On Discrete Poisson–Mirra Distribution: Regression, INAR(1) Process and Applications

Radhakumari Maya 1, Muhammed Rasheed Irshad 2, Christophe Chesneau 3,*, Soman Latha Nitin 4 and Damodaran Santhamani Shibu 4

1 Department of Statistics, Government College for Women, Thiruvananthapuram 695 014, Kerala, India; publicationsofmaya@gmail.com
2 Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India; irshadmr@cusat.ac.in
3 Department of Mathematics, Université de Caen Basse-Normandie, LMNO, UFR de Sciences, F-14032 Caen, France
4 Department of Statistics, University College, Thiruvananthapuram 695 034, Kerala, India; nitinstat24@gmail.com (S.L.N.); statshibu@gmail.com (D.S.S.)

* Correspondence: christophe.chesneau@unicaen.fr

Abstract: Several pieces of research have spotlighted the importance of count data modelling and its applications in real-world phenomena. In light of this, a novel two-parameter compound-Poisson distribution is developed in this paper. Its mathematical functionalities are investigated. The two unknown parameters are estimated using both maximum likelihood and Bayesian approaches. We also offer a parametric regression model for the count datasets based on the proposed distribution. Furthermore, the first-order integer-valued autoregressive process, or INAR(1) process, is also used to demonstrate the utility of the suggested distribution in time series analysis. The unknown parameters of the proposed INAR(1) model are estimated using the conditional maximum likelihood, conditional least squares, and Yule–Walker techniques. Simulation studies for the suggested distribution and the INAR(1) model based on this innovative distribution are also undertaken as an assessment of the long-term performance of the estimators. Finally, we utilized three real datasets to depict the new model’s real-world applicability.

Keywords: compounding; over-dispersion; Bayesian estimation; count time series; COVID-19 data; earthquake data

MSC: 60E05; 62E15; 62M10

1. Introduction

Many studies have underlined the importance of modelling count data as well as the time series of counts, which has sparked considerable interest in a variety of sectors, including medical science, earth science, physics, finance, and insurance. For modelling count datasets, the Poisson distribution is the most frequently used, but it has the drawback of not being able to model over-dispersed datasets, while the negative binomial distribution is used for modelling over-dispersed data. When investigating counting data, the existence of at least one overdispersion warrants extra consideration when selecting a count model (for details, see [1,2]). As a consequence, the conventional Poisson regression model is rarely used in these situations.

Furthermore, one of the most widely utilized methodologies for analyzing time-dependent data are time series analysis. The corresponding time series is known as a time series of counts or, equivalently, an integer-valued time series when the data are generated by a random counting procedure. The demand for modelling count time series is seen in diverse real-life situations—for example, the monthly number of earthquakes in...
a specific place, the monthly number of automobile sales, the monthly number of deaths due to a specific disease, the monthly number of traffic accidents, and the number of living cells (see [3]). Refs. [4–6] proposed the class of first-order non-negative integer-valued autoregressive, shortly INAR(1), processes with Poisson innovation based on the binomial thinning operator to model time series of counts. INAR models are frequently built using the binomial thinning operator.

Moreover, the INAR(1) model can be altered by changing the innovation distributions according to various real-life situations. According to [7], the INAR(1) model with respect to negative binomial innovations (NB-INAR(1)) is effective for generating overdispersion. Since the common occurrence in overdispersion is that the incidence of zero counts is higher than expected from the Poisson distribution, Ref. [8] proposed an INAR(1) process with geometric distribution as the innovation model. Ref. [9] proposed the PL-INAR(1) model, which combines the INAR(1) model with Poisson–Lindley innovations. Then, Ref. [10] introduced the INAR(1) process with Poisson–Bilal innovations. These are a few of the most significant studies on the overdispersed INAR(1) process. Even while these methods give fine solutions for over-dispersed time series count datasets, they still have significant drawbacks that can pose computing challenges in real-world applications. Thus, discovering more INAR(1) models will provide more possibilities for superiorly fitting the real datasets by choosing those models according to the situations.

In this paper, we construct a novel discrete two-parameter distribution by mixing the Poisson and Mirra distributions, which can be used to model overdispersed count data sets. Hereafter, we call this new distribution the Poisson–Mirra distribution (PMiD). The two-parameter Mirra distribution by [11] is considered to be the generalization of the Xgamma distribution with one parameter, which is proposed by [12]. In addition, Ref. [13] considers the Poisson–Xgamma distribution to be an alternative or rival to the well-known one-parameter Poisson–Lindley distribution proposed by [14]. Interestingly, we discovered that the proposed PMiD is the generalization of the Poisson–Xgamma distribution. As a result, more research into the PMiD is unavoidable, both theoretically and in terms of applied aspects, which is the primary motivation for this work. The suggested distribution has the advantages of having a simple probability mass function (pmf) and cumulative distribution function (cdf), as well as explicit probability and moment generating functions, and can monitor over-dispersed count datasets, which are common in real-world data modelling. By using the PMiD as an innovation process, we also emphasize the significance of the PMiD in the INAR(1) process.

The following is how the rest of the article is sorted. The detailed description of the Mirra distribution is covered in Section 2. The definition of the new distribution, its factorial moments, moments about the origin, skewness, kurtosis, mode, generating functions, entropy measures, and other details are presented in Section 3. In Section 4, the maximum likelihood (ML) estimation and the Bayesian estimation technique are defined to estimate the unknown parameters of the new distribution, and the performance of the PMiD parameters for the ML estimation is also studied using simulation technique in the same section. A regression model with respect to the new distribution is discussed in Section 5. Section 6 develops the INAR(1) model construction with PMiD innovations; the PMiD-INAR(1) process is created. The methods for the estimation of parameters, which include the conditional maximum likelihood, conditional least squares, and the Yule–Walker estimation procedures for the PMiD-INAR(1) process are discussed in Section 7, and the simulation studies based on these three estimation procedures are also conducted in the same section. The applications and the empirical studies based on the new model concerning three real datasets are conducted in Section 8. A detailed discussion of the article is presented in Section 9. Then, Section 10 finishes with the decisive concluding words.

2. The Mirra Distribution

An inaugural approach to the Mirra distribution (MiD) was implemented in the literature by [11]. Let \( T \) be a continuous random variable which follows the MiD with
parameters $\alpha$ and $\theta$. Then, the probability density function ($pdf$), and the survival function ($sf$) of the random variable $T$ are respectively given by

$$f(t) = \frac{\theta^3}{\theta^2 + \alpha} \left( 1 + \frac{\alpha t^2}{2} \right) e^{-\theta t}$$

and

$$S(t) = \frac{\theta^2}{\theta^2 + \alpha} \left( 1 + \frac{\alpha}{\theta^2} + \frac{\alpha t}{\theta} + \frac{\alpha t^2}{2} \right) e^{-\theta t},$$

where $t > 0$ and the parameters $\alpha, \theta > 0$. This distribution is denoted by $Mi(\alpha, \theta)$. The plots in Figure 1 portray the graphical representation of the pdf.

