Three-Inflated Poisson Distribution and its Application in Suicide Cases of India During Covid-19 Pandemic

Tousifur Rahman[1] · Partha Jyoti Hazarika[1] · M. Masoom Ali[2] ·
Manash Pratim Barman[1]

Abstract

Inflated models are generally used whenever there is an excess number of frequencies at particular count. In this study, a three-inflated Poisson (ThIP) distribution is proposed by mixing the Poisson distribution and a distribution to a point mass at three. Some of its distribution properties and reliability characteristics are studied. A simulation study is carried out to see the performance of the MLEs. In India Covid-19 implications on mental health have been abysmal. Covid-19 related suicide data of India during lockdown to the first gradual relaxation of the terms of the total lockdown (unlocking 1.0) are used to examine the appropriateness of the proposed distribution. Likelihood ratio test is used for discriminating between Poisson and the proposed distribution.

Keywords Three-inflated Poisson distribution · Inflated · MLE · Health · Covid-19

Mathematics Subject Classification 60E05 · 60E10 · 62F10

1 Introduction

Statistical modelling is an essential part of data science in various areas of scientific research or decision-making situations. To carry out this approach, selecting an appropriate distribution is one of the significant tasks. Distributional properties are quite important while dealing with huge data in data science (see [1–3] for example). In medical and social sciences, modeling count variables is an everyday exercise [4]. Poisson distribution plays a foremost part in count data analysis. Count data such as number of suicide attempts, number of heart attacks, number of unprotected sexual

✉ Manash Pratim Barman
manash_stats@dibru.ac.in

1 Department of Statistics, Dibrugarh University, Dibrugarh, Assam 786004, India

2 Department of Mathematical Sciences, Ball State University, Muncie, IN 47306, USA

encounters, number of days of alcohol drinking, the number of days of missing primary activities, number of cigarettes smoked, number of hospitalizations, or number of unhealthy days during a period are common in medical and psychological research [5–7]. Poisson distribution is widely used to model such type of count data [6]. As a matter of fact, what happens because of data clustering or certain other factors is a kind of heterogeneity in study populations leading to the creation of extra variability which in turn results in variance that is greater than the mean [6]. In such cases Poisson distribution is inappropriate for data modelling. Medical and public health research often used zero-inflated models when there is a large proportion of zeros [5]. Whenever there is an excess number of frequencies at particular count, inflated models are used. The following are the situations in which too many observed frequencies at particular count data point may occur: (I) All the participants of the study area contemplated are not affected by the Poisson process, so inflation occurs at a particular count; (II) The increase or decrease of the participants of an area at a particular count into the sample may be due to some certain unavoidable problem in the sampling, which leads to inflation or deflation at that particular count; (III) There is no possibility that all the participants in the sample would come into a particular count as an utmost case of situation (II) and this is called truncation at that particular count; (IV) Leading to the data generating procedure of Poisson distribution, we have a sub population as an amalgamation of situations (I) and (III), whereas the part of the population out of this subpopulation in contention not affected by the distribution process furnished excess count in that particular count. Thus the inflated distribution is a mixture between a distribution to a point mass at a particular count and any other count distribution supported by non-negative integers [8].

In statistical literature, the issue of zero moderation in count data has a long history. Neyman [9] and Feller [10] first introduced the concept of zero-inflation when there is a problem of extravagant zeros in the data. Mullahy [11] introduced the zero-inflated Poisson (ZIP) distribution as a mixture of Poisson distribution and a distribution to a point mass at zero, with mixing probability γ, denoted by $ZIP(\lambda, \gamma)$ and the probability mass function of this distribution is given by

$$P(z; \lambda, \gamma) = \begin{cases} \gamma + (1 - \gamma)e^{-\lambda}; & z = 0 \\ (1 - \gamma)\frac{e^{-\lambda}\lambda^z}{z!}; & z > 0 \end{cases}$$

where γ is a zero-inflation parameter $(0 < \gamma < 1)$, $\lambda \geq 0$ and if $\gamma = 0$, ZIP distribution reduces to Poisson distribution.

Using inflated Poisson distribution Pandey [12] narrates a situation of the number of flowers of primulaveris and he exhibits the persistence of Poisson distribution inflated at the point eight, not at zero with the extravagant number of plants with eight flowers. Keeping this example in view, inflated discrete distribution should be studied at any point say k, $(k \neq 0)$.

Johnson et al. [13] described the zero-inflated distribution as a mixture of any count distribution hold up on non-negative integers and distribution at a point mass at zero and is defined as follows: A random variable Z is said to be a zero-inflated distribution if its probability mass function is given by

$$g(z) = \begin{cases} \gamma + (1 - \gamma)h(z; \Theta); \; z = 0 \\ (1 - \gamma)h(z; \Theta); \quad z > 0 \end{cases}$$

where γ is a zero-inflation parameter $(0 < \gamma < 1)$ and $h(z; \Theta)$ is the pmf of Z with a vector of parameter, $\Theta = \{\phi_1, \phi_2, ..., \phi_n\}$.

Gupta et al. [14] studied the structural properties and attained the MLEs of discrete distributions inflated at the point zero. Murat and Szynal [15] studied the discrete distributions inflated at any point j, $(j \geq 0)$, which was extended by the results of Gupta et al. [14]. Najundan et al. [16] estimate the parameters of zero-inflated Poisson model using the method of moments and compared with the maximum likelihood estimators. Using some natural calamities data Beckett et al. [17] studied zero-inflated Poisson model and juxtapose MLEs and MMEs regarding standardized bias and standardized mean squared error. Zero-inflated Poisson distribution, Zero-inflated binomial distribution, Zero-inflated negative binomial distribution and Zero-inflated geometric distribution are characterized by Najundan et al. [18], Najundan et al. [19], Suresh et al. [20] and Nagesh et al. [21].

Alshkaki [22] introduced zero–one-inflated Poisson (ZOIP) distribution and defined as: A random variable Z is said to be a zero–one inflated Poisson distribution, denoted by $ZOIP(\lambda, \gamma, \psi)$, if its probability mass function is given by

$$g(z) = \begin{cases} \gamma + (1 - \gamma - \psi)e^{-\lambda}; & z = 0 \\ \psi + (1 - \gamma - \psi)\lambda e^{-\lambda}; & z = 1 \\ (1 - \gamma - \psi)\frac{e^{-\lambda}\lambda^z}{z!}; & z > 1 \end{cases}$$

where $\gamma \; and \; \psi$ are zero and one inflation parameter $(0 < \gamma < 1, \; 0 < \psi < 1, \; 0 < \gamma + \psi < 1)$ and if $\psi = 0$ ZOIP reduces to ZIP and if $\gamma = 0 \; and \; \psi = 0$ ZOIP reduces to Poisson distribution. He studied its structural properties and estimates its parameters by method of maximum likelihood and method of moments. Using three real data sets constituting, a stillbirths of rabbit's data, an accident insurance claims data and a heavy vehicle traffic accident data, he shows that the zero–one inflated Poisson distribution gives better fitting than the zero inflated Poisson distribution and also MLE provides better estimates than MME.

Singh et al. [23] introduced two-inflated binomial distribution to investigate the mechanism of son preference through the modeling of the pattern of male children in Uttar Pradesh, where family size and sex composition are dominated by strong son reference. Mwalili et al. [24] studied a zero-inflated negative binomial model to gratify extravagant zeros, an extension of negative binomial distribution.

For fitting a data set of excessive zeros, how a zero-inflated Poisson regression is better than a Poisson regression was demonstrated by Lambert [25]. Lambert [25] used a dataset of the number of manufacturing defects on writing boards to juxtapose the models. Many extensions and implementations of zero-inflated Poisson regression were described (for details see [26–30] and among others). Hall [31] procuring a zero-inflated binomial model considered a situation of upper bound count data, altering Lambert [25] methodology, with an example from horticulture. Famoye and Singh [32] propounded a zero-inflated generalized Poisson regression model, an extension

of generalized Poisson regression model. Mwalili et al. [24] studied a zero-inflated negative binomial model to gratify extravagant zeros, an extension of negative binomial distribution.

