case of any positive integer g and thus the proof follows. □

Proposition 13. *The pgf of a rv following the* $LZTPMD_g$ *with pmf given in (17) is given by*

$$G(u) = \sum_{i=1}^{g} l_i \frac{(1-\lambda_i)(e^{z\theta_i}-1)}{(e^{\theta_i}-1)(1-\lambda_i z)},$$

where $z = ue^{\lambda_i(z-1)}$.

Proof. Given the pgf of the LZTPD(θ, λ) stated in (13), the proof follows directly from Definition 2. □

5. Estimation

In this section , we explore two popular methods of estimation, namely the method of moments (MM) and maximum likelihood (ML), for estimating the parameters of the LZTPD(θ, λ).

5.1. Maximum Likelihood

Let $Y_1, Y_2, \ldots, Y_n$ be n iid rvs derived from a rv Y following the LZTPD(θ, λ) (so with the pmf given in (6)), and $y_1, y_2, \ldots, y_n$ be n observations. Then, the likelihood function of the parameter vector $\Theta = (\theta, \lambda)$ is given by

$$L_n = L_n(\Theta) = \frac{(1-\lambda)^n e^{-\lambda \sum_{i=1}^{n} y_i} \prod_{i=1}^{n} [(\theta + \lambda y_i)^{y_i} - (\lambda y_i)^{y_i}]}{(e^\theta - 1)^n}.$$

The log-likelihood function is given by

$$\begin{aligned}\mathcal{L}_n = \mathcal{L}_n(\Theta) = \log(L_n) &= n\log(1-\lambda) - \lambda \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} \log[(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}] \\ &\quad - n\log(e^\theta - 1).\end{aligned} \tag{24}$$

The first partial derivatives of $\mathcal{L}_n$ with respect to the parameters θ and λ are, respectively, given by

$$\frac{\partial \mathcal{L}_n}{\partial \theta} = \sum_{i=1}^{n} \frac{y_i(\theta + y_i)^{y_i - 1}}{[(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}]} - \frac{n\, e^\theta}{e^\theta - 1}$$

and

$$\frac{\partial \mathcal{L}_n}{\partial \lambda} = \sum_{i=1}^{n} \frac{y_i^2[(\theta+\lambda y_i)^{y_i-1} - (\lambda y_i)^{y_i-1}]}{(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}} - \sum_{i=1}^{n} y_i - \frac{n}{1-\lambda}.$$

The ML estimate (MLE) (vector) of Θ, say $\hat{\Theta} = (\hat{\theta}, \hat{\lambda})$, is obtained by the solutions of the likelihood equations $\frac{\partial \mathcal{L}_n}{\partial \theta} = 0$ and $\frac{\partial \mathcal{L}_n}{\partial \lambda} = 0$ with respect to θ and λ. With these notations, $\hat{\theta}$ and $\hat{\lambda}$ are also called MLEs of θ and λ, respectively. Analytical solutions to the likelihood equations are not possible. However, one can still compute the MLEs numerically by maximizing the log-likelihood function given in (24) by using the optim function in the *R* programming language under the L-BFGS-B algorithm. The associated *R*-code is provided in Appendix A.

In order to obtain the asymptotic confidence intervals for the parameters θ and λ, we consider the second partial derivatives of $\mathcal{L}_n$ taken at $\hat{\Theta} = (\hat{\theta}, \hat{\lambda})$. In this way, the Hessian matrix of the LZTPD(θ, λ) can be obtained, and it is given by

$$H(\hat{\Theta}) = \left. \begin{pmatrix} \frac{\partial^2 \mathcal{L}_n}{\partial \theta^2} & \frac{\partial^2 \mathcal{L}_n}{\partial \theta \partial \lambda} \\ \\ \frac{\partial^2 \mathcal{L}_n}{\partial \lambda \partial \theta} & \frac{\partial^2 \mathcal{L}_n}{\partial \lambda^2} \end{pmatrix} \right|_{\Theta = \hat{\Theta}},$$

where

$$\frac{\partial^2 \mathcal{L}_n}{\partial \lambda^2} = -\frac{n}{(1-\lambda)^2} + \sum_{i=1}^{n} y_i^3 \left\{ \frac{(y_i-1)[(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}]\left[(\theta+\lambda y_i)^{y_i-2} - (\lambda y_i)^{y_i-2}\right]}{[(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}]^2} \right\}$$
$$- \sum_{i=1}^{n} y_i^3 \left\{ \frac{y_i\left[(\theta+\lambda y_i)^{y_i-1} - (\lambda y_i)^{y_i-1}\right]^2}{[(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}]^2} \right\},$$
$$\frac{\partial^2 \mathcal{L}_n}{\partial \theta^2} = \left\{ \frac{y_i(\theta+\lambda y_i)^{y_i-2}[(y_i-1)\{(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}\} - (\theta+\lambda y_i)^{y_i}]}{[(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}]^2} \right\} + \frac{n}{e^{\theta}(1-e^{-\theta})^2}$$

and

$$\frac{\partial^2 \mathcal{L}_n}{\partial \theta \partial \lambda} = \sum_{i=1}^{n} y_i^2 (\theta+\lambda y_i)^{y_i-1} \left\{ \frac{\{(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}\}(\theta+\lambda y_i)^{-1} - y_i\{(\theta+\lambda y_i)^{y_i-1} - (\lambda y_i)^{y_i-1}\}}{[(\theta+\lambda y_i)^{y_i} - (\lambda y_i)^{y_i}]^2} \right\}.$$

Therefore, the observed Fisher information matrix $J(\hat{\Theta})$ can be obtained by the negative of the Hessian matrix. That is, $J(\hat{\Theta}) = -H(\hat{\Theta})$. Moreover, the variance-covariance matrix of the MLEs is the inverse of the observed Fisher information matrix. It is given as follows:

$$\Sigma = J^{-1}(\hat{\Theta}) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

and $\Sigma_{ij} = \Sigma_{ji}$ for $i \neq j = 1, 2$.