![Figure 1. Plots of the pdf of the MiD.](image)

Now, the hazard function ($hf$) of the MiD is given as

$$h(t) = \frac{\theta \left( 1 + \frac{\alpha t^2}{2} \right)}{1 + \frac{\alpha}{\theta^2} + \frac{\alpha t}{\theta} + \frac{\alpha t^2}{2}}.$$  

As indicated by [11], the $hf$ of the MiD is shaped as a bathtub, decreasing for $t < \sqrt{\frac{2}{\alpha}}$, and increasing for $t > \sqrt{\frac{2}{\alpha}}$.

As a special case, for $\alpha = \theta = \lambda$, the Xgamma distribution (XGD) is obtained with the $pdf$ given by

$$f(t) = \frac{\lambda^2}{\lambda + 1} \left( 1 + \frac{\lambda t^2}{2} \right) e^{-\lambda t},$$

where $t > 0$ and the parameter $\lambda > 0$. For more information on the XGD, see [12].

3. The Poisson–Mirra Distribution

3.1. Presentation

Through the following proposition, a novel mixed-Poisson distribution is introduced by compounding the Poisson and Mirra distributions.

Proposition 1. Assume that $X$ follows the new compound Poisson–Mirra distribution (PMiD), which has the stochastic representation as follows:

$$X|\lambda \sim \text{Poisson}(\lambda)$$

$$\lambda|\alpha, \theta \sim Mi(\alpha, \theta),$$
where \( \lambda, \alpha \) and \( \theta > 0 \). Then, the pmf of \( X \) is given by

\[
p(x; \alpha, \theta) = \Pr(X = x) = \frac{\theta^3}{(\theta^2 + \alpha)(1 + \theta)^{x+1}} \left\{ 1 + \frac{\alpha(x + 1)(x + 2)}{2(1 + \theta)^2} \right\}, \quad x = 0, 1, 2, \ldots \tag{1}
\]

This distribution is denoted as PMiD\((\alpha, \theta)\), and one can note \( X \sim \text{PMiD}(\alpha, \theta) \) to inform that \( X \) follows the PMiD with parameters \( \alpha \) and \( \theta \).

**Proof.** The pmf of \( X \) can be derived using the general compounding formula as follows:

\[
p(x; \alpha, \theta) = \int_0^\infty \Pr(X = x|\lambda) f(\lambda|\alpha, \theta) \, d\lambda
\]

\[
= \int_0^\infty \frac{e^{-\lambda} \lambda^x}{x!} \frac{\theta^3}{\theta^2 + \alpha} \left( 1 + \frac{\alpha \lambda^2}{2} \right) e^{-\theta \lambda} \, d\lambda
\]

\[
= \frac{\theta^3}{x!(\theta^2 + \alpha)} \left\{ \int_0^\infty e^{-\lambda(1+\theta)} \lambda^x \, d\lambda + \frac{\alpha}{2} \int_0^\infty e^{-\lambda(1+\theta)} \lambda^{x+2} \, d\lambda \right\}
\]

\[
= \frac{\theta^3}{x!(\theta^2 + \alpha)} \left\{ \frac{\Gamma(x + 1)}{(1 + \theta)^{x+1}} + \frac{\alpha}{2} \frac{\Gamma(x + 3)}{(1 + \theta)^{x+3}} \right\}
\]

\[
= \frac{\theta^3}{(\theta^2 + \alpha)(1 + \theta)^{x+1}} \left\{ 1 + \frac{\alpha(x + 1)(x + 2)}{2(1 + \theta)^2} \right\}.
\]

We have employed the gamma function defined by \( \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \, dt \), with the relation \( \Gamma(m) = (m - 1)! \) for any positive integer \( m \). The proof is completed. \( \square \)

Now, for \( \alpha = \theta = \lambda \), the pmf of the PMiD reduces to

\[
p(x; \lambda) = \frac{\lambda^2}{2(1 + \lambda)^{x+1}} \left\{ 2(1 + \lambda)^2 + \lambda(x + 1)(x + 2) \right\}, \quad x = 0, 1, 2, \ldots \tag{2}
\]

The expression in Equation (2) is the pmf of Poisson–Xgamma distribution (PXGD), which was introduced by [13]. Thus, the PXGD is a special case of the PMiD.

Now, the possible pmf plots for various values of the parameters of the PMiD are portrayed in Figure 2.
Figure 2. Plots of the pmf of the PMiD.

The pmf appears to be declining, growing, and unimodal, with some fluctuation in the mode and tails.

3.2. Mode

The next result clarifies the mode analysis of the PMiD.

Proposition 2. Let $X$ be a random variable following a PMiD. Then, if $\alpha \geq \frac{8\theta^2(1+\theta)^2}{(\theta+2)^2}$, the mode of $X$, denoted by $x_m$, exists, and lies in either of these two cases:

$$\begin{align*}
\max\{a_1(\alpha, \theta), b_2(\alpha, \theta)\} & \leq x_m \leq b_1(\alpha, \theta) \\
\text{or} \quad a_1(\alpha, \theta) & \leq x_m \leq \min\{b_1(\alpha, \theta), a_2(\alpha, \theta)\},
\end{align*}$$

where

$$\begin{align*}
a_1(\alpha, \theta) &= \frac{\alpha(2-\theta) - \sqrt{\epsilon}}{2a\theta}, \quad b_1(\alpha, \theta) = \frac{\alpha(2-\theta) + \sqrt{\epsilon}}{2a\theta}, \\
a_2(\alpha, \theta) &= \frac{\alpha(2-3\theta) - \sqrt{\epsilon}}{2a\theta}, \quad b_2(\alpha, \theta) = \frac{\alpha(2-3\theta) + \sqrt{\epsilon}}{2a\theta},
\end{align*}$$

(3)
with \( e = a^2(\theta + 2)^2 - 8a\theta^2(1 + \theta)^2 \). The condition \( a \geq \frac{8\theta^2(1+\theta)^2}{(\theta+2)^2} \) is to ensure that \( e \geq 0 \), such that \( \sqrt{e} \) has sense.