In this paper, the researchers propose a Three-inflated Poisson (ThIP) distribution along with its distributional properties, reliability characteristics and consider the method of moment estimation (MM) and maximum likelihood estimation (MLE) to estimate its parameters. A simulation study has been conducted to see the behavior of the MLEs. In the application part a real-life data set of Covid-19 related suicides during Lockdown to the first gradual relaxation of the terms of the total lockdown (Unlocking 1.0) is used to examine the pertinence of the proposed distribution. The proposed distribution is compared with PD, ZIPD and ZOIPD using log-likelihood, the Akaike Information Criterion (AIC) [33], the Bayesian Information Criterion (BIC) [34] for model selection and the Kolmogorov–Smirnov (K-S) test [35] P-values for the goodness of fit. Likelihood ratio test is provided to discriminate between Poisson distribution and our proposed distributions.

2 Three Inflated Poisson Distribution (ThIPD)

Definition 1 A random variable Z is said to be a three-inflated Poisson distribution, denoted by $ThIPD(\lambda, \alpha)$ if its probability mass function is given by

$$P(z; \lambda, \alpha) = P(Z = z) = \begin{cases} \alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{3!}; & z = 3 \\ (1 - \alpha)\frac{e^{-\lambda}\lambda^z}{z!}; & z \in \{0, 1, 2, 4, 5, 6, 7...\} \end{cases} \tag{1}$$

where α is a three-inflation parameter $(0 < \alpha < 1)$ and $\lambda \geq 0$.

Some particular cases: When.
(i) $\alpha \to 0$, $ThIPD(\lambda, \alpha)$ reduces to $Poisson(\lambda)$.
(ii) $\alpha \to 0$ and $\lambda \to \infty$, $ThIPD(\lambda, \alpha)$ reduces to Normal distribution.

Theorem 1 The probability of three in ThIPD is larger than that of a general PD.

Proof We know that $\alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{3!} > 0$ and $\frac{e^{-\lambda}\lambda^3}{3!} > 0$.

Hence $1 - \frac{e^{-\lambda}\lambda^3}{3!} > 0 \Rightarrow \frac{e^{-\lambda}\lambda^3}{3!} < 1$.
Now multiplying $-\alpha$ and adding α to both sides, we get

$$\alpha - \alpha\frac{e^{-\lambda}\lambda^3}{3!} > 0$$

Finally, adding both sides by $\frac{e^{-\lambda}\lambda^3}{3!}$, we get

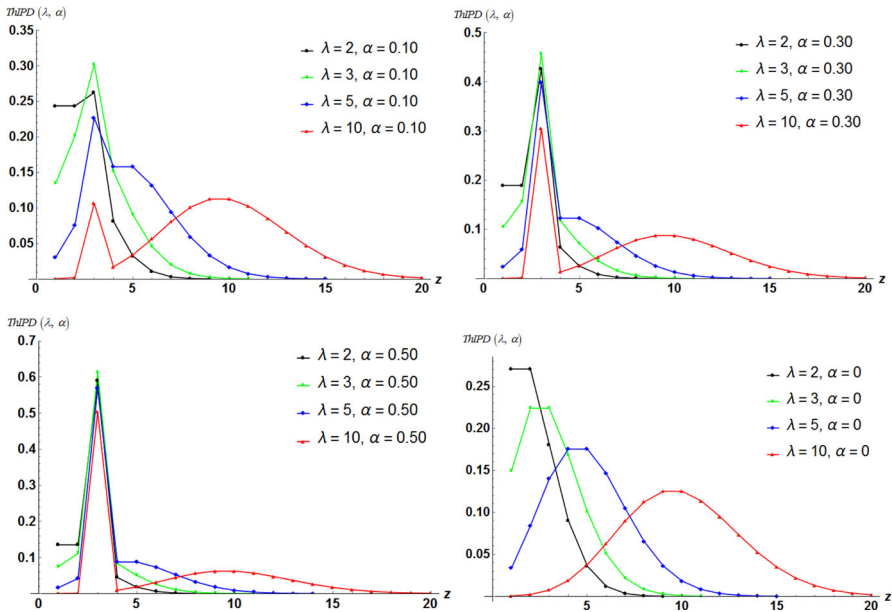$$\alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{3!} > \frac{e^{-\lambda}\lambda^3}{3!}$$

Fig. 1 Shapes of $ThIPD(\lambda, \alpha)$ with different choices of parameter

Hence proved.

The pmf plots of $ThIPD(\lambda, \alpha)$ with different choice of parameters values of λ *and* α to study the variety of shapes are provided in Fig. 1.

It is observed from the plots of $ThIPD(\lambda, \alpha)$ in Fig. 1 that as α increases the curve is peak at $z = 3$ and as α decreases and λ increases the curve tends to normal curve.

3 Distributional Properties

3.1 Moments

Theorem 2 *If* $Z \sim ThIPD(\lambda, \alpha)$, *then its* r^{th} *order moments about zero is as follows*

$$\mu_r' = E(z^r) = 3^r \alpha + (1 - \alpha) \sum_{j=0}^{r} S(r, j) \lambda^j \qquad (2)$$

Proof If $Z \sim ThIPD(\lambda, \alpha)$, then the rth order moments about zero is

$$\mu'_r = E(Z^r)$$

$$= \sum_z z^r p(\lambda, \alpha)$$

$$= 3^r \alpha + (1 - \alpha) \sum_z z^r \frac{e^{-\lambda} \lambda^z}{z!}$$

$$= 3^r \alpha + (1 - \alpha) \sum_{j=0}^{r} S(r, j) \lambda^j \left[\because \sum_z z^r \frac{e^{-\lambda} \lambda^z}{z!} = \sum_{j=0}^{r} S(r, j) \lambda^j \text{ for details see [13]} \right]$$

where $S(r, j)$ is the second kind of Stirling number. Hence proved.

In particular

$$\mu'_1 = E(Z) = 3\alpha + (1 - \alpha)\lambda \tag{3}$$

$$\mu'_2 = E\left(Z^2\right) = 9\alpha + (1 - \alpha)\left(\lambda^2 + \lambda\right) \tag{4}$$

$$\mu'_3 = E\left(Z^3\right) = 27\alpha + (1 - \alpha)\left(\lambda^3 + 3\lambda^2 + \lambda\right) \tag{5}$$

$$\mu'_4 = E\left(Z^4\right) = 81\alpha + (1 - \alpha)\left(\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda\right) \tag{6}$$

Therefore,

$$V(Z) = (1 - \alpha)\left[\alpha(\lambda - 3)^2 + \lambda\right] \tag{7}$$

The plots of mean and variance of the proposed distribution with different choice of parameters to study their variations are shown in the Figs. 2 and 3. From the Figs. 2 and 3 it is clear that as α tends to 0, the mean and variance of the proposed distribution tends to λ, which is the mean and variance of general Poisson distribution.
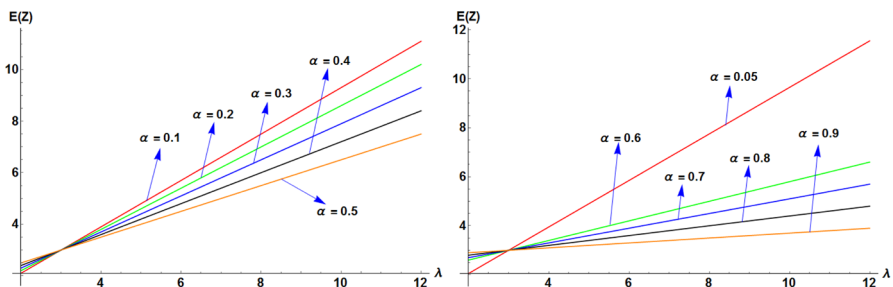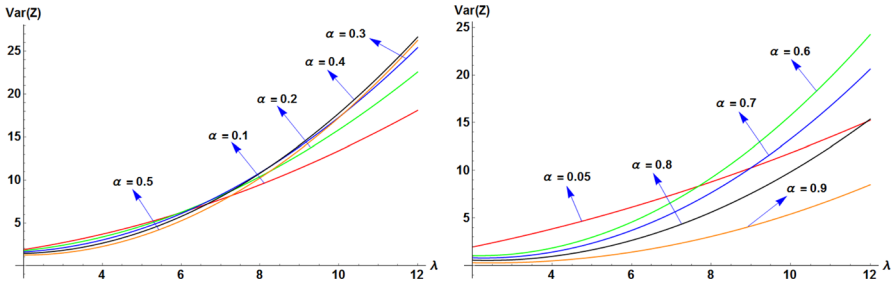

Fig. 2 Plots of mean against λ for different values of α

Fig. 3 Plots of variance against λ for different values of α

3.1.1 Coefficient of Skewness

If $Z \sim ThIPD\,(\lambda,\,\alpha)$, then the Pearson's β_1 coefficient is as follows

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\left[\mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3\right]^2}{\left[\mu_2' - \mu_1'^2\right]^3}$$

$$= \frac{\left[\lambda + \alpha(\lambda - 3)\left(2\alpha(\lambda - 3)^2 - (\lambda - 9)\lambda\right) - 9\right]^2}{(1 - \alpha)\left(\alpha(\lambda - 3)^2 + \lambda\right)^3} \tag{8}$$

The plots of coefficient of skewness of the proposed distribution for different choice of parameters are shown in the Fig. 4

From the Fig. 4 it is observed that β_1 increases as α increases and β_1 tends to zero as λ decreases and increases for $0 < \alpha < 1$.