The asymptotic distribution of the random version of $\hat{\Theta}$ follows a normal distribution that has been thoroughly established under the regularity constraints. That is, the random version of $\hat{\Theta} - \Theta$ follows the multivariate normal distribution $N_2(0, \Sigma)$ asymptotically. For $\rho \in (0, 1)$, we calculate the $100 \times (1-\rho)\%$ asymptotic confidence intervals for parameters using the following formulae:

$$\theta \in \left\{ \hat{\theta} \mp Z_{\rho/2}\sqrt{\Sigma_{11}} \right\},\ \lambda \in \left\{ \hat{\lambda} \mp Z_{\rho/2}\sqrt{\Sigma_{22}} \right\},$$

where Z_ρ is the upper ρ_{th} percentile of the standard normal distribution.

5.2. Method of Moments

In this portion, the parameters of the LZTPD are estimated by means of the method of moments (MM). This method's concept is to solve theoretical moments using empirical moments. So we use the first and second sample moments, say m_1 and m_2, respectively. Using this idea, we have

$$m_1 = \mu_1' = \frac{\lambda}{(1-\lambda)^2} + \frac{\theta}{(1-e^{-\theta})(1-\lambda)} \tag{25}$$

and

$$m_2 = \mu_2' = \frac{\lambda+\lambda^2}{(1-\lambda)^4} + \frac{\theta^2(1-\lambda)+\theta}{(1-e^{-\theta})(1-\lambda)^3} - \frac{\theta^2}{(1-e^{-\theta})^2(1-\lambda)^2} + \frac{\lambda}{(1-\lambda)^2} + \frac{\theta}{(1-e^{-\theta})(1-\lambda)}. \tag{26}$$

Solving the above two equations, we get the method of moment estimates (MMEs) of θ and λ, say $\hat{\theta}_{MM}$ and $\hat{\lambda}_{MM}$, respectively, governed by the following equations:

$$\hat{\theta}_{MM} = \frac{m_1(1-\hat{\lambda}_{MM})^2 - \hat{\lambda}_{MM}}{(1-\hat{\lambda}_{MM})^5} \left[(m_2 - m_1^2)(1-\hat{\lambda}_{MM})^4 - (m_1 + \hat{\lambda}_{MM}^2)(1-\hat{\lambda}_{MM})^2 + (m_1(1-\hat{\lambda}_{MM})^2 - \hat{\lambda}_{MM})^2 \right]$$

and

$$\hat{\lambda}_{MM} = \frac{\sqrt{(1+2m_1-q)^2 - 4m_1(m_1-q)} + (1+2m_1-q)}{2m_1},$$

where $q = \frac{\hat{\theta}_{MM}}{1-e^{-\hat{\theta}_{MM}}}$.

6. Generalized Likelihood Ratio Test

In this section, we test the significance of an additional parameter included in the LZTPD using the generalized likelihood ratio test (GLRT). For the details, see [29].

Thus, in order to test the significance of the additional parameter λ of the LZTPD(θ, λ), we take over the GLRT procedure. Here, the null hypothesis is $H_0 : Y$ follows the ZTPD, against the alternative hypothesis: $H_1 : Y$ follows the LZTPD(θ, λ).

In the case of the GLRT, the test statistic is

$$-2\log\Lambda = -2\log\left(\frac{\mathcal{L}_n(\hat{\Theta}^*)}{\mathcal{L}_n(\hat{\Theta})}\right), \tag{27}$$

where $\hat{\Theta}$ is the MLE of $\Theta = (\theta, \lambda)$ with no restrictions and $\hat{\Theta}^*$ is the MLE of Θ under H_0. The random version of the test statistic given in (27) is asymptotically distributed as the chi-square distribution with one degree of freedom.

7. Simulation Study

In this section, we present a brief simulation exercise to assess the limited sample performance of estimates derived using the ML estimation approach by the *R* programming language (see [30]). Here, the iteration process is repeated $N = 1000$ times. The specification of the parameter values is as follows: $\theta = 1$ and $\lambda = 0.5$ (over-dispersion case), and $\theta = 0.5$ and $\lambda = 0.1$ (under-dispersion case). Thus, we computed the average of the mean square error (MSE), and bias using the MLEs. The results obtained are summarized in Table 2 corresponding to the samples of sizes 25, 50, 200, 500, and 1000.

The average bias of the simulated estimates equals $\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\Theta}_i - \Theta)$ and the average MSE of the simulated estimates equals $\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\Theta}_i - \Theta)^2$, in which i is the rank of the considered iteration, $\Theta \in \{\theta, \lambda\}$ and $\hat{\Theta}$ is the MLE of Θ.

Table 2. Estimates of the parameters and the corresponding bias and MSE.