**Proof.** We have to find the integer \( x = x_m \), for which \( p(x; \alpha, \theta) \) takes its maximum value. That is, we aim to solve \( p(x; \alpha, \theta) \geq p(x-1; \alpha, \theta) \) and \( p(x; \alpha, \theta) \geq p(x+1; \alpha, \theta) \), which is equivalent to solving:

\[
a\theta x^2 + a(\theta - 2)x + 2[\theta(1 + \theta)^2 - a] \leq 0,
\]

and

\[
a\theta x^2 + a(3\theta - 2)x + 2[a(\theta - 2) + \theta(1 + \theta)^2] \geq 0,
\]

respectively. On solving the quadratic inequality in Equation (4), we obtain

\[
\left\{ \begin{array}{l}
\max\{a_1(\alpha, \theta), b_2(\alpha, \theta)\} \leq x_m \leq b_1(\alpha, \theta) \\
\text{or}
\end{array} \right.
\]

\[
a_1(\alpha, \theta) \leq x_m \leq \min\{b_1(\alpha, \theta), a_2(\alpha, \theta)\}.
\]

This completes the proof. \( \square \)

### 3.3. Cdf and Hf

The corresponding cdf of the PMiD is given by

\[
F(x; \alpha, \theta) = \Pr(X \leq x) = \frac{(\theta + 1)^{-x-3}}{2(\alpha + \theta^2)} \left[ a[2(\theta + 1)^{x+3} - \theta(x+3)(\theta(x+2) + 2) - 2]ight.
\]

\[
+ 2\theta^2(\theta + 1)^2[(\theta + 1)^{x+1} - 1] \right],
\]

and the hf of the PMiD is given by

\[
H(x; \alpha, \theta) = \frac{p(x; \alpha, \theta)}{1 - F(x; \alpha, \theta)},
\]

where \( p(x; \alpha, \theta) \) and \( F(x; \alpha, \theta) \) are respectively given in Equations (1) and (6). Furthermore, plots in Figure 3 refer to the shapes of the hf of the PMiD.
The \( hf \) is found to have all of the typical shapes, such as increasing, decreasing, bathtub, and upside-down bathtub shapes.

### 3.4. Moments

Some aspects of the PMiD are now being studied using various moment measures.
Proposition 3. The \( r \)-th factorial moment of \( X \sim \text{PMiD}(\alpha, \theta) \) is given by

\[
\mu_r = E(X(X-1) \ldots (X-r+1)) = \frac{\theta^3}{(\theta^2 + \alpha)} \frac{r!}{\theta^{r+1}} \left[ 1 + \frac{\alpha(r+1)(r+2)}{2\theta^2} \right].
\]  

(7)

Proof. Based on the compound-Poisson theory (see [14]), the \( r \)-th factorial moment of \( X \) can be obtained as follows:

\[
\mu_r = \int_0^\infty \lambda^r \frac{\theta^3}{\theta^2 + \alpha} \left( 1 + \frac{\alpha \lambda^2}{2} \right) e^{-\theta \lambda} d\lambda
\]

\[
= \frac{\theta^3}{\theta^2 + \alpha} \left\{ \int_0^\infty e^{-\theta \lambda} \lambda^r d\lambda + \frac{\alpha}{2} \int_0^\infty e^{-\theta \lambda} \lambda^{r+2} d\lambda \right\}
\]

\[
= \frac{\theta^3}{(\theta^2 + \alpha) \theta^{r+1}} \left[ 1 + \frac{\alpha(r+1)(r+2)}{2\theta^2} \right].
\]

Thus, the proof is complete. \( \square \)

The first four factorial moments of the PMiD can be obtained by substituting \( r = 1, 2, 3, \) and 4 in Equation (7). That is,

\[
\mu_1 = \frac{\theta}{\theta^2 + \alpha} \left( 1 + \frac{3\alpha}{\theta^2} \right), \quad \mu_2 = \frac{2}{\theta^2 + \alpha} \left( 1 + \frac{6\alpha}{\theta^2} \right),
\]

\[
\mu_3 = \frac{6}{\theta (\theta^2 + \alpha)} \left( 1 + \frac{10\alpha}{\theta^2} \right), \quad \text{and} \quad \mu_4 = \frac{24}{\theta^2 (\theta^2 + \alpha)} \left( 1 + \frac{15\alpha}{\theta^2} \right).
\]

Now, the first four moments about the origin of the PMiD are obtained by utilizing the general relationship between factorial moments and moments about the origin. We get

\[
\text{Mean} = \mu'_1 = E(X) = \mu_1 = \frac{\theta}{\theta^2 + \alpha} \left( 1 + \frac{3\alpha}{\theta^2} \right),
\]

(8)

\[
\mu'_2 = E(X^2) = \frac{1}{\theta^2 + \alpha} \left\{ 2 \left( 1 + \frac{6\alpha}{\theta^2} \right) + \theta \left( 1 + \frac{3\alpha}{\theta^2} \right) \right\},
\]

\[
\mu'_3 = E(X^3) = \frac{1}{\theta^2 + \alpha} \left\{ 6 \theta \left( 1 + \frac{10\alpha}{\theta^2} \right) + 6 \left( 1 + \frac{6\alpha}{\theta^2} \right) + \theta \left( 1 + \frac{3\alpha}{\theta^2} \right) \right\}
\]

and

\[
\mu'_4 = E(X^4)
\]

\[
= \frac{1}{\theta^2 + \alpha} \left\{ 24 \theta \left( 1 + \frac{15\alpha}{\theta^2} \right) + 36 \frac{\theta}{\theta^2} \left( 1 + \frac{10\alpha}{\theta^2} \right) + 14 \left( 1 + \frac{6\alpha}{\theta^2} \right) + \theta \left( 1 + \frac{3\alpha}{\theta^2} \right) \right\}.
\]

Therefore, the variance of the PMiD is obtained as

\[
\text{Var}(X) = \frac{1}{\theta^2 (\theta^2 + \alpha)} \left\{ 2(\theta^2 + 6\alpha) + \theta (\theta^2 + 3\alpha) \left[ 1 - \frac{\theta}{\theta^2 + \alpha} \left( 1 + \frac{3\alpha}{\theta^2} \right) \right] \right\}.
\]

(9)

The dispersion index of the PMiD is given by

\[
DI = 1 + \frac{2}{\theta} \left( \frac{\theta^2 + 6\alpha}{\theta^2 + 3\alpha} \right) - \frac{\theta^2 + 3\alpha}{(\theta^2 + \alpha)}.
\]

(10)
Since $\alpha$ and $\theta > 0$, one can establish that the PMiD’s dispersion index is greater than one, i.e., $DI > 1$, indicating that the PMiD is over-dispersed. The following formulae can be used to get explicit expressions for the skewness and kurtosis of the PMiD:

$$Skewness(X) = \frac{\mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3}{[\text{Var}(X)]^{3/2}}$$

and

$$Kurtosis(X) = \frac{\mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4}{[\text{Var}(X)]^2}.$$ 

Now, Table 1 shows some numerical values for the mean, variance, DI, skewness, and kurtosis for the PMiD for various parameter settings.

| Table 1. Values of some moment measures of the PMiD for various values of $\alpha$ and $\theta$. |
|---------------------------------------------|
| $\alpha = 0.5$ and Various Values of $\theta$ |
| Measures | $\theta = 1.5$ | $\theta = 3.5$ | $\theta = 5.5$ | $\theta = 7.5$ | $\theta = 9.5$ |
| Mean | 0.9091 | 0.3081 | 0.1877 | 0.1357 | 0.1064 |
| Variance | 1.7796 | 0.4085 | 0.2240 | 0.1544 | 0.1179 |
| DI | 1.9576 | 1.3256 | 1.1931 | 1.1379 | 1.1076 |
| Skewness | 2.1407 | 2.5913 | 2.9319 | 3.2482 | 3.5398 |
| Kurtosis | 9.3872 | 11.8878 | 13.6830 | 15.5978 | 17.5592 |
| $\alpha = 1.5$ and Various Values of $\theta$ |
| Measures | $\theta = 1.5$ | $\theta = 3.5$ | $\theta = 5.5$ | $\theta = 7.5$ | $\theta = 9.5$ |
| Mean | 1.2 | 0.3481 | 0.1990 | 0.1403 | 0.1087 |
| Variance | 2.4267 | 0.4792 | 0.2411 | 0.1608 | 0.1209 |
| DI | 2.0222 | 1.3769 | 1.2117 | 1.1462 | 1.1118 |
| Skewness | 1.8289 | 2.5314 | 2.9032 | 3.2258 | 3.5212 |
| Kurtosis | 7.4713 | 11.5402 | 13.5972 | 15.5194 | 17.4747 |

**Proposition 4.** The probability generating function (pgf) of $X \sim PMiD(\alpha, \theta)$ is obtained as

$$G(s) = E(s^X) = \frac{\theta^3}{(\theta^2 + \alpha)(\theta + 1 - s)} \left[ 1 + \frac{\alpha}{(\theta + 1 - s)^2} \right],$$

(11)

for $s \in (-1, 1)$.