Remark 1 As $\alpha \to 0\,and\,\lambda \to \infty$, the coefficient of skewness $\beta_1 \to 0$ i.e. the proposed distribution tends to symmetric distribution.

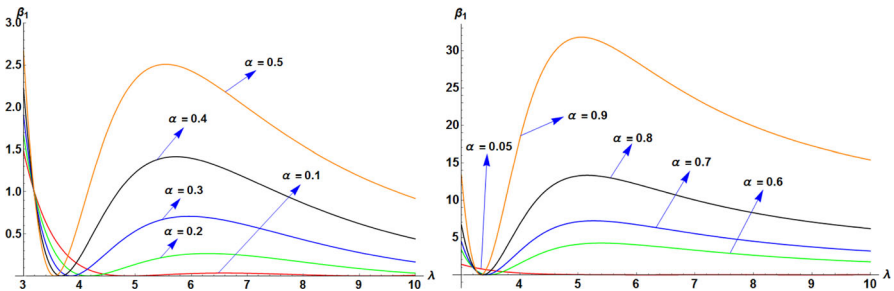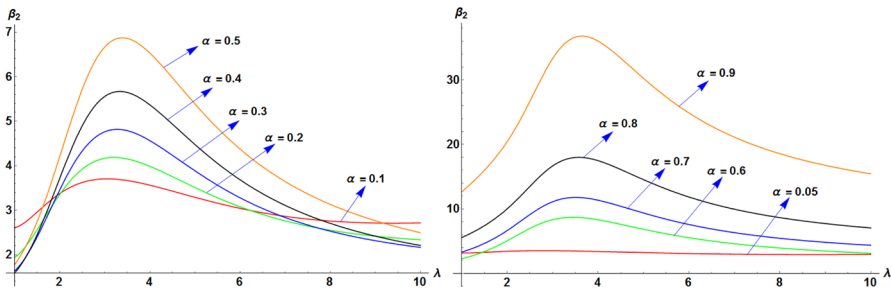**Fig. 4** Plots of β_1 against λ for different values of α

Fig. 5 Plots of β_2 against λ for different values of α

3.1.2 Coefficient of Kurtosis

If $Z \sim ThIPD(\lambda, \alpha)$, then the Pearson's β_2 coefficient is as follows

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\left[\mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1\right]}{\left[\mu'_2 - \mu'^2_1\right]^2}$$

$$= \frac{\alpha(\lambda - 3)\left(3\alpha^2(\lambda - 3)^3 - 3\alpha(\lambda - 3)(9 + (\lambda - 8)\lambda) + \lambda(31 + (\lambda - 9)\lambda) - 27\right) + \lambda^2\left(3 + \frac{1}{\lambda}\right)}{(1 - \alpha)\left(\alpha(\lambda - 3)^2 + \lambda\right)^2} \quad (9)$$

The plots of coefficient of Kurtosis of the proposed distribution for different choice of parameters are shown in the Fig. 5.

From the Fig. 5 it is observed that $\beta_2 > 3$ as α increases.

Remark 2 As $\alpha \to 0$ *and* $\lambda \to \infty$, the coefficient of kurtosis $\beta_2 = 3$ i.e. the proposed distribution tends to normal.

3.2 Probability Generating Function

Theorem 3 *If* $Z \sim ThIPD(\lambda, \alpha)$, *then its Probability Generating Function(p.g.f)*, $P_z(S)$ *is as follows*

$$P_z(S) = \alpha s^3 + (1 - \alpha)e^{\lambda(s-1)} \quad (10)$$

Proof If $Z \sim ThIPD(\lambda, \alpha)$, then the probability generating function $P_z(S)$ is

$$P_z(S) = E\left(S^z\right)$$

$$= \sum_{z=0}^{\infty} P(z, \lambda, \alpha)s^z$$

$$= \alpha s^3 + (1 - \alpha)\sum_{z=0}^{\infty} s^z \frac{e^{-\lambda}\lambda^z}{z!}$$

$$= \alpha s^3 + (1 - \alpha)e^{\lambda(s-1)}$$

Hence proved.

Remark 3 Putting $S = e^t$ in Eq. (10), the Moment Generating Function (m.g.f), $M_z(t)$ of $Z \sim ThIPD(\lambda, \alpha)$ is as follows

$$M_z(t) = \alpha e^{3t} + (1 - \alpha)e^{\lambda(e^t - 1)} \tag{11}$$

Remark 4 Putting $S = e^{it}$ in Eq. (10), the Characteristic Function, $\varphi_z(t)$ of $Z \sim ThIPD(\lambda, \alpha)$ is as follows

$$\varphi_z(t) = \alpha e^{3it} + (1 - \alpha)e^{\lambda(e^{it} - 1)} \tag{12}$$

3.3 Cumulative Distribution Function (CDF)

Theorem 4 If $Z \sim ThIPD(\lambda, \alpha)$, then its CDF of Z is as follows

$$F(z) = P(Z \le z) = \alpha + (1 - \alpha)\frac{\Gamma(z + 1, \lambda)}{\Gamma(z + 1)} \tag{13}$$

Proof If $Z \sim ThIPD(\lambda, \alpha)$, then its CDF is as follows

$$F(z) = P(Z \le z)$$

$$= \sum_{t=0}^{z} P(Z = t)$$

$$= \alpha + \sum_{t=0}^{z} (1 - \alpha)\frac{e^{-\lambda}\lambda^z}{z!}$$

$$= \alpha + (1 - \alpha)\frac{\Gamma(z + 1, \lambda)}{\Gamma(z + 1)} \left[\sum_{t=0}^{z} \frac{e^{-\lambda}\lambda^z}{z!} = \frac{\Gamma(z + 1, \lambda)}{\Gamma(z + 1)} \text{ for detail see [36]} \right]$$

where $\Gamma(z + 1, \lambda)$ and $\Gamma(z + 1)$ is an upper incomplete gamma function.

Hence proved.

The plots of CDF of $ThIPD(\lambda, \alpha)$ with different choice of parameters λ and α are provided in Fig. 6.

4 Reliability Characteristics

4.1 Survival Function (*SF*)

Theorem 5 If $Z \sim ThIPD(\lambda, \alpha)$, then its Survival Function (SF) of Z is as follows

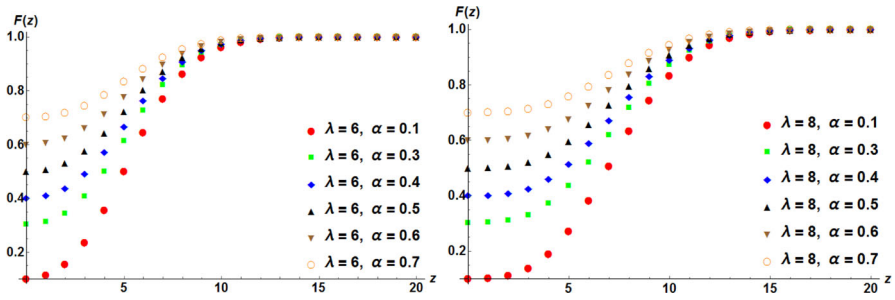$$S(z) = P(Z > z) = \alpha + (1 - \alpha)\frac{\gamma(z + 1, \lambda)}{\Gamma(z + 1)} \tag{14}$$

Fig. 6 Shapes of CDF with different choices of parameter λ *and* α

Proof If $Z \sim ThIPD(\lambda, \alpha)$, then its Survival Function *(SF)* is as follows

$$S(z) = P(Z > z)$$

$$= \sum_{t=z+1}^{\infty} P(Z = t)$$

$$= \alpha + \sum_{t=z+1}^{\infty} (1 - \alpha) \frac{e^{-\lambda} \lambda^z}{z!}$$

$$= \alpha + (1 - \alpha) \frac{\gamma(z+1, \lambda)}{\Gamma(z+1)} \left[\sum_{t=z+1}^{\infty} \frac{e^{-\lambda} \lambda^z}{z!} = \frac{\gamma(z+1, \lambda)}{\Gamma(z+1)} \text{ for detail see [36]} \right]$$

where $\gamma(z+1, \lambda)$ a lower incomplete gamma is function and $\Gamma(z+1)$ is an upper incomplete gamma function.