Parameter Set	Sample Size	Parameters	Estimates	MSE	Bias
$\theta = 1, \lambda = 0.5$	$n = 25$	θ	0.7554	0.0598	−0.2464
		λ	0.0379	0.0031	−0.0590
	$n = 50$	θ	0.7592	0.0579	−0.2488
		λ	0.4490	0.0026	−0.0510
	$n = 200$	θ	0.7801	0.0483	−0.2199
		λ	0.4601	0.0015	−0.0399
	$n = 500$	θ	0.8023	0.0390	−0.1977
		λ	0.4780	0.0004	−0.0220
	$n = 1000$	θ	0.9567	0.0018	−0.0433
		λ	0.4901	0.0001	−0.0099
$\theta = 0.5, \lambda = 0.1$	$n = 25$	θ	0.3646	0.0183	−0.1354
		λ	0.0300	0.0049	−0.0700
	$n = 50$	θ	0.3699	0.0169	−0.1301
		λ	0.0542	0.0020	−0.0458
	$n = 200$	θ	0.3978	0.0104	−0.1022
		λ	0.0801	0.0003	−0.0199
	$n = 500$	θ	0.4736	0.0006	−0.0264
		λ	0.0891	0.0001	−0.0109
	$n = 1000$	θ	0.4983	0.00001	−0.0017
		λ	0.0940	0.00003	−0.0060

From Table 2, it can be observed that the MSE in both the cases of the parameter sets is in decreasing order as the sample size increases, and also, the MLEs of the parameters come closer to the original parameter values as the sample size increases.

8. Lagrangian Zero Truncated Poisson Regression Model

The first thought that comes to mind when modelling a discrete response variable with associated covariates is a Poisson regression model. Except in the case of equi-dispersion, it can be seen that the Poisson regression provides erroneous findings when the response variable is over-dispersed or under-dispersed. Several models have been proposed to deal with these dispersions, including mixed Poisson models, generalized Poisson models, etc. However, we frequently encounter cases in which count data has no zeros; refs. [31,32] provided examples of length of hospital stay data. In this case, the ZTPRM performs well. Here, we introduce a novel count regression model called the LZTPRM, which is based on the LZTPD and provides additional options for predicting over-dispersed and under-dispersed zero truncated counts. Finally, we see that the LZTPRM performs well compared to the ZTPRM, ZTGPRM, and IPRM in the case of length of hospital stay data.

To link the covariates to the mean of the response rv X, we use the log-link function such that

$$\mu = \mathrm{E}(X) = e^{y^T\alpha}, \;\; i = 1, 2, \ldots, n, \tag{28}$$

where $y^T = (y_1, y_2, \ldots, y_k)$ is the vector of k covariates and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is the unknown vector of regression coefficients. Now, by considering the notations involved for the LZTPD(θ, λ) and the following re-parametrization:

$$\lambda = \frac{\sqrt{\left[\frac{\theta}{1-e^{-\theta}} - (2\mu+1)\right]^2 - 4\mu\left(\mu - \frac{\theta}{1-e^{-\theta}}\right)} - \left[\frac{\theta}{1-e^{-\theta}} - (2\mu+1)\right]}{2\mu},$$

the pmf of the LZTPD can be re-expressed as

$$h(x|\theta,\mu) = \frac{1-V}{(e^{\theta}-1)x!}\, e^{-Vx}\,[(\theta+Vx)^x - (Vx)^x], \tag{29}$$

where

$$V = \frac{\sqrt{\left[\frac{\theta}{1-e^{-\theta}} - (2\mu+1)\right]^2 - 4\mu\left(\mu - \frac{\theta}{1-e^{-\theta}}\right)} - \left[\frac{\theta}{1-e^{-\theta}} - (2\mu+1)\right]}{2\mu},$$

$\theta > 0$ and $\mu \geq \max\left\{\frac{\theta}{1-e^{-\theta}}, \frac{1}{4}\left(\frac{2\theta}{1-e^{-\theta}} - \frac{\theta^2}{(1-e^{-\theta})^2} - 1\right)\right\}$. Based on n independent observations of the regression model, say $(x_1, y_1^T), (x_2, y_2^T), \ldots, (x_n, y_n^T)$, and substituting (28) in (29), $X_i|y_i^T$ follows the LZTPRM(θ, μ_i), where $y_i^T = (y_{i1}, y_{i2}, \ldots, y_{ik})$, with the following pmf:

$$h(x_i|y_i^T,\theta) = \frac{1-W_i}{(e^{\theta}-1)x_i!}\, e^{-W_i\, x_i}\,[(\theta+W_i\, x_i)^{x_i} - (W_i\, x_i)^{x_i}],$$

where

$$W_i = \frac{\sqrt{\left[\frac{\theta}{1-e^{-\theta}} - (2\, e^{y_i^T\alpha}+1)\right]^2 - 4\, e^{y_i^T\alpha}\left(e^{y_i^T\alpha} - \frac{\theta}{1-e^{-\theta}}\right)} - \left[\frac{\theta}{1-e^{-\theta}} - (2e^{y_i^T\alpha}+1)\right]}{2e^{y_i^T\alpha}}.$$

The log-likelihood function of the LZTPRM based on a sample of n independent observations $(x_1, y_1^T), (x_2, y_2^T), \ldots, (x_n, y_n^T)$ is expressed as

$$\begin{aligned}\log L = &\sum_{i=1}^{n}\{\log(1-W_i) - W_i x_i + \log[(\theta+W_i x_i)^{x_i} - (W_i x_i)^{x_i}] - \log\, x_i!\}\\ &- n\log(e^{\theta}-1).\end{aligned} \tag{30}$$

As in Section 5, for finding the MLEs of the parameters, we use the optim function in the *R* programming language under the L-BFGS-B algorithm.