**Proof.** Owing to the well-known compound-Poisson theory, the pgf of the PMiD is obtained as follows:

$$G(s) = \int_0^\infty e^{\lambda(s-1)} \frac{\theta^3}{\theta^2 + \alpha} \left( 1 + \frac{\alpha\lambda^2}{2} \right) e^{-\theta\lambda} d\lambda$$

$$= \frac{\theta^3}{\theta^2 + \alpha} \left[ \int_0^\infty e^{-\lambda(\theta+1-s)} d\lambda + \frac{\alpha}{2} \int_0^\infty e^{-\lambda(\theta+1-s)} \lambda^2 d\lambda \right]$$

$$= \frac{\theta^3}{\theta^2 + \alpha} \left[ \Gamma(1) \frac{\Gamma(1)}{\theta + 1 - s} + \frac{\alpha}{2} \frac{\Gamma(3)}{(\theta + 1 - s)^3} \right]$$

$$= \frac{\theta^3}{(\theta^2 + \alpha)(\theta + 1 - s)} \left[ 1 + \frac{\alpha}{(\theta + 1 - s)^2} \right].$$

Thus, the proof is complete. □
When $s$ is replaced by $e^t$ and $e^{it}$ in Equation (11), we obtain the moment generating function (mgf) and characteristic function (cf) of the PMiD, respectively. They are respectively given by

$$M(t) = \frac{\theta^3}{(\theta^2 + \alpha)(\theta + 1 - e^t)} \left[ 1 + \frac{\alpha}{(\theta + 1 - e^t)^2} \right],$$
for $t \leq 0$, and

$$\phi(t) = \frac{\theta^3}{(\theta^2 + \alpha)(\theta + 1 - e^{it})} \left[ 1 + \frac{\alpha}{(\theta + 1 - e^{it})^2} \right],$$
for $t \in \mathbb{R}$.

### 3.5. Rényi and Shannon Entropies

Entropy is a measure of uncertainty fluctuation in a stochastic situation, in which higher entropy indicates less information. The most popular entropy measures are Rényi entropy and Shannon entropy, which are among the most accessible in the literature.

For every discrete distribution with pmf $p(x)$, the related Rényi entropy is defined by

$$H_\gamma = \frac{1}{1 - \gamma} \log \sum_x p^\gamma(x),$$

for $\gamma > 0$ and $\gamma \neq 1$.

In the context of the PMiD, by using Equation (1), we obtain

$$\sum_{x=0}^{\infty} p^\gamma(x; \alpha, \theta) = \frac{\theta^{3\gamma}}{(\theta^2 + \alpha)^\gamma} \sum_{x=0}^{\infty} \left\{ \frac{1}{(1 + \theta)^{x+1}} + \frac{\alpha(x+1)(x+2)}{2(1 + \theta)^{x+3}} \right\}^{\gamma}.$$

Thus, the Rényi entropy of the PMiD is simplified to the following formula:

$$H_\gamma = \frac{1}{1 - \gamma} \left\{ \log \left[ \frac{\theta^{3\gamma}}{(\theta^2 + \alpha)^\gamma} \right] + \log \sum_{x=0}^{\infty} \left\{ \frac{1}{(1 + \theta)^{x+1}} + \frac{\alpha(x+1)(x+2)}{2(1 + \theta)^{x+3}} \right\}^{\gamma} \right\}.$$

Now, the Shannon entropy for a discrete distribution with pmf $p(x)$ is given by

$$H_1 = -\sum_{x=0}^{\infty} p(x) \log p(x).$$

Hence, the Shannon entropy for the PMiD can be expressed as

$$H_1 = -\log \left( \frac{\theta^3}{\theta^2 + \alpha} \right) - \sum_{x=0}^{\infty} p(x) \log \left[ \frac{1}{(1 + \theta)^{x+1}} + \frac{\alpha(x+1)(x+2)}{2(1 + \theta)^{x+3}} \right].$$

### 4. Estimation of the Parameters

Hereafter, we perform the estimation of parameters of the PMiD using two well-known estimation approaches: ML and Bayesian methods.

#### 4.1. Maximum Likelihood Estimation

Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from $X \sim PMiD(\alpha, \theta)$ (so $n$ independent and identically distributed (iid) random variables with the PMiD), with unknown $\alpha$ and $\theta$, and $x_1, x_2, \ldots, x_n$ be observations of $X_1, X_2, \ldots, X_n$. Then, the log-likelihood function is
\[
\log L_n = 3n \log(\theta) - n \log(\theta^2 + \alpha) - \log(1 + \theta) \sum_{i=1}^{n} (x_i + 1) \\
+ \sum_{i=1}^{n} \log \left[ 1 + \frac{\alpha(x_i + 1)(x_i + 2)}{2(1 + \theta)^2} \right].
\]

The maximization of \(\log L_n\) with respect to the parameters give their ML estimates (MLEs).

The following approach can be considered. The score function associated with this log-likelihood function is

\[
U = \left( \frac{\partial \log L_n}{\partial \alpha}, \frac{\partial \log L_n}{\partial \theta} \right)^T.
\]

Now, by solving \(\frac{\partial \log L_n}{\partial \alpha} = 0\), and \(\frac{\partial \log L_n}{\partial \theta} = 0\), we obtain the associated nonlinear log-likelihood equations. They are respectively given by

\[
\sum_{i=1}^{n} \frac{(x_i + 1)(x_i + 2)}{2(1 + \theta)^2 + \alpha(x_i + 1)(x_i + 2)} - \frac{n}{\theta^2 + \alpha} = 0,
\]

and

\[
\frac{3n}{\theta} - \frac{2n\theta}{\theta^2 + \alpha} - \sum_{i=1}^{n} \left( \frac{x_i + 1}{1 + \theta} \right) \left[ 1 - \frac{2\alpha(x_i + 2)}{2(1 + \theta)^2 + \alpha(x_i + 1)(x_i + 2)} \right] = 0.
\]

The solutions of these two equations give the MLEs. We obtained the MLEs numerically using the \texttt{fitdistrplus} package of the \texttt{R} software (see [15]). For more details on the \texttt{fitdistrplus} package, one should go through the link \url{https://CRAN.R-project.org/package=fitdistrplus} accessed on 14 February 2021. For the detailed R-code for finding the MLEs of the PMiD, see Appendix A of this article.