Hence proved.

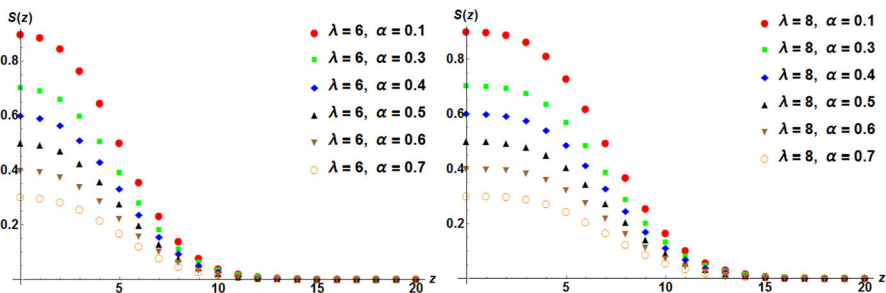The plots of Survival Function *(SF)* of $ThIPD(\lambda, \alpha)$ with different choice of parameters λ *and* α are provided in Fig. 7.

Fig. 7 Shapes of SF with different choices of parameter λ *and* α
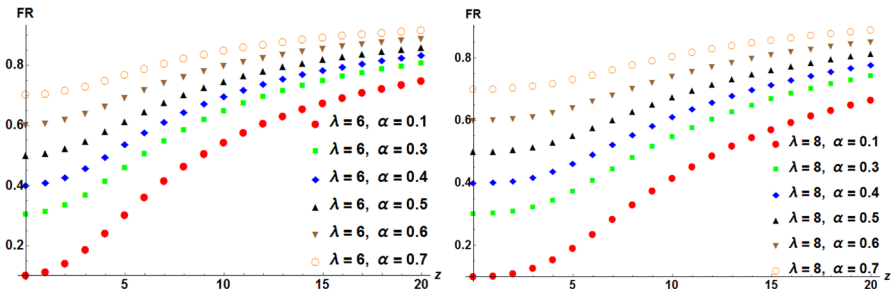
Fig. 8 Shapes of FR with different choices of parameter λ and α

4.2 Failure Rate *(FR)*

Let $z_1, z_2, z_3, ..., z_n$ be a random sample from three inflated Poisson distribution as given by Eq. (1)

Define Y be the number of $Z_i's$ taking the value three. Then Eq. (1) can be inscribed as follows

$$P(Z = z_i) = \left[\alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{3!}\right]^Y \left[(1 - \alpha)\frac{e^{-\lambda}\lambda^{z_i}}{z_i!}\right]^{1-Y}$$

and using $S(z)$ from Eq. (14)

The failure rate of $ThIPD\,(\lambda,\ \alpha)$ is given by

$$R(z) = \frac{P(z)}{S(z)} = \frac{\Gamma(z+1)\left[\left(\alpha + (1-\alpha)\frac{e^{-\lambda}\lambda^3}{3!}\right)^Y\left((1-\alpha)\frac{e^{-\lambda}\lambda^{z_i}}{z_i!}\right)^{1-Y}\right]}{\Gamma(z+1)\alpha + (1-\alpha)\gamma(z+1,\lambda)} \tag{15}$$

The plots of Failure Rate *(FR)* of $ThIPD\,(\lambda,\ \alpha)$ with different choice of parameters λ and α are provided in Fig. 8.

5 Parameter Estimation

5.1 Method of Moment Estimation (MM)

The parameters λ and α of (1) can be obtained using method of moments as follows:
Considering the first two moments from Eqs. (3) and (4)

$$\hat{\alpha} = \frac{\mu_1' - \lambda}{(3 - \lambda)} \tag{16}$$

$$\mu_2' - 3\mu_1' = (1 - \alpha)\left(\lambda^2 - 2\lambda\right) \tag{17}$$

Putting the value of α from Eq. (16), the Eq. (17) reduces to

$$\mu_2' - 3\mu_1' = \left(1 - \frac{\mu_1' - \lambda}{(3 - \lambda)}\right)\left(\lambda^2 - 2\lambda\right)$$

$$\frac{\mu_2' - 3\mu_1'}{3 - \mu_1'} = \frac{\lambda^2 - 2\lambda}{3 - \lambda} = M \text{ (say)} \tag{18}$$

Then

$$M = \frac{\lambda^2 - 2\lambda}{3 - \lambda}$$

$$\lambda^2 - \lambda(2 - M) - 3M = 0 \tag{19}$$

Solving the quadratic Eq. (19), we can estimate the value of λ, which has been used in Eq. (16) to estimate the value of α.

5.2 Maximum Likelihood Estimation (MLE)

The parameters λ and α of Eq. (1) can be obtained using method of maximum likelihood as follows:

Let $z_1, z_2, z_3, ..., z_n$ be a random sample from three inflated Poisson distribution as given by Eq. (1) and let for $i = 1, 2, 3, ..., n$

$$a_i = \begin{cases} 1, & if \ z_i = 3 \\ 0, & otherwise \end{cases}$$

Then for $i = 1, 2, 3, ..., n$, Eq. (1) can be described as follows

$$P(Z = z_i) = \left[\alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{3!}\right]^{a_i}\left[(1 - \alpha)\frac{e^{-\lambda}\lambda^{z_i}}{z_i!}\right]^{1 - a_i}$$

Hence the likelihood function, $L = L(\lambda, \alpha; z_1, z_2, z_3, ..., z_n)$ will be

$$L = \prod_{i=1}^{n}\left[\alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{3!}\right]^{a_i}\left[(1 - \alpha)\frac{e^{-\lambda}\lambda^{z_i}}{z_i!}\right]^{1 - a_i}$$

$$= \left[\alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{6}\right]^{n_0}\prod_{i=1}^{n}\left[(1 - \alpha)\frac{e^{-\lambda}\lambda^{z_i}}{z_i!}\right]^{k_i}$$

where $k_i = 1 - a_i$, $n_0 = \sum_{i=1}^{n} a_i$. Note that n_0 represents the number of three's in the sample.

Therefore,

$$\log L = n_0 \log \left[\alpha + (1 - \alpha) \frac{e^{-\lambda}\lambda^3}{6} \right] + (n - n_0) \log (1 - \alpha)$$

$$- (n - n_0)\lambda + \sum_{i=1}^{n} k_i z_i \log \lambda - \sum_{i=1}^{n} k_i \log (z_i!)$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{n_0 \left(6 - e^{-\lambda}\lambda^3 \right)}{6\alpha + (1 - \alpha)e^{-\lambda}\lambda^3} - \frac{(n - n_0)}{(1 - \alpha)} \tag{20}$$

Similarly,

$$\frac{\partial \log L}{\partial \lambda} = \frac{n_0 \left(3e^{-\lambda}\lambda^2 - e^{-\lambda}\lambda^3 \right)(1 - \alpha)}{6\alpha + (1 - \alpha)e^{-\lambda}\lambda^3} + \frac{\sum_{i=1}^{n} k_i z_i}{\lambda} - (n - n_0) \tag{21}$$

Let,

$$p = \alpha + (1 - \alpha)\frac{e^{-\lambda}\lambda^3}{6} \tag{22}$$

Now, let $\frac{\partial \log L}{\partial \alpha} = 0$. Then from Eq. (20) and using Eq. (22)

$$1 - \alpha = \frac{6(n - n_0)}{\frac{n_0}{p}\left(6 - e^{-\lambda}\lambda^3 \right)} \tag{23}$$

And letting $\frac{\partial \log L}{\partial \lambda} = 0$ from Eq. (21) and using Eq. (22)

$$\frac{n_0}{6p}\left(3e^{-\lambda}\lambda^2 - e^{-\lambda}\lambda^3 \right)(1 - \alpha) + \frac{\sum_{i=1}^{n} k_i z_i}{\lambda} - (n - n_0) = 0 \tag{24}$$

Now, if we replace p by their sample relative frequencies, i.e. by their sample estimates, the proportion of three's in the sample, i.e. $\hat{p} = {}^{n_0}/_{n}$ and then Eqs. (23) and (24) reduces to

$$1 - \alpha = \frac{6(n - n_0)}{n\left(6 - e^{-\lambda}\lambda^3 \right)} \tag{25}$$

and

$$\frac{n}{6}\left(3e^{-\lambda}\lambda^2 - e^{-\lambda}\lambda^3 \right)(1 - \alpha) + \frac{\sum_{i=1}^{n} k_i z_i}{\lambda} - (n - n_0) = 0 \tag{26}$$

Using Eq. (25), Eq. (26) reduces to

$$\sum_{i=1}^{n} k_i z_i \left(6 - e^{-\lambda}\lambda^3\right) + (n - n_0)\left(3e^{-\lambda}\lambda^3 - 6\lambda\right) = 0$$

$$C(\lambda) = 0 \tag{27}$$

where $C(\lambda) = \sum_{i=1}^{n} k_i z_i \left(6 - e^{-\lambda}\lambda^3\right) + (n - n_0)\left(3e^{-\lambda}\lambda^3 - 6\lambda\right)$.