9. Applications and Empirical Study

This section contains three real datasets to demonstrate the empirical importance of the LZTPD. The first two datasets are used to compare the LZTPD's modeling ability to that of some competing models, while the third dataset is for the regression study. The form of the hrf of the datasets is determined using a graphical method based on Total Time on Test (TTT). If the empirical TTT plot is convex, concave, convex then concave, and concave then convex, then the form of the associated hrf is decreasing, increasing, bathtub shape, and upside-down bathtub shape, respectively (see [33]). The following distributions are considered to demonstrate the potential advantage of the LZTPD:

- The ZTPD proposed by [1], and defined by the following pmf:

$$h_6(y) = \frac{\theta^y}{y!(e^\theta - 1)}, \quad y = 1, 2, \ldots,$$

  with $\theta > 0$.
- The ZTGPD proposed by [21], and specified by the following pmf:

$$h_7(y) = \frac{\theta(\theta + \lambda y)^{y-1} e^{-\lambda y}}{y!(e^\theta - 1)}, \quad y = 1, 2, \ldots,$$

  wih $\theta, \lambda > 0$.
- The IPD elaborated by [3], and indicated by the following pmf:

$$h_8(y) = \frac{[(1+\varphi)^y - \varphi^y]\zeta^y}{e^{\varphi\zeta}(e^\zeta - 1)y!}, \quad y = 1, 2, \ldots$$

  with $\zeta > 0$ and $\varphi \geq 0$.
- The zero truncated discrete Shankar distribution (ZTDSD) proposed by [34], and defined by the following pmf:

$$h_9(y) = \frac{(\theta^2 + 1 + \theta y)(1 - e^{-\theta}) - \theta e^{-\theta}}{(\theta^2 + \theta + 1)}, \quad y = 1, 2, \ldots,$$

  with $\theta > 0$.
- The two-parameter zero truncated Poisson-Lindley distribution (ZTPLD) introduced by [35], and indicated by the following pmf:

$$h_{10}(y) = \frac{\theta^2}{\theta^2 + 2\theta\alpha + \theta + \alpha} \frac{\alpha y + \theta + \alpha + 1}{(\theta + 1)^y}, \quad y = 1, 2, \ldots,$$

  with $\theta, \alpha > 0$.
- The zero truncated generalized negative binomial distribution (ZTGNBD) proposed by [36], and defined by the following pmf:

$$h_{11}(y) = \frac{\theta}{\theta + \lambda y} \binom{\theta + \lambda y}{y} \frac{\alpha^y (1-\alpha)^{\theta + \lambda y - y}}{1 - (1-\alpha)^\theta}, \quad y = 1, 2, \ldots,$$

  with θ, $\lambda > 0$ and $0 < \alpha < 1$.

The MLEs of the parameters, negative estimated Log Likelihood ($-\hat{\mathcal{L}}_n$), Akaike information criterion (AIC), Bayesian information criterion (BIC), and the χ^2 statistic value are now calculated for the first two datasets. The best model is the one with minimum values for its model adequacy measures, such as the AIC and BIC, and the best fitted model is the one having a minimum value for the goodness of fit statistic (χ^2).

9.1. University Course Enrollments

Ref. [37] provided the following data on student enrollments in selected senior mathematics and statistics courses at the University of Calgary over a five-year period:
1, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8, 9, 9, 9, 9, 9, 9, 13, 13, 14, 16, 16, 17, 17, 17, 18, 19, 20, 24, 24, 24, 24, 27, 31, 33, 35, 37.

Table 3 shows the descriptive measures of the data, which include sample size n, minimum (min), first quartile (Q_1), median (Md), third quartile (Q_3), maximum (max), and interquartile range (IQR).

Table 3. Descriptive statistics for the university course enrollments data.

Statistics	n	min	Q_1	Md	Q_3	max	IQR
Values	56	1	4	7	17	37	13

The empirical *IOD* of the data is equal to 7.7131. As a result, our model employed to describe the current data set is capable of dealing with over-dispersion. In addition, Figure 4 shows an empirical TTT plot of the data and it reveals an increasing hrf.

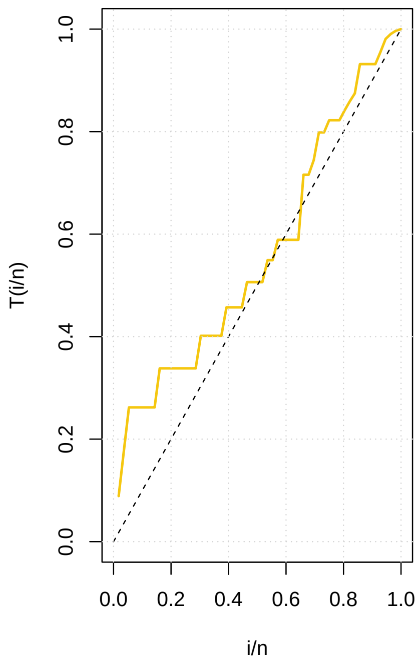

Figure 4. TTT plot for the university course enrollments data.

The MLEs, model adequacy measures, and χ^2-value for the data are computed. They are given in Table 4. From Table 4, it is clear that the LZTPD has better performance compared to all the other competing models considered here since it has the smallest χ^2-value and model adequacy measures.