4.2. Bayesian Estimation

The Bayesian estimation technique is used to estimate the PMiD parameters in this subsection. That is, each parameter of PMiD must have some prior densities. For both of the parameters \(\alpha\) and \(\theta\), the half-Cauchy (hC) distribution is used as the prior densities. The hereunder is the pdf of the hC distribution with scale parameter \(\delta\):

\[
f_{hC}(u) = \frac{2}{\pi(\delta^2 + u^2)}, \quad u > 0, \quad \delta > 0.
\]

There is no mean and variance for the hC distribution. Other than that, its mode is equal to 0. The hC distribution with the value of \(\delta\) equals 25 is the preferable alternative to the uniform distribution, if more information is needed, according to [16]. Thus, as a noninformative prior distribution for the parameters \(\alpha\) and \(\theta\), we utilize hC distribution with its \(\delta\) value fixed to 25. That is, we use

\[\alpha, \theta \sim hC(25).\] (12)

Thus, using Equation (12), the joint posterior pdf is given by

\[
\psi(\alpha, \theta|x) \propto L_n \times \psi(\alpha) \times \psi(\theta),
\] (13)

where \(L_n\) is the likelihood function for the PMiD, and \(\psi(x)\) is the pdf of the hC distribution with \(\delta = 25\). It is clear from Equation (13) that there is no analytical solution for determining the Bayesian estimates. As a consequence, we adopt the Metropolis–Hastings algorithm (MHA) of the Markov Chain Monte Carlo (MCMC) technique, which is a phenomenal simulation method, using the \texttt{R} software.
4.3. Performance of the PMiD Parameters Using Simulation Study

For some finite sample sizes, we execute the simulation studies to test the long-run accuracy of the MLEs of the PMiD parameters. We have generated samples of sizes \( n = 100, 250, 500, 750, \) and 1000 from the PMiD using various sets of parameter values. The R-code for generating the PMiD random samples for the specified parameter settings are given in Appendix A. The iteration is conducted 1001 times in this case. As a consequence, we calculated the average of the biases, mean squared errors (MSEs), coverage probabilities (CPs), and average lengths (ALs) of each parameter estimate for all iterations in the relevant sample sizes. The results are reported in Table 2. It can be seen that, as the sample size increases, the MSEs and ALs were associated with each estimate decrease. Interestingly, the CPs of the confidence intervals (CIs) for each parameter are relatively close to the nominal 95 percent level. This illustrates the steady performances of the MLEs.

Table 2. The simulation results for \((\alpha = 2.5, \theta = 0.5)\).

| Parameters | \(n\) | MLE | Bias | MSE | CP | AL |
|------------|------|-----|------|-----|----|----|
| \(\alpha\) | 100  | 2.8296 | 0.3296 | 5.4612 | 0.8302 | 18.9348 |
|           | 250  | 2.9959 | 0.4959 | 4.3959 | 0.8701 | 12.1873 |
|           | 500  | 2.9402 | 0.4402 | 2.7254 | 0.9131 | 7.6664 |
|           | 750  | 2.8896 | 0.3896 | 2.1199 | 0.9171 | 5.9141 |
|           | 1000 | 2.8283 | 0.3283 | 1.5774 | 0.9610 | 4.8197 |
| \(\theta\) | 100  | 0.4922 | -0.0078 | 0.0019 | 0.9570 | 0.1855 |
|           | 250  | 0.4973 | -0.0027 | 0.00075 | 0.9630 | 0.1161 |
|           | 500  | 0.4989 | -0.0011 | 0.00038 | 0.9640 | 0.0816 |
|           | 750  | 0.4998 | -0.00023 | 0.00025 | 0.9710 | 0.0666 |
|           | 1000 | 0.5003 | 0.00031 | 0.0002 | 0.9610 | 0.0578 |

5. PMiD Regression Model

In this section, we define a new count regression model based on the PMiD known as a PMiD regression model. Let \(Y \sim PMiD(\alpha, \theta)\). By considering the re-parametrization \(\alpha = \theta^2(\theta + 1)/(3 - \theta^2)\), the \(pmf\) of the PMiD can be expressed in terms of the mean \(E(Y) = \mu > 0\) and \(\theta > 0\), and it is given by

\[
P(Y = y | \mu, \theta) = \frac{\theta (\theta - 1)}{2(1 + \theta)^{y+1}} \left[ 1 + \frac{\theta^2(\theta - 1)(y + 1)(y + 2)}{2(3 - \theta^2)(1 + \theta)^2} \right], \quad y = 0, 1, 2, \ldots \quad (14)
\]

The \(pmf\) in Equation (14) that defines a distribution is denoted as \(PMiD(\mu, \theta)\). Thus, we have \(Y \sim PMiD(\mu, \theta)\). Assume that the response variable \(Y\) satisfies \(Y \sim PMiD(\mu, \theta)\), and the covariates are in relation with the \(i^{th}\) mean by the log-link function given as follows:

\[
\mu_i = E(Y_i) = \exp\left( \mathbf{v}_i^T \mathbf{\tau} \right), \quad i = 1, 2, \ldots, n, \quad (15)
\]

where \(\mathbf{v}_i = (v_{i1}, v_{i2}, \ldots, v_{ip})\) is the vector of covariates and \(\mathbf{\tau} = (\tau_0, \tau_1, \ldots, \tau_p)^T\) is the unknown vector of regression coefficients. The log-likelihood function of the PMiD regression model is derived by inserting the log-link function of Equation (15) in Equation (14), and is given by

\[
\log L(\Theta) = n \log(\theta/2) + \sum_{i=1}^{n} \log(3 - \theta \mu_i) - \log(1 + \theta) \sum_{i=1}^{n} (y_i + 1) \left[ 1 + \frac{\theta^2(\theta \mu_i - 1)(y_i + 1)(y_i + 2)}{2(3 - \theta^2)(1 + \theta)^2} \right], \quad (16)
\]

where \(y_i\) is the \(i^{th}\) observations of \(Y\), \(\mu_i\) is given in Equation (15) for \(i = 1, 2, \ldots, n\), and \(\Theta = (\mathbf{\tau}, \theta)\) are obtained by maximizing Equation (16) using \textit{optim} function in the \textit{R} software.
6. INAR(1) Model with PMiD Innovations

The novel distribution is particularly well suited for modelling integer-valued time series data.

The stochastic process, \( \{ X_t \} \), is an INAR(1) process if it is given by

\[
X_t = p \circ X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z}
\]

where \( p \in [0, 1) \) and \( \epsilon_t \) is represented as the innovation process, thus composed of iid integer-valued random variables, with mean \( E(\epsilon_t) = \mu_\epsilon \) and variance \( Var(\epsilon_t) = \sigma^2_\epsilon \). The operator symbol ‘\( \circ \)’ represents the binomial thinning operator, which is defined as

\[
p \circ X_{t-1} = \sum_{j=1}^{X_{t-1}} B_j,
\]

where \( \{ B_j \} \) is a sequence of iid Bernoulli random variables with success probability, \( p \).