Hence Eq. (27) can be solved by any numerical procedure, say, Newton Rapson method, to obtain $\hat{\lambda}$ numerically, i.e. $C\left(\hat{\lambda}\right) = 0$.

Similarly using Eqs. (22) and (25), α can be estimated as

$$\hat{\alpha} = \frac{1}{n}\left[n_0 - \frac{(n - n_0)e^{-\lambda}\lambda^3}{6 - e^{-\lambda}\lambda^3}\right] \tag{28}$$

Therefore, the maximum likelihood estimates (MLE) of the parameter λ and α can be estimated by solving (27) numerically to find $\hat{\lambda}$, and Eq. (28) gives $\hat{\alpha}$ respectively.

For the reckoning of the asymptotic variance–covariance matrix of the estimates the second order differentiations of the log-likelihood function are furnished here

$$\frac{\partial^2 \log L}{\partial \alpha^2} = -\frac{n_0\left(6 - e^{-\lambda}\lambda^3\right)^2}{\left[6\alpha + (1 - \alpha)e^{-\lambda}\lambda^3\right]^2} - \frac{(n - n_0)}{(1 - \alpha)^2}$$

$$\frac{\partial^2 \log L}{\partial \lambda^2}$$
$$= n_0(1 - \alpha)\left[\frac{3\lambda e^{-\lambda}(2 - \lambda)}{6\alpha + (1 - \alpha)e^{-\lambda}\lambda^3} - \frac{\lambda^2 e^{-\lambda}(3 - \lambda)\left[6\alpha + (1 - \alpha)3\lambda^2 e^{-\lambda}\right]}{\left(6\alpha + (1 - \alpha)e^{-\lambda}\lambda^3\right)^2}\right]$$
$$- \frac{\sum_{i=1}^{n} k_i z_i}{\lambda^2}$$

$$\frac{\partial \log L}{\partial \lambda \partial \alpha} = \frac{6n_0 e^{-\lambda}\lambda^2(\lambda - 3)}{\left(6\alpha + (1 - \alpha)e^{-\lambda}\lambda^3\right)^2}$$

The asymptotic variance–covariance matrix of the maximum likelihood estimates of λ and α for $ThIPD\,(\lambda, \alpha)$, can be acquired by inverting the Fisher information matrix (I), given by

$$I = \begin{bmatrix} E\left(-\frac{\partial^2 \log L}{\partial \alpha^2}\right) & E\left(-\frac{\partial^2 \log L}{\partial \alpha \partial \lambda}\right) \\ E\left(-\frac{\partial^2 \log L}{\partial \lambda \partial \alpha}\right) & E\left(-\frac{\partial^2 \log L}{\partial \lambda^2}\right) \end{bmatrix}$$

The ingredient of the above Fisher information matrix can be acquired as.

$$E\left(-\frac{\partial^2 \log L}{\partial \alpha^2}\right) = \left(-\frac{\partial^2 \log L}{\partial \alpha^2}\right)\Big|_{\alpha=\hat{\alpha}, \lambda=\hat{\lambda}}, \text{ and so on.}$$

The asymptotic distribution of the maximum likelihood estimator $\left(\hat{\lambda}, \hat{\alpha}\right)^{/}$ is given by

$$\sqrt{n}\begin{pmatrix}\hat{\lambda}\\\hat{\alpha}\end{pmatrix}_{MLE} \xrightarrow{L} AN\left(\begin{pmatrix}\hat{\lambda}\\\hat{\alpha}\end{pmatrix}, I^{-1}\right), \text{ as } n \to \infty$$

6 Simulation Study

In this section a simulation study has been conducted to see the performance of the estimated parameters. Here, to generate random numbers Z from $ThIPD(\lambda, \alpha)$ we have applied acceptance rejection sampling [37]. By applying this method random samples are generated of size $n = 30, 50, 100$ and 200 with different combinations of true values of parameters λ and α and finally, MLEs are computed using optim function of R software. Bias and MSE of the parameters given in the Table 1 are calculated using the following formulae.

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

$$MSE\left(\hat{\theta}\right) = E\left(\hat{\theta} - \theta\right)^2$$

$\hat{\theta} = $ estimated parameter and $\theta = $ true parameter

$$\theta = (\lambda, \alpha), \hat{\theta} = \left(\hat{\lambda}, \hat{\alpha}\right)$$

Here $r(= $ number of replication$) = 1000$.

From the values of the MSE and biases of the simulation study given in Table 1, it is observed that the bias and MSE of the estimators are small and as the sample size increases the estimated bias and MSE also gradually decreases which is as expected.

Furthermore we have checked the normality of the MLEs by normal Q-Q plot for all the parameters of each run. One such Q-Q plot is presented here obtained for the case when $(\lambda, \alpha) = (5, 0.3)$ as a demonstration (Fig. 9). From the Fig. 9 it is observed that the MLEs of all the parameters follow approximately normal distribution.

7 Real-life Applications

Covid-19 virus started from Wuhan, China and has blazed the trail of a new world order [38]. This new world order necessitated that the global community drop and dissolve all culture differences and brainstorms to locate mitigating measures especially in

Table 1 Results of Simulation

$\alpha \downarrow$	$n \rightarrow$	$\lambda \rightarrow$ 3				5				7			
		30	50	100	200	30	50	100	200	30	50	100	200
0.1	Bias$(\hat{\lambda})$	0.0141	− 0.0119	0.0072	− 0.0002	0.0139	− 0.0110	0.0098	− 0.0011	0.0134	− 0.0117	0.0090	0.0004
	MSE$(\hat{\lambda})$	0.0885	0.0685	0.0398	0.0199	0.0828	0.0625	0.0585	0.0292	0.0884	0.0803	0.0779	0.0384
	Bias$(\hat{\alpha})$	0.0040	− 0.0009	− 0.0007	0.0001	0.0033	0.0013	0.0010	0.0003	0.0029	0.0020	0.0008	0.0003
	MSE$(\hat{\alpha})$	0.0043	0.0028	0.0016	0.0007	0.0032	0.0019	0.0011	0.0005	0.0031	0.0018	0.0009	0.0005
0.2	Bias$(\hat{\lambda})$	− 0.0055	− 0.0043	0.0023	− 0.0009	0.0078	− 0.0058	0.0052	− 0.0037	0.0057	0.0050	0.0041	− 0.0035
	MSE$(\hat{\lambda})$	0.0801	0.0615	0.0448	0.0233	0.0825	0.0734	0.0675	0.0331	0.0947	0.0901	0.0893	0.0452
	Bias$(\hat{\alpha})$	− 0.0018	− 0.0015	− 0.0010	0.0004	0.0047	− 0.0038	0.0031	0.0004	0.0037	− 0.0033	0.0030	0.0004
	MSE$(\hat{\alpha})$	0.0076	0.0042	0.0021	0.0011	0.0057	0.0031	0.0017	0.0008	0.0055	0.0031	0.0017	0.0008
0.3	Bias$(\hat{\lambda})$	0.0115	0.0097	0.0059	− 0.0040	0.0151	− 0.0111	0.0051	− 0.0046	0.0149	0.0129	0.0121	0.0076
	MSE$(\hat{\lambda})$	0.0865	0.0624	0.0483	0.0268	0.0921	0.0825	0.0749	0.0383	0.1325	0.1098	0.1045	0.0532
	Bias$(\hat{\alpha})$	0.0027	0.0002	0.0001	0.0001	0.0037	− 0.0027	0.0017	0.0005	0.0039	0.0034	0.0031	0.0023
	MSE$(\hat{\alpha})$	0.0083	0.0051	0.0025	0.0013	0.0072	0.0045	0.0020	0.0010	0.0070	0.0043	0.0021	0.0010