Now, the Hessian matrix related to the MLEs is given as

$$H(\hat{\Theta}) = \begin{pmatrix} 1392.5626 & 57.7502 \\ 57.7502 & 4.8318 \end{pmatrix}$$

and the quadrated estimated variance-covariance matrix is

$$\Sigma = \begin{pmatrix} 0.0014 & -0.0170 \\ -0.0170 & 0.4103 \end{pmatrix}.$$

Table 4. MLEs, model adequacy measures and χ^2 value for the university course enrollments data.

Model	MLEs	$-\hat{\mathcal{L}}_n$	χ^2	AIC	BIC
ZTPD	$\theta = 11.25$	300.1443	2997.476	602.2886	604.3140
IPD	$\varphi = 11.249$ $\zeta = 3.3197 \times 10^{-8}$	300.1443	2998.871	604.2886	608.3393
ZTDSD	$\theta = 0.1709$	187.8862	144.3541	377.7725	381.7978
ZTPLD	$\theta = 0.1785$ $\alpha = 0.2975$	187.856	144.7294	379.7121	383.7628
ZTGPD	$\theta = 3.8053$ $\lambda = 0.6540$	186.8137	154.0068	377.6273	381.6780
ZTGNBD	$\theta = 20.5554$ $\lambda = 4.0554$ $\alpha = 0.1687$	186.8197	153.9518	379.6394	385.7154
LZTPD	$\theta = 2.5878$ $\lambda = 0.6113$	186.7358	126.0983	377.4716	381.5223

It is noticed that the determinant value of the observed information matrix $J(\hat{\Theta})$ is non-zero and hence meets the non-singularity conditions of the information matrix.

For comparison purposes only, we compute the MMEs of the LZTPD parameters for the university course enrollments data, and they are obtained as $\hat{\theta}_{MM} = 2.0154$ and $\hat{\lambda}_{MM} = 0.5815$. It is concluded that MMEs and MLEs are approximately equal.

In the case of the GLRT, the calculated value based on the test statistic (27) is $2(-186.7358$ + 300.1443) = 226.817 (p-value $=$ 0.00002). As a result, at any level >0.00002, the null hypothesis is rejected in favor of the alternative hypothesis. Hence, we conclude that the additional parameter λ in the LZTPD is significant in the light of the test procedure outlined in Section 6. Furthermore, the approximate 95% confidence intervals for θ and λ are given by $(2.3321, 2.7431)$ and $(0.5374, 0.6853)$, respectively.

9.2. Demographic Data

Secondly, we utilize the data available in [38] as a demographic study, which represent the number of mothers of the completed fertility having experienced at least one child death. Table 5 provides the descriptive measures of the data, such as n, min, Q_1, Md, Q_3, max and IQR.

Table 5. Descriptive statistics for the demographic data.

Statistics	n	min	Q_1	Md	Q_3	max	IQR
Values	135	1	1	1	2	6	1

The empirical *IOD* of the data is equal to 0.6787. As a result our model is employed to explain this data set. Furthermore, Figure 5 shows an empirical TTT plot of the data. It can be concluded that the data have an increasing hrf.

Table 6 displays the MLEs, model adequacy measures, and χ^2-value for the data. From Table 6, it is clear that the LZTPD has better performance compared to all the other competing models considered here since it has the smallest χ^2-value and model adequacy measures.

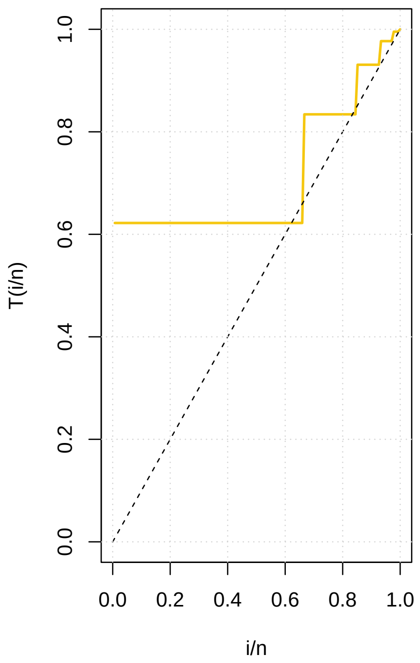

Figure 5. TTT plot for the demographic data.

Table 6. MLEs, model adequacy measures and χ^2-value for the demographic data.

Model	MLEs	$-\hat{\mathcal{L}}_n$	χ^2	AIC	BIC
ZTPD	$\theta = 1.0381$	150.0619	7.9012	302.12	305.029
IPD	$\varphi = 1.0382$ $\zeta = 4.5998 \times 10^{-10}$	150.0619	14.863	304.70	309.93
ZTDSD	$\theta = 0.9999$	148.8624	12.1887	299.7248	302.6301
ZTPLD	$\theta = 1.6466$ $\alpha = 0.00038$	143.8747	2.8235	291.7493	297.5599
ZTGPD	$\theta = 0.2838$ $\lambda = 0.2855$	143.3546	1.746	290.709	296.520
ZTGNBD	$\theta = 0.2041$ $\lambda = 1.0002$ α=0.5281	143.2366	1.5554	292.4731	301.189
LZTPD	$\theta = 4.5593 \times 10^{-8}$ $\lambda = 0.2112$	143.0373	1.304	290.6747	296.4852

Now, the Hessian matrix related to the MLEs is obtained as

$$H(\hat{\Theta}) = \begin{pmatrix} 500.0923 & 95.5925 \\ 95.5925 & 13.5116 \end{pmatrix}$$

and the quadrated estimated variance-covariance matrix is

$$\Sigma = \begin{pmatrix} 0.0020 & -0.0108 \\ -0.0108 & 0.0740 \end{pmatrix}.$$

Here again, the observed information matrix's determinant value $J(\hat{\Theta})$ is non-zero, indicating that the non-singularity condition is satisfied.