For \( p \in [0, 1) \), the INAR(1) process is stationary, while, for the case \( p = 1 \), the process is non-stationary. The INAR(1) process, according to [4,5], is a homogeneous Markov chain with 1-step transition probabilities given by

\[
P(X_t = k | X_{t-1} = l) = \min(k, l) \sum_{i=0}^{l} \binom{l}{i} p^i (1-p)^{l-i} \Pr(\epsilon_t = k - i), \quad k, l \geq 0,
\]

where \( p \in (0, 1) \). Therefore, in general, the mean, variance, and dispersion index of the INAR(1) process are respectively obtained as

\[
E(X_t) = \frac{\mu_\epsilon}{1-p},
\]

\[
Var(X_t) = \frac{p \mu_\epsilon + \sigma^2_\epsilon}{1-p^2},
\]

and

\[
DI_{X_t} = \frac{D\mu_\epsilon + p}{1+p},
\]

where \( \mu_\epsilon, \sigma^2_\epsilon \), and \( DI_\epsilon \) are given in Equations (8)–(10), respectively. For the detailed information on the INAR process, one can go through [17].

Now, in this section, based on the PMiD innovations, we present a new over-dispersed INAR(1) model, and we call it the PMiD-INAR(1) process.

**Definition 1.** Assume that the innovation process \( \{ \epsilon_t \} \) follows a PMiD. Then, the (one-step) transition probability of the PMiD-INAR(1) process is given by

\[
P_{k,l} = \Pr(X_t = k | X_{t-1} = l) = \min(k,l) \sum_{i=0}^{l} \binom{l}{i} p^i (1-p)^{l-i} \frac{\theta^3}{(\theta^2 + \alpha)(1+\theta)k^{i+1}} \left\{ 1 + \frac{\alpha(k-i+1)(k-i+2)}{2(1+\theta)^2} \right\}.
\]

Now, using Equations (17)–(19), the mean, variance, and dispersion index of the PMiD-INAR(1) process are respectively obtained as

\[
\mu_{X_t} = \frac{\theta}{(\theta^2 + \alpha)(1-p)} \left( 1 + \frac{3\alpha}{\theta^2} \right),
\]

\[
\sigma^2_{X_t} = \frac{\theta}{(\theta^2 + \alpha)(1-p^2)} \left\{ \left( 1 + \frac{3\alpha}{\theta^2} \right) \left[ p + 1 - \frac{1}{(\theta^2 + \alpha)} \left( 1 + \frac{3\alpha}{\theta^2} \right) \right] + 2 \left( 1 + \frac{6\alpha}{\theta^2} \right) \right\},
\]
and
\[ DI_{X_t} = 1 + \frac{2}{(1+p)\theta} \left[ \frac{\theta^2 + 6\alpha}{\theta^2 + 3\alpha} - \frac{\theta^2 + 3\alpha}{2\theta(\theta^2 + \alpha)} \right]. \] (20)

Since \( DI_{X_t} \) is clearly greater than 1, the process is appropriate for over-dispersed integer valued autoregressive time series data. The conditional expectation and variance of the PMiD-INAR(1) process are given by
\[
E(X_t|X_{t-1}) = E[p \circ X_{t-1}|X_{t-1}] + E[\epsilon_t|X_{t-1}]
= pX_{t-1} + \frac{\theta}{\theta^2 + \alpha} \left( 1 + \frac{3\alpha}{\theta^2} \right),
\] (21)
and
\[
Var(X_t|X_{t-1}) = Var[p \circ X_{t-1}|X_{t-1}] + Var[\epsilon_t|X_{t-1}]
= p(1-p)X_{t-1} + \sigma^2_t,
\] (22)
where \( \sigma^2_t \) is given in Equation (9) (see [17,18]).

7. Estimation of the Parameters: PMiD-INAR(1) Process

In this section, the conditional maximum likelihood (CML), conditional least squares (CLS), and Yule–Walker (YW) methods are used to investigate the inference of the PMiD-INAR(1) process. To examine the efficiency of these three strategies, a simulation study is also carried out.

7.1. Conditional Maximum Likelihood (CML) Estimation

Let \( X_1, X_2, \ldots, X_T \) be a random sample of the stationary PMiD-INAR(1) process, and \( x_1, x_2, \ldots, x_T \) be observations of \( X_1, X_2, \ldots, X_T \). Then, the conditional log-likelihood function is given by
\[
L(p, \alpha, \theta) = \sum_{t=2}^{T} \log P(X_t = x_t|X_{t-1} = x_{t-1})
= \sum_{t=2}^{T} \log \left\{ \min_{t=2}^{T} \left[ \frac{x_{t-1}}{\theta^2 + \alpha}(1+\theta)^{x_{t-1}^2+1} \right] \left[ 1 + \frac{\theta}{\theta^2 + \alpha} \left( 1 + \frac{3\alpha}{\theta^2} \right) \right] \right\}. \] (23)

By applying the direct maximization technique on Equation (23), the respective CML estimates, say \( \hat{p}_{\text{cls}}, \hat{\alpha}_{\text{cls}}, \text{ and } \hat{\theta}_{\text{cls}} \) corresponding to the parameters of the PMiD-INAR(1) process, \( p, \alpha, \text{ and } \theta \) can be obtained using the \texttt{optim} function of the \texttt{R} software.

7.2. Conditional Least Squares (CLS) Estimation

By using Equation (21), the CLS estimates of the PMiD-INAR(1) process parameters \( \zeta = (p, \alpha, \theta) \) are obtained by minimizing the function given below:
\[
S(\zeta) = \sum_{t=2}^{T} \left[ x_t - E(X_t|X_{t-1} = x_{t-1}) \right]^2
= \sum_{t=2}^{T} \left[ x_t - px_{t-1} - \frac{\theta}{\theta^2 + \alpha} \left( 1 + \frac{3\alpha}{\theta^2} \right) \right]^2.
\]
As a result, by solving $\frac{\partial}{\partial \zeta} S(\zeta) = 0$, the system of equations based on the estimates can be derived. The CLS estimate of $\alpha$ is obtained as

$$\hat{\alpha}_{CLS} = \frac{(T - 1) \sum_{t=2}^{T} x_t x_{t-1} - \sum_{t=2}^{T} x_t \sum_{t=2}^{T} x_{t-1}}{(T - 1) \sum_{t=2}^{T} x_t^2 - \left( \sum_{t=2}^{T} x_{t-1} \right)^2}.$$  

Hence, in $S(\zeta)$, $p$ is switched with $\hat{p}_{CLS}$, and the resultant function $S(\alpha, \theta)$ should be minimized by concerning $\alpha$ and $\theta$ to obtain their CLS estimates. Since $\hat{\alpha}_{CLS}$ and $\hat{\theta}_{CLS}$ do not have the closed-form expression, we utilize the `optim` function of the R software to obtain it numerically by minimizing $S(\alpha, \theta)$.