Table 1 (continued)

| $\alpha \downarrow$	$n \rightarrow$	$\lambda \rightarrow$ 3				5				7			
		30	50	100	200	30	50	100	200	30	50	100	200
0.4	Bias$(\hat{\lambda})$	− 0.0127	0.0102	− 0.0056	0.0006	− 0.0139	0.0113	0.0094	− 0.0003	− 0.0138	− 0.0118	0.0085	− 0.0009
	MSE$(\hat{\lambda})$	0.0965	0.0812	0.0587	0.0297	0.1021	0.0968	0.0897	0.0429	0.1412	0.1327	0.1219	0.0566
	Bias$(\hat{\alpha})$	0.0053	− 0.0029	− 0.0023	− 0.0003	0.0008	0.0003	0.0004	0.0000	0.0013	− 0.0006	− 0.0003	0.0000
	MSE$(\hat{\alpha})$	0.0094	0.0057	0.0026	0.0013	0.0082	0.0049	0.0025	0.0012	0.0081	0.0051	0.0023	0.0012
0.5	Bias$(\hat{\lambda})$	− 0.0286	− 0.0139	− 0.0125	0.0000	0.0056	0.0041	− 0.0038	− 0.0001	0.0065	0.0057	0.0047	− 0.0011
	MSE$(\hat{\lambda})$	0.1002	0.0956	0.0711	0.0367	0.1512	0.1290	0.1101	0.0535	0.1766	0.1505	0.1465	0.0694
	Bias$(\hat{\alpha})$	− 0.0059	0.0029	− 0.0024	0.0005	0.0019	− 0.0007	− 0.0005	0.0004	0.0017	− 0.0010	− 0.0005	0.0004
	MSE$(\hat{\alpha})$	0.0092	0.0057	0.0028	0.0014	0.0079	0.0048	0.0025	0.0012	0.0077	0.0050	0.0024	0.0012
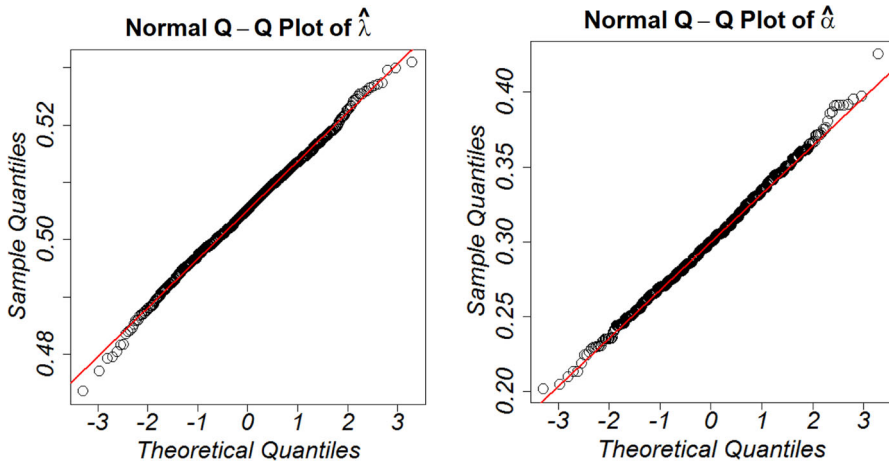
Fig. 9 Q-Q plots of different parameter estimates

respect of providing sustenance to the economy (for details see [39–41]). Something that came as a bolt from the blue, India was unprepared to defend the onslaughts of Covid-19. In India, the first case of Covid-19 was reported on 30th January, 2020 [42]. The disease accelerated to such a level that is prompted the Govt. of India to enforce and clamp an emergency like Lockdown (lockdown denotes a clamp down on almost all human transactions and activities in an emergency), the fallouts of which have been discernible in an emphatic manner [43]. The Lockdown in India starts from 25th March 2020 to 7th June 2020 [44] in a phased manner. The first phase of lockdown in India starts from 25th March 2020 for a period of twenty one days [45]. After the first lockdown, the next three phases are announced with conditional relaxations and restrictions- the second phase with effect from 15th April 2020 to 3rd May 2020, the third phase from 4th May 2020 to 17th May 2020 and the fourth phase from 18th May 2020 to 31st May 2020 [45]. The fourth phase of lockdown extended to 7th June 2020.The hardest hit because of the clamping of lockdown/s have been those living on the edge and on the margins like the daily wage earners, private job seekers who lost their job, the farmers who could not locate markets to sell their agricultural produce, the migrant workers who were left stranded like anything, the large chunks of underprivileged students and those who opted for reverse migration [43, 46]. The most discernible corollary has therefore, been the conditions of poverty and starvation which in turn have sapped the vitality and jolted the psychological and cognitive get-up of the hugely dense Indian populace [43]. Billed as a biomedical disease with negative cognitive responses, it has been unfortunate that Asian countries have been facing the brunt in terms of the exponential growth of the transmission of Sars-Cov-2 in densely populated areas of internal migrants. There have alarming instances of socially irresponsible behavior and panic attacks among internal migrant workers who are in desperate need of psycho-social support [47].

After completion of the fourth phase of lockdown, the Government of India started unlocking (unlocking denotes relaxations on the imposed lockdown in the event of any

emergency) the nation in a phased manner with restriction to containment zones from 8th June, 2020 [48]. The first phase of unlocking i.e. unlocking 1.0 starts from 8th June 2020 to 30th June 2020 [48]. Covid-19 has a diverse array of effects. The worst being that of committing suicides primarily triggered by uncertainty regarding living from hand to mouth in all aspects of life. Given the circumstances, the situation has come to such a pass that suicides, more often than not have occasionally hogged the limelight. The first Indian suicide related to Covid-19 took place on February 12, 2020 [49], followed by two more such suicides [50]. In addition, the first Covid-19 related Student Suicide case was reported on June 2, 2020 [51]. The reasons being financial distress, fear of infection, freezing of employment opportunities, lack of freedom of movement, withdrawal etc. the major cause of occurrences of suicide [47, 49, 50].

The data set of 298 Covid-19 pandemic related suicide cases during Lockdown to Unlocking 1.0 in India are collected from the web portal https://www.kaggle.com/.. The age and sex distribution of the individuals who committed Suicides during Lockdown and Unlocking 1.0 are presented in (Fig. 10) which manifest that, throughout both lockdown and unlocking1.0, the highest percentage of suicides has been committed by individuals of the age group 21–40 and male individuals. The causes of suicide during Lockdown and Unlocking 1.0 are presented in (Fig. 11), where it can be observed that due to financial distress and fear of infection maximum suicide occurred during lockdown and unlocking 1.0. The occupation of the individuals who committed suicide during Lockdown and Unlocking 1.0 are presented in (Fig. 4), where it can be
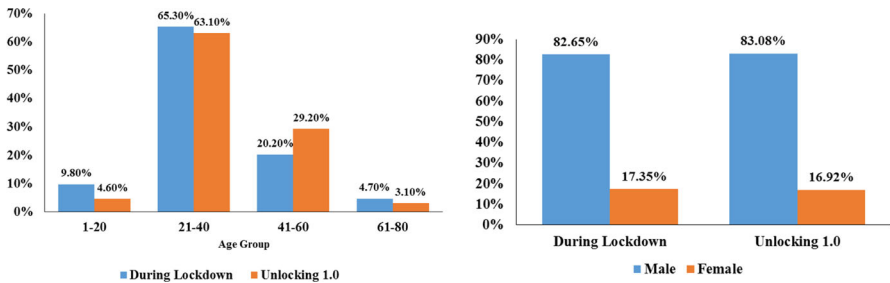