For comparison purposes only, we compute the MMEs of the LZTPD parameters for the demographic data, and they are obtained as $\hat{\theta}_{MM} = 4.1979 \times 10^{-8}$ and $\hat{\lambda}_{MM} = 0.2361$. It is concluded that MMEs and MLEs are approximately equal.

In the case of GLRT, the computed value based on the test statistic (27) is $2(-143.0373 + 150.0619) = 14.0492$ (p-value $= 0.0001$). As a result, the null hypothesis is rejected in favor of the alternative hypothesis at any level >0.0001. Hence, we conclude that the additional parameter λ in the LZTPD is significant in the light of the test procedure outlined in Section 6. The approximate 95% confidence intervals for θ and λ are given by $(8.7193 \times 10^{-9}, 9.7470 \times 10^{-9})$ and $(0.5787, 0.7276)$, respectively.

9.3. Biological Science

The third data set, which is included in the 'azpro' package of the *R* software (also, available in the 'COUNT' package of the *R* software), is about a 1991 Arizona cardiovascular patient. We have 3589 patients and the aim is to model the length of hospital stay, say x_i for the ith patient, with the following covariates: y_{i1} represents the cardiovascular procedure (the variable takes the value 1 for CABG procedure and 0 for PTCA procedure), y_{i2} represents the sex of the patients (the variable takes the value 1 for male and 0 for female patients), y_{i3} represents the type of the admission (the variable takes the value 1 for urgent and 0 for electiive), and y_{i4} represents the age of the patients (the variable takes the value 1 for the age > 75 and 0 for the age ≤ 75).

The empirical *IOD* of the response variable is calculated as 5.432. So the response variable is over-dispersed. Therefore, the LZTPRM is able to handle this over-dispersion, with the configuration

$$\mu_i = e^{\alpha_0 + \alpha_1 y_{i1} + \alpha_2 y_{i2} + \alpha_3 y_{i3} + \alpha_4 y_{i4}}.$$

The following regression models are used to compare the LZTPRM:

- the ZTPRM given in [39];
- the ZTGPRM given in [40];
- the IPRM given in [18].

Table 7 compares the LZTPRM's performance to that of the ZTPRM, ZTGPRM, and IPRM, as well as summaries based on the real data set, such as standard errors (SEs), p-values, negative log-likelihood (-logL), and AIC values. According to Table 7, the LZTPRM has the lowest values across all model selection criteria, indicating that it is the best count regression model among the ZTPRM, ZTGPRM, and IPRM.

Table 7. The results of the regression models for the length of hospital stay data (SEs in brackets).

Covariates	ZTPRM		ZTGPRM		IPRM		LZTPRM	
	Estimate	***p*-Value**	**Estimate**	***p*-Value**	**Estimate**	***p*-Value**	**Estimate**	***p*-Value**
α_0	1.2367 (0.0213)	<0.001	1.1961 (0.0160)	<0.001	1.1981 (0.0019)	<0.001	2.0181 (0.0021)	<0.001
α_1	0.5609 (0.0305)	<0.001	0.5931 (0.0280)	<0.001	0.5751 (0.0345)	<0.001	0.1361 (0.0145)	<0.001
α_2	−0.0739 (0.0365)	<0.001	−0.0781 (0.0156)	<0.001	−0.0766 (0.0019)	<0.001	−0.0141 (0.0232)	<0.001
α_3	0.1452 (0.0168)	<0.001	0.1499 (0.0255)	<0.001	0.2908 (0.0217)	<0.001	0.0982 (0.0251)	<0.001
α_4	0.0934 (0.0134)	<0.001	0.0991 (0.0346)	<0.001	0.1352 (0.0109)	<0.001	0.0142 (0.0019)	<0.001
−logL	6921.34		6629.25		6579.37		6494.85	
AIC	13854.68		13272.50		13172.74		13003.70	

10. Discussion

10.1. Brief Summary

In the case of over-dispersion and under-dipersion, new count models must be discovered, which could provide additional options for better fitting real datasets by selecting the appropriate models for the situation. We developed a new over-dispersed count model and analysed its regression characteristics in this regard. The primary reason for creating this model is also explained. We discovered that the suggested model outperformed the compared models in every way, including its main competitors: the ZTPD and IPD.

10.2. This Work

A novel discrete distribution, the LZTPD, is developed and thoroughly examined. The median, mode, pgf, cgf, factorial moments, mean, variance, CV, skewness, and kurtosis were all given precise formulations. The distribution parameters were estimated using the classical ML and MM techniques. A simulation study on the MLEs was also conducted. On the basis of a real data set, a new LZTPD-based regression model for count data is developed and compared to competitive regression models. The new model was demonstrated using three real-world datasets: one with university course enrollments data, another with demographic data, and the third with healthcare data.

10.3. Contributions and Limitations

A new discrete distribution with its own count model and its own regression model is proposed in this work. We feel the proposed models are suitable for the in-depth analysis of data in the domains of demography and health care, and we hope that they can be applied to other fields of study as well. The proposed distribution's potential shortcoming is the inability to display a bimodal character.

10.4. Future Work

Based on the demand of applied scientists for our proposed LZTPD, one may explore more features, such as its generalizations using conventional ideas as well as its bivariate and multivariate extensions. This requires significant investigation, which we will delegate to further research.