7.3. Yule–Walker (YW) Estimation

The concept of the YW approach is to synchronously solve the theoretical moments with the empirical ones. Because of the autocorrelation function (ACF) of the INAR(1) process at lag $\tau$ is $\rho_x(\tau) = p^\tau$, the YW estimate of $p$ is given by

$$\hat{p}_{YW} = \frac{\sum_{t=2}^{T} (x_t - \bar{x})(x_{t-1} - \bar{x})}{\sum_{t=1}^{T} (x_t - \bar{x})^2}.$$  

Now, the theoretical mean and dispersion index of the PMiD-INAR(1) process is then solved with their empirical equivalents to derive the YW estimates of $\alpha$ and $\theta$. When the theoretical mean is equated to the empirical mean, the YW estimate $\hat{\alpha}_{YW}$ of $\alpha$ is calculated as follows:

$$\hat{\alpha}_{YW} = \frac{\hat{\beta}_{YW}^2 \left[ \bar{x}_{YW} (1 - \hat{p}_{YW}) - 1 \right]}{3 - \bar{X}_{YW} (1 - \hat{p}_{YW})}, \hspace{1cm} (24)$$

where $\bar{X} = \frac{1}{N} \sum_{t=1}^{T} x_t$. Then, substituting $\hat{\alpha}_{YW}$ with Equation (24) in Equation (20) and equating Equation (20) to the empirical value of the dispersion index ($\hat{D}_1 X_t$), we obtain the YW estimate $\hat{\theta}_{YW}$ of $\theta$. Since $\hat{\theta}_{YW}$ does not have the closed-form expression, we utilize the `uniroot` function available in the R software to obtain it numerically.

7.4. Simulation: PMiD-INAR(1) Process

In this section, a simulation study is conducted for the comprehensive assessment of the long-standing performances of CML, YW, and CLS estimates of the PMiD-INAR(1) process parameters. We have generated samples of sizes $n = 100, 250, 500, 750,$ and $1000$ using values of the parameter setting ($p = 0.5, \alpha = 0.6, \theta = 0.7$), and the iteration process is conducted 1001 times. The simulation results are analyzed using estimated biases, MSEs, and mean relative errors (MREs), and are presented in Table 3. The following formulae are used to determine these values:

$$\text{Average bias} = \frac{1}{1001} \sum_{i=1}^{1001} (\hat{\zeta}_i - \zeta), \hspace{1cm} \text{Average MSE} = \frac{1}{1001} \sum_{i=1}^{1001} (\hat{\zeta}_i - \zeta)^2,$$

$$\text{Average MRE} = \frac{1}{1001} \sum_{i=1}^{1001} \frac{\hat{\zeta}_i}{\zeta},$$

where $\zeta = (p, \alpha, \theta)$. 
Table 3. Simulation results for \((p = 0.5, \alpha = 0.6, \theta = 0.7)\).

| \(\zeta\) | \(n\) | CML | CLS | YW |
| --- | --- | --- | --- | --- |
| Bias | MSE | MRE | Bias | MSE | MRE | Bias | MSE | MRE |
| \(p\) | 100 | -0.0029 | 0.0021 | 0.9943 | 0.0366 | 0.0079 | 0.9269 | -0.0674 | 0.0181 | 0.8653 |
| 250 | -0.0017 | 0.00095 | 0.9966 | -0.0137 | 0.0030 | 0.9725 | -0.0246 | 0.0047 | 0.9508 |
| 500 | -0.0015 | 0.00054 | 0.9984 | -0.0089 | 0.0016 | 0.9862 | -0.0101 | 0.0019 | 0.9799 |
| 750 | -0.0012 | 0.00035 | 0.9966 | -0.0056 | 0.0011 | 0.9888 | -0.0070 | 0.0013 | 0.9859 |
| 1000 | -0.0011 | 0.00029 | 0.9979 | -0.0027 | 0.0008 | 0.9946 | -0.0037 | 0.0008 | 0.9927 |
| \(\alpha\) | 100 | 0.0701 | 0.1251 | 1.1169 | 0.0826 | 0.0248 | 1.1377 | 1.8696 | 20.5181 | 4.1161 |
| 250 | 0.0628 | 0.0957 | 1.1047 | 0.0517 | 0.0107 | 1.0862 | 0.9482 | 7.7421 | 2.5804 |
| 500 | 0.0483 | 0.0740 | 1.0805 | 0.0456 | 0.0071 | 1.0760 | 0.4234 | 1.8246 | 1.7056 |
| 750 | 0.0456 | 0.0617 | 1.0760 | 0.0427 | 0.0053 | 1.0712 | 0.2460 | 0.7204 | 1.4099 |
| 1000 | 0.0421 | 0.0566 | 1.0785 | 0.0384 | 0.0044 | 1.0641 | 0.1495 | 0.2579 | 1.2492 |
| \(\theta\) | 100 | -0.0203 | 0.0120 | 0.9710 | -0.0151 | 0.0049 | 0.9784 | 0.0291 | 0.1533 | 1.0416 |
| 250 | -0.0092 | 0.0058 | 0.9868 | -0.0016 | 0.0020 | 0.9979 | -0.0083 | 0.0236 | 0.9881 |
| 500 | -0.0057 | 0.0038 | 0.9918 | 0.0014 | 0.0011 | 1.0022 | -0.0047 | 0.0091 | 0.9933 |
| 750 | -0.0032 | 0.0021 | 0.9954 | 0.0011 | 0.0008 | 1.0038 | -0.0040 | 0.0044 | 0.9943 |
| 1000 | -0.0011 | 0.0018 | 0.9984 | 0.0009 | 0.0006 | 1.0067 | -0.0021 | 0.0035 | 0.9970 |

8. Applications and Empirical Study

To show the usage of the PMiD model, we utilize three real datasets in this section: the first is COVID-19 data, the second is hospital length of stay, and the third is the number of earthquakes data.

8.1. COVID-19 Data: Armenia

First, we utilize the dataset of the daily new deaths in Armenia due to the COVID-19 disease. The data are available at https://www.worldometers.info/coronavirus/country/armenia/ accessed on the 10 January 2021 and are also studied by [19]. They contain the daily new COVID cases between 15 February 2020 and 4 October 2020. To demonstrate the PMiD’s potential benefit, the distributions given in Table 4 are considered for comparison.