Fig. 10 Age and Sex of the individuals who Suicides during Lockdown and Unlocking 1.0
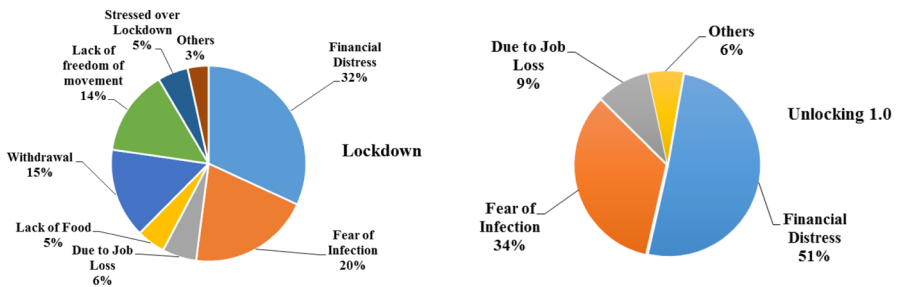

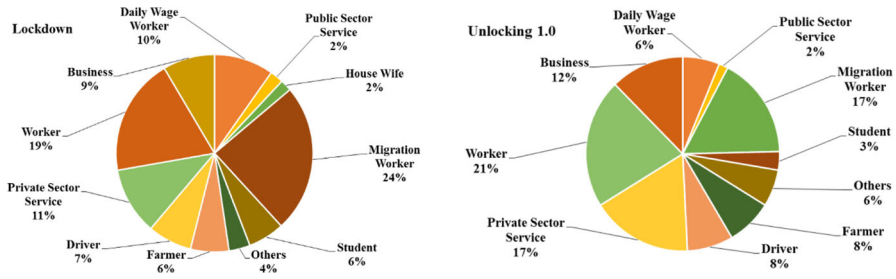**Fig. 11** Causes of suicide during Lockdown and Unlocking 1.0

Fig. 12 Occupation of the individuals who suicides during Lockdown and Unlocking 1.0

observed that maximum individuals who committed suicides during lockdown and unlocking1.0 are migration workers, worker and private sector service (Fig. 12).

In order to study the pattern of the suicide cases, Poisson distribution is used. Since during lockdown to unlocking 1.0 the proportion of three (3) deaths per day is inflated than the others, so we used three-inflated Poisson distribution $ThIPD(\lambda, \alpha)$ for fitting the real life data set of Covid-19 related suicides along with zero inflated Poisson distribution $ZIPD(\lambda, \gamma)$ [11], and zero–one inflated Poisson distribution $ZOIP(\lambda, \gamma, \psi)$[22] and standard Poisson distribution $PD(\lambda)$. The values of the MLEs of the parameters for different distributions are estimated using optim function of R language. The log-likelihood, Akaike information criterion (AIC), Bayesian information criterion (BIC) and the Kolmogorov–Smirnov test (KS test) with *p*-values are summarized in Table 2 for the number of suicides cases during the 98 days of lockdown and Unlocking 1.0 in India during Covid-19 pandemic.

From the Table 2 it is seen that the value of AIC and BIC of ThIPD is smaller than PD, ZIPD, ZOIPD and highest P-value of the KS statistics of ThIPD and also the expected frequencies of ThIPD are closer to the observed frequencies.

In Fig. 13 the observed histogram and estimated pmf's of PD, ZIPD, ZIOPD and ThIPD are plotted which also validate our findings and in Fig. 14 the observed Ogive and estimated cdf's of PD, ZIPD, ZOIPD and ThIPD are plotted for visual comparisons.

The proposed three-inflated Poisson distribution (ThIPD) provides better fit to the data set under consideration of all criteria.

8 Likelihood Ratio Test

Since $Poi(\lambda)$ and $ThIPD(\lambda, \alpha)$ are nested models, the likelihood ratio (LR) test is used to discriminate between them. The LR test is carried out to test the hypothesis: $H_0 : \alpha = 0$, that is the sample is drawn from $Poi(\lambda)$; against the alternative $H_0 : \alpha \neq 0$, that is the sample is drawn from $ThIPD(\lambda, \alpha)$. The value of LR test statistic for the above dataset is given below in Table 3.

The value of the LR test statistic for the dataset is respectively 28.542 which exceeds the critical value at 5% level of significance for one (1) degrees of freedom, i.e., 3.841.

Table 2 Distribution of number of suicides per day during Lockdown to Unlocking 1.0 in 98 observed Days of India

Number of suicides per day during Lockdown to Unlocking 1.0	Observed Number of Days	$P(\lambda)$	$ZIPD(\lambda, \gamma)$	$ZOIPD(\lambda, \gamma, \psi)$	$ThIPD(\lambda, \alpha)$
0	8	4.684	8.003	7.958	3.143
1	6	14.243	12.487	12.497	9.615
2	8	21.656	19.807	19.820	14.710
3	46	21.950	20.945	20.956	45.999
4	23	16.686	16.611	16.618	11.476
5	2	10.148	10.540	10.543	7.023
6	2	5.143	5.573	5.574	3.581
7	1	2.234	2.526	2.526	1.565
8	1	0.849	1.001	1.001	0.599
9	1	0.287	0.353	0.353	0.204
Estimated parameters		$\hat{\lambda} = 3.041$	$\hat{\lambda} = 3.172$	$\hat{\lambda} = 3.172$	$\hat{\lambda} = 3.059$
			$\hat{\gamma} = 0.042$	$\hat{\gamma} = 0.041$	$\hat{\alpha} = 0.316$
				$\hat{\psi} = 7.044 \times 10^{-12}$	
	$Log\ L$	-182.313	-181.061	-181.061	-168.042
	$KS\ test$	0.190	0.187	0.184	**0.061**
	p-value	0.029	0.032	0.036	**0.694**
	AIC	366.626	366.122	368.122	**340.084**
	BIC	369.211	371.282	375.877	**345.254**

Fig. 13 Plots of observed histogram and estimated pmf of PD, ZIPD, ZOIPD and ThIPD for the number of suicides cases during the 98 days of lockdown and Unlocking 1.0 in India during Covid-19 Pandemic
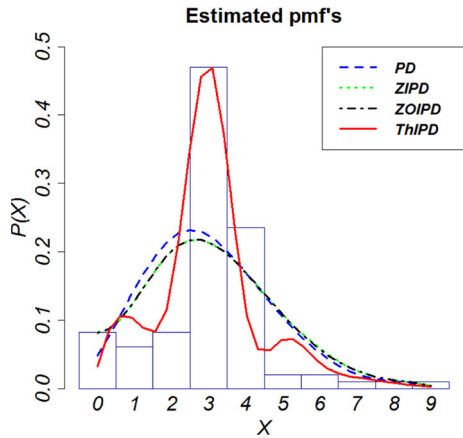
Fig. 14 Plots of observed Ogive and estimated cdf's of PD, ZIPD, ZOIPD and ThIPD for the number of suicides cases during the 98 days of lockdown and Unlocking 1.0 in India during Covid-19 Pandemic
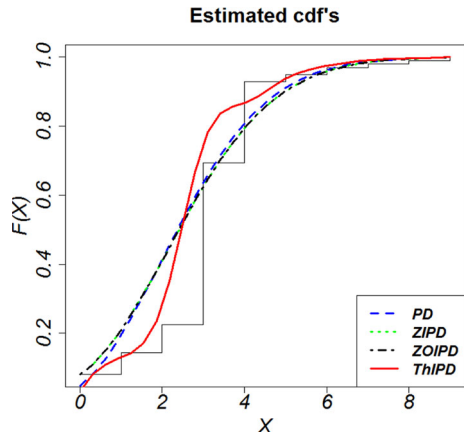

Table 3 The value of the LR test statistic for the Covid-19 related suicide dataset

	Dataset	Degrees of freedom	Critical value
LR test statistic	28.542	1	3.841

Thus the evidence is in support of the alternative hypothesis that the sample data comes from $ThIPD(\lambda, \alpha)$ and not from $Poi(\lambda)$.

9 Conclusion

A three-inflated Poisson distribution (ThIPD) is proposed and we studied its distributional properties and reliability characteristics. A simulation study has been conducted to see the behavior of the MLEs. The appropriateness of fitting the distribution is carried out based on the goodness of fit test and some information criteria. The usefulness of the proposed distribution is exemplified by the data of number of suicides occurred during lockdown to unlocking 1.0 in India. The real life data set of Covid-19 related suicides considered here has shown that the proposed three-inflated Poisson distribution (ThIPD) provides better fit in comparison to the other known distributions viz. ZIPD, ZOIPD and general PD under considerations in terms of model selection criteria, namely AIC and BIC and goodness of fit test, namely KS-test. The plots presented above also validate our findings. Moreover from the LR test it is observed that the sample comes from ThIPD, not from PD. Thus our proposed distribution provide better fitting in comparison to the other competitor distributions.