11. Concluding Remarks

The LZTPD is suitable in both under-dispersed and over-dispersed count datasets, whereas the IPD is only useful in under-dispersed cases. Several key LZTPD features have been determined, and it has been observed that they are more flexible than those of the IPD. The LZTPD is compared to the well-known IPD and a few other competing distributions, and it is discovered that the LZTPD outperforms competing models for the datasets under consideration.

Author Contributions: Conceptualization, M.R.I., C.C., D.S.S., M.M. and R.M.; methodology, M.R.I., C.C., D.S.S., M.M. and R.M.; software, M.R.I., C.C., D.S.S., M.M. and R.M.; validation, M.R.I., C.C., D.S.S., M.M. and R.M.; formal analysis, M.R.I., C.C., D.S.S., M.M. and R.M.; investigation, M.R.I., C.C., D.S.S., M.M. and R.M.; resources, M.R.I., C.C., D.S.S., M.M. and R.M.; data curation, M.R.I., C.C., D.S.S., M.M. and R.M.; writing—original draft preparation, M.R.I., C.C., D.S.S., M.M. and R.M.; writing—review and editing, M.R.I., C.C., D.S.S., M.M. and R.M.; visualization, M.R.I., C.C., D.S.S., M.M. and R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: On the initial version of the manuscript, we appreciate the constructive comments from the three reviewers and the associate editor.

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

Abbreviations

The following abbreviations are used in this manuscript:

LZTPD	Lagrangian Zero Truncated Poisson Distribution
ZTPD	Zero Truncated Poisson Distribution
IPD	Intervened Poisson Distribution
LF	Lagrangian Family
DLF	Discrete Lagrangian Family
LZTPRM	Lagrangian Zero Truncated Poisson Regression Model
ZTPRM	Zero Truncated Poisson Regression Model
ZTGPRM	Zero Truncated Generalized Poisson Regression Model
IPRM	Intervened Poisson Regression Model
pmf	Probability Mass Function
hrf	Hazard Rate Function
IOD	Index Of Dispersion
pgf	Probability Generating Function
mgf	Moment Generating Function
cgf	Cumulant Generating Function
CV	Coefficient of Variation
iid	independent and identically distributed
rv	random variable
ML	Maximum Likelihood
MLEs	Maximum Likelihood Estimates
MM	Method of Moments
MMEs	Method of Moments Estimates
GLRT	Generalized Likelihood Ratio Test
MSE	Mean Squared Error
LZTPMD_g	Lagrangian Zero Truncated Poisson Mixture Distribution with g components
TTT	Total Time on Test
ZTGPD	Zero Truncated Generalized Poisson Distribution
ZTDSD	Zero Truncated Discrete Shankar Distribution
ZTPLD	Zero Truncated Poisson Lindley Distribution
ZTGNBD	Zero Truncated Generalized Negative Binomial Distribution
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
IQR	Inter Quartile Range
Md	Median
min	Minimum
max	Maximum
SE	Standard Error

Appendix A

Below is the main *R*-code for determining the MLEs of the LZTPD parameters, as well as the model adequacy measures.

```
library(fitdistrplus)

dfn <- function(y, theta, lambda){
d <- (exp(-lambda*y)/(factorial(y)*(exp(theta)-1)))
* (theta *((theta+(lambda*y)^(y))-(lambda*y)^(y)))
return(d)
}
```

```
pfn <- function(q,theta,lambda){
cumsum(dfn(q,theta,lambda))
}
#
pfn(x,3,0.4)
#
pre <- prefit(x, "fn", "mle", list(theta=0.1, lambda=0.1),
lower=c(0, 0), upper = c(Inf, 1))

fit.fn <- fitdist(x, "fn",
start = list(theta = pre$theta, lambda = pre$lambda),
optim.method = "L-BFGS-B", lower=c(0, 0), upper = c(Inf, 1),
discrete = TRUE)

summary(fit.fn)
gofstat(fit.fn)
```