Author's contributions All authors equally contributed towards preparing the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Statement We hereby declare that this manuscript is the result of our independent creation under the reviewer's comments. Except for the quoted contents, this manuscript does not contain any research achievements that have been published or written by other individuals or groups. We are the only authors of this manuscript. The legal responsibility of this statement shall be borne by us.

References

1. Olson DL, Shi Y, Shi Y (2007) Introduction to business data mining. McGraw-Hill/Irwin, New York, pp 2250–2254
2. Shi Y, Tian YJ, Kou G, Peng Y, Li JP (2011) Optimization based data mining: theory and applications. Springer, Berlin
3. Tien JM (2017) Internet of things, real-time decision making, and artificial intelligence. Annal Data Sci 4(2):149–178
4. Nani L (2019) Modeling Zero-inflated and overdispersed count data: Application to IN-Hospital mortality data. PhD Thesis, The University of Tennessee at Chattanooga, Chattanooga TN.
5. Lam KF, Xue H, Cheung YB (2006) Semiparametric analysis of zero-inflated count data. Biometrics 62:996–1003
6. He H, Tang W, Wang W, Crits-Christoph P (2014) Structural zeroes and zero-inflated models. Shanghai Arch Psy 26(4):236–242
7. Yang S, Harlow LL, Puggioni G, Redding CA (2017) A comparison of different methods of zero-inflated data analysis and an application in health surveys. J Mod Appl Stat Methods 16(1):518–543
8. Bodhisuwan W, Samutwachirawong S, Payakkapong P (2018) The zero-inflated negative binomial-Erlang distribution: An application to highly pathogenic avian influenza H5N1 in Thailand. Songklanakarin J Sci Technol 40(6):1428–1436
9. Neyman J (1939) On a new class of contagious distributions applicable in entomology and bacteriology. Ann Math Stat 10:35–57
10. Feller W (1945) On a general class of contagious distributions. Ann Math Stat 12:389–400
11. Mullahy J (1986) Specification and testing of some modified count data models. J Econometrics 33(3):341–365
12. Pandey KN (1964) Generalized inflated Poisson distribution. J Sci Res 15(2):157–162
13. Johnson NL, Kotz S, Kemp AW (1992) Univariate discrete distributions, 2nd edn. John Wiley and Sons, New York
14. Gupta PL, Gupta RL, Tripathi RC (1995) Inflated modified power series distributions with applications. Commun Stat Theory Methods 24(9):2355–2374
15. Murat M, Szynal D (1998) Non–Zero-Inflated modified power series distributions. Commun Stat-Theory Methods 27(12):3047–3064
16. Nanjundan G, Naika TR (2012) Asymptotic Comparison of Method of Moments estimators and maximum likelihood estimators of parameters in Zero-Inflated Poisson model. Appl Math 3:610–616
17. Beckett S, Jee J et al (2014) Zero-inflated Poisson (ZIP) distribution: parameter estimation and applications to model data from natural calamities. Involve J Math 7(6):751–767
18. Nanjundan G, Pasha S (2015) A note on the characterization of Zero-inflated Poisson model. Open J Stat 5:140–142
19. Nanjundan G, Pasha S (2015) A characterization of Zero-inflated binomial model. Int J Math Comp Res 3(10):1188–1190

20. Suresh R, Nanjundan G et al (2015) On a characterization of Zero-inflated negative binomial distribution. Open J Stat 5:511–513

21. Nagesh S, Nanjundan G et al (2015) A Characterization of Zero-inflated geometric model. Int J Math Trends Technol 23(1):71–73

22. Alshkaki RSA (2016) On the Zero-One Inflated Poisson distribution. Int J Stat Distributions Appl 2(4):42–48

23. Singh BP, Maheshwari S, Gupta PK (2015) A probability model for sex composition of children in the presence of son preference. Demography India 44:50–57

24. Mwalili SM, Lesaffre E, Declerck D (2008) The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. Stat Methods Med Res 17(2):123–139

25. Lambert D (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34(1):1–14

26. Bohning D (1998) Zero-inflated Poisson models and C.A.MAN: A tutorial collection of evidence. Biom J 40:833–843

27. Bohning D, Dietz E, Schlattmann P, Mendonca L, Kirchner U (1999) The zero inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. J R Stat Soc Ser A 162:195–209

28. Ridout M, Hinde J, Demetrio CGB (2001) A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. Biometrics 57:219–223

29. Lee AH, Wang K, Yau KKW (2001) Analysis of zero-inflated Poisson data incorporating extent of exposure. Biom J 43:963–975

30. Yau KKW, Lee AH (2001) Zero-inflated Poisson regression with random effects to evaluate an occupational injury prevention programme. Stat Med 20:2907–2920

31. Hall DB (2000) Zero-inflated Poisson and binomial regression with random effects: A case study. Biometrics 56:1030–1039

32. Famoye F, Singh KP (2006) Zero-inflated generalized Poisson regression model with an application to domestic violence data. J Data Sci 4(1):117–130

33. Akaike H (1974) A new look at the statistical model identification. IEEE-TAC 19(6):716–723

34. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464

35. Rohatgi VK, Saleh AKMdE (2003) An Introduction to Probability and Statistics. John Willy and Sons, New York

36. Haight FA (1967) Handbook of the Poisson distribution. John Wiley & Sons Inc, New York

37. Ross SM (2012) Simulation. Elsevier Academic Press, USA

38. Barman MP, Rahman T, Bora K, Borgohain C (2020) COVID-19 pandemic and its recovery time of patients in India: A pilot study. Diabetes Metab Syndr 14:1205–1211

39. Guan C, Liu W, Cheng JYC (2021) Using social media to predict the stock market crash and rebound amid the pandemic: the digital 'Haves' and 'Have-mores.' Annals Data Sci. https://doi.org/10.1007/s40745-021-00353-w

40. Li J, Guo K, Herrera VE, Lee H, Liu J, Zhong Z, Gomes L, Filip FG, Fang SC, Özdemir MS, Liu XH, Lu G, Sh Y (2020) Culture vs policy: more global collaboration to effectively combat COVID-19. The Innovation. https://doi.org/10.1016/j.xinn.2020.100023

41. Liu Y, Gu Z, Xia S, Shi B, Zhou X, Shi Y, Liu J (2020) What are the underlying transmission patterns of COVID-19 outbreak? An Age-Specific Social Contact Characterization. EClincialMedicine 22:100354

42. Kumar S (2020) Monitoring Novel Corona Virus (COVID-19) infections in India by cluster analysis. Annals of Data Science 7(3):417–425

43. Golechha M (2020) COVID-19, India, lockdown and psychosocial challenges: What next? Int J Soc Psychiatry 66(8):830–832

44. Rahman T, Bora K, Barman MP, Goswami K, Borgoain C (2020) Unlocking India during Covid-19 pandemic: a data driven investigation. Int J Stat Reliab Eng 7(3):445–450

45. Rai B, Shukla A, Dwivedi LK (2020) Dynamics of COVID-19 in India: A review of different phases of lockdown. Population Med 21(2):1–2

46. Rehman U, Shahnawaz MG, Khan NH, Kharshiing KD, Khursheed M, Gupta K, Kashyap D, Uniyal R (2020) Depression, Anxiety and Stress among Indians in times of Covid-19 Lockdown. Community Ment Health J 57(1):42–48

47. Choudhari R (2020) COVID 19 pandemic: Mental health challenges of internal migrant workers of India. Asian J Psy 54:102254

48. Saha J, Chouhan P (2021) Lockdown and unlock for the COVID-19 pandemic and associated residential mobility in India. Int J Infect Dis 104:382–389

49. Goyal K, Chauhan P, Chhikara K, Gupta P, Singh MP (2020) Fear of COVID 2019: first suicidal case in India! Asian J Psy 49:e101989
50. Sahoo S, Bharadwaj S, Parveen S, Singh AP, Tandup C, Mehra A, Grover S (2020) Self-harm and COVID-19 Pandemic: an emerging concern–a report of 2 cases from India. Asian J Psychiatry 51:e1022104
51. The Hindu (2020) Kerala class X girl ends life allegedly over lack of access to online classes. https://www.thehindu.com/news/national/kerala/kerala-class-x-girl-ends-life-allegedly-over-lack-of-access-to-online-classes/article31728470.ece. Accessed 11 June 2020