References

1. Cohen, A.C. Estimating parameters in a conditional Poisson distribution. *Biometrics* **1960**, *16*, 203–211. [CrossRef]
2. Singh, J. A characterization of positive Poisson distribution and its application. *SIAM J. Appl. Math.* **1978**, *34*, 545–548. [CrossRef]
3. Shanmugam, R. An intervened Poisson distribution and its medical application. *Biometrics* **1985**, *41*, 1025–1029. [CrossRef] [PubMed]
4. Shanmugam, R. An inferential procedure for the Poisson intervention parameter. *Biometrics* **1992**, *48*, 559–565. [CrossRef] [PubMed]
5. Kumar, C.S.; Shibu, D.S. Modified intervened Poisson distribution. *Statistica* **2011**, *71*, 489–499.
6. Singh, B.P.; Dixit, S.; Shukla, U. An alternative to intervened Poisson distribution for prevalence reduction. *J. Math. Stat. Sci.* **2016**, *2016*, 730–740.
7. Lagrange, J.L. *Mécanique Analytique*; Jacques Gabay: Paris, France, 1788.
8. Consul, P.C.; Shenton, L.R. Use of Lagrange expansion for generating generalized probability distributions. *SIAM J. Appl. Math.* **1972**, *23*, 239–248. [CrossRef]
9. Consul, P.C.; Shenton, L.R. Some interesting properties of Lagrangian distributions. *Commun. Stat.* **1973**, *2*, 263–272. [CrossRef]
10. Mohanty, S.G. On a generalized two- coin tossing problem. *Biom. Z.* **1966**, *8*, 266–272. [CrossRef]
11. Consul, P.C.; Famoye, F. Lagrangian Katz family of distributions. *Commun. Stat. Theory Methods* **1996**, *25*, 415–434. [CrossRef]
12. Berg, K.; Nowicki, K. Statistical inference for a class of modified power series distribution with applications to random mapping theory. *J. Stat. Plan. Inference* **1991**, *28*, 247–261. [CrossRef]
13. Li, S.; Black, D.; Lee, C.; Famoye, F. Dependence models arising from the Lagrangian probability distributions. *Commun. Stat. Theory Methods* **2010**, *29*, 1729–1742. [CrossRef]
14. Innocenti, A.R.; Fox, O.; Chibbaro, S. A Lagrangian probability density-function model for collisional turbulent fluid particle flows i. Model derivation. *J. Fluid Mech.* **2019**, *862*, 449–489. [CrossRef]
15. Dobson, A.J.; Dobson, A. *An Introduction to Generalized Linear Models*, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2001.
16. Long, J.S.; Freese, J. *Regression Models for Categorical Dependent Variables Using Stata*, 2nd ed.; Stata Press: College Station, TX, USA, 2005.
17. Shaw, D. On-site samples regression problems of non-negative integers, truncation, and endogenous stratification. *J. Econom.* **2005**, 37, 211–223. [CrossRef]
18. Shibu, D.S. On Intervened Poisson Distribution and Its Generalization. An Unpublished Ph. D. Thesis Submitted to the University of Kerala, Thiruvananthapuram, India, 2013.
19. Janardan, K.G.; Rao, B.R. Lagrangian distributions of second kind and weighted distributions. *SIAM J. Appl. Math.* **1983**, *43*, 302–313. [CrossRef]
20. Janardan, K.G. A wider class of Lagrange distributions of the second kind. *Commun. Stat. Theory Methods* **1997**, *26*, 2087–2097. [CrossRef]
21. Consul, P.C.; Famoye, F. The truncated generalized Poisson distribution and its estimation. *Commun. Stat. Theory Methods* **1989**, *18*, 3635–3648. [CrossRef]
22. Jain, G.C. A Linear function Poisson distribution. *Biom. J.* **1975**, *17*, 501–506. [CrossRef]
23. Consul, P.C.; Famoye, F. *Lagrangian Probability Distributions*; Birkhaüser: New York , NY, USA, 2006.
24. Consul, P.C.; Jain, G.C. A generalization of the Poisson distribution. *Technometrics* **1973**, *15*, 791–799. [CrossRef]

25. Gupta, R.C. Modified power series distribution and some of its applications. *Sankhya Ser. B* **1974**, *35*, 288–298.
26. McLachlan, G.; Peel, D. *Finite Mixture Models*; Wiley: Hoboken, NJ, USA, 2000.
27. Kumar, C.S.; Shibu, D.S. On finite mixtures of modified intervened Poisson distribution and its applications. *J. Stat. Theory Appl.* **2014**, *13*, 344–355. [CrossRef]
28. Titterington, D.M.; Smith, A.F.; Markov, U.E. *Statistical Analysis of Finite Mixture Distributions*; Wiley: Hoboken, NJ, USA, 1985.
29. Rao, C.R. Minimum variance and the estimation of several parameters. *Math. Proc. Camb. Philos.* **1947**, *43*, 280–283. [CrossRef]
30. R Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria. Available online: https://www.R-project.org/ (accessed on 6 September 2021).
31. Hardin, J.; Hilbe, J. *Generalized Linear Models and Extensions*, 2nd ed.; StatCorp LP Texas, A Stata Press Publication: College Station, TX, USA, 2007.
32. Hilbe, J.M. *Negative Binomial Regression*; Cambridge University Press: Cambridge, UK, 2007.
33. Aarset, M.V. How to identify a bathtub hazard rate. *IEEE Trans Reliab.* **1987**, *36*, 106–108. [CrossRef]
34. Borah, M.; Saikaia, K.R. Zero- tuncated discrete Shankar distribution and its applications. *Biom. Biostat. Int. J.* **2017**, *5*, 232–237.
35. Shanker, R.; Shukla, K.K. A zero- truncated two-parameter Poisson-Lindley distribution with an application to biological science. *Turk. Klin. J. Biostat.* **2017**, *9*, 85–95. [CrossRef]
36. Famoye, F.; Consul, P.C. The truncated generalized negative binomial distribution. *J. Appl. Stat. Sci.* **1993**, *1*, 141–157.
37. Huang, M.; Fung, K.Y. Intervened truncated Poisson distribution. *Sankhya Ser. B* **1989**, *51*, 302–310.
38. Shanker, R.; Fesshaye, H.; Selvaraj, S.; Yemane, A. On zero-truncation of Poisson and Poisson-Lindley distributions and their applications. *Biom. Biostat. Int. J.* **2015**, *2*, 168–181. [CrossRef]
39. Wang, Y.; Ong, M.; Liu, H. Compare Predicted Counts between Groups of Zero Truncated Poisson Regression Model based on Recycled Predictions Method; In Proceedings of the Section on Statistics in Epidemiology—JSM 2011, Miami Beach, FL, USA, 30 July–4 August 2011.
40. Zhao, W.; Feng, Y.; Li, Z. Zero-truncated generalized Poisson regression model and its score tests. *J. East China Norm. Univ. Sci.* **2010**, *1*, 17–23.