Table 4. The considered competitive distributions.

| Distribution | Abbreviation | Reference |
| --- | --- | --- |
| Discrete generalized Lindley | DGLi | [19] |
| Poisson–Xgamma | PXGD | [13] |
| Discrete Poisson–Lindley | DPLi | [9] |
| Discrete Lindley | DLi | [20] |
| New Poisson-weighted exponential | NPWE | [21] |
| Poisson-transmuted exponential | PTE | [22] |
| Poisson | P | - |

We compare the competitive distributions to the suggested distribution using the statistical techniques provided, namely the negative log-likelihood (\(-\log L\)), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), \(\chi^2\) statistic, and \(p\)-value. Tables 5 and 6 display the corresponding MLEs (with standard errors (SEs) and CIs) and goodness-of-fit (GOF) results, respectively. The PMiD’s GOF statistics values are less than the other examined distributions, as shown in these tables. As a result, the suggested model is the most appropriate for modelling the given COVID-19 data. Now, the empirical mean, variance, and DI of this Armenia dataset are 4.1931, 18.7944, and 4.4822, respectively, and the theoretical values for the mean, variance, and DI measures of the PMiD are 4.1931, 19.6659, and 4.6900. It is observed that the empirical and the theoretical means are exactly the same, and the empirical and the theoretical values of variances, and DIs are the closest to each other. The empirical cdf, pmf, and P-P plots for the Armenia dataset are respectively given in Figures 4–6. It again gives some better-shaped curves for those fitted in the PMiD.
### Table 5. Armenia dataset: MLEs of the parameters.

| Distribution | α         | θ         | CI          | α         | θ         | CI          |
|--------------|-----------|-----------|-------------|-----------|-----------|-------------|
| PMiD         | 0.1029    | 0.0586    | (−0.0121, 0.2178) | 0.4162    | 0.0463    | (0.3254, 0.5070) |
| PXGD         | 0.5431    | 0.0275    | (0.4892, 0.5969) | -         | -         | -           |
| DGLi         | 0.2477    | 0.0843    | (0.0825, 0.4128) | 0.7763    | 0.0316    | (0.7144, 0.8382) |
| DPLi         | 0.41001   | 0.0227    | (0.3656, 0.4545) | -         | -         | -           |
| DLi          | 0.6914    | 0.0121    | (0.6677, 0.7151) | -         | -         | -           |
| NPWE         | 0.2167    | 3.1158    | (−5.8903, 6.3236) | 0.1008    | 15.8322   | (−30.9297, 31.1313) |
| PTE          | 0.0001    | 0.2144    | (−0.4202, 0.4202) | 0.2385    | 0.0303    | (0.1790, 0.2980) |
| P            | 4.1931    | 0.1342    | (−3.9302, 4.4561) | -         | -         | -           |

### Table 6. Armenia dataset: Goodness-of-fit test.

| X | OF | Expected Frequency |
|---|----|-------------------|
|   | PMiD | PXGD | DGLi | DPLi | DLi | NPWE | PTE | P |
| 0 | 56   | 45.1654 | 35.4419 | 43.6875 | 33.6739 | 28.4819 | 44.8674 | 44.8672 | 3.5181 |
| 1 | 31   | 35.0039 | 31.4999 | 35.8015 | 33.0940 | 36.2276 | 36.2274 | 14.7516 |
| 2 | 22   | 28.0128 | 28.7071 | 29.2575 | 30.9936 | 32.1465 | 29.2515 | 29.2514 | 30.9278 |
| 3 | 25   | 22.8835 | 25.7701 | 23.8497 | 26.9655 | 28.6318 | 23.6187 | 23.6186 | 43.2281 |
| 4 | 11   | 18.8974 | 22.5056 | 19.3972 | 22.6594 | 24.2248 | 19.0706 | 19.0705 | 45.3153 |
| 5 | 14   | 15.6646 | 19.0995 | 15.7433 | 18.5775 | 19.6108 | 15.3983 | 15.3982 | 38.0026 |
| 6 | 14   | 12.9730 | 15.7909 | 12.7535 | 14.9535 | 15.8141 | 12.4331 | 12.4331 | 26.5583 |
| 7 | 10   | 10.7033 | 12.7614 | 10.3135 | 11.8663 | 12.3974 | 10.0389 | 10.0390 | 15.9089 |
| 8 | 11   | 8.7834  | 10.1133 | 8.3269  | 9.3101  | 9.5833  | 8.1058  | 8.1059  | 8.3385  |
| 9 | 3    | 7.1637  | 7.8812  | 6.7131  | 7.2372  | 7.3255  | 6.5449  | 6.5450  | 3.8850  |
| 10| 10   | 5.8053  | 6.0535  | 5.4045  | 5.5826  | 5.5485  | 5.2846  | 5.2846  | 1.6290  |
| 11| 7    | 4.6746  | 4.5919  | 4.3455  | 4.2783  | 4.1706  | 4.2670  | 4.2670  | 6.2100  |
| 12| 4    | 3.7409  | 3.4454  | 3.4898  | 3.2605  | 3.1147  | 3.4453  | 3.4453  | 0.2170  |
| 13| 5    | 2.9762  | 2.5607  | 2.7995  | 2.4729  | 2.3133  | 2.7819  | 2.7818  | 0.0700  |
| 14| 2    | 2.3547  | 1.8870  | 2.2434  | 1.8676  | 1.7099  | 2.2462  | 2.2462  | 0.0210  |
| 15| 2    | 1.8534  | 1.3802  | 1.7960  | 1.4053  | 1.2586  | 1.8136  | 1.8140  | 0.0005  |
| ≥16| 6   | 6.3439  | 3.5106  | 7.0775  | 4.1044  | 3.3744  | 7.6048  | 7.6049  | 0.0020  |
| Total | 233 | 233 | 233 | 233 | 233 | 233 | 233 | 233 |

$-\log L$  590.3751  596.7075  592.6174  598.9318  605.3913  592.7991  592.7991  827.4472
AIC  1184.750  1195.415  1189.235  1199.864  1212.783  1189.598  1189.598  1656.894
BIC  1191.652  1198.866  1196.137  1203.315  1216.234  1198.598  1198.598  1660.345
χ²  9.3187  25.1615  12.3710  28.1328  45.7236  12.0550  12.0550  483.1912
df  6  6  6  6  6  6  6  7
p value 0.1564 < 0.001 0.0542 0.0004 < 0.001 0.0608 0.0607 < 0.001
Figure 4. Empirical cdfs of the fitted distributions for the Armenia dataset.
Figure 5. Empirical pmfs of the fitted distributions for the Armenia dataset.
The next goal is to use the Bayesian technique to estimate the parameters of the PMiD using the above-mentioned COVID-19 dataset. The analysis was carried out using the MHA of the MCMC technique with 1000 iterations in this perspective. For comparison purposes, both the Bayesian and MLE estimates of the PMiD parameters for the real dataset are given in Table 7. R programming is used to perform the numerical computations.

Table 7. MLEs and Bayes estimates of the PMiD parameters on the COVID-19 dataset.

| Parameter | ML   | Bayes  |
|-----------|------|--------|
| $\alpha$  | 0.1029 | 0.1688 |
| $\theta$  | 0.4162 | 0.4515 |
8.2. Length of Hospital Stay

By applying the PMiD regression model to an actual dataset, we can confirm its prominence. The dataset is about the 1991 Arizona cardiovascular patients (AZPRO data), which is available in the COUNT package of the R software. The goal of this study is to investigate the combined relationship between the length of hospital stay \((y_i)\) with the covariates \(x_{i1}\) - the cardiovascular procedures (CABG = 1, PTCA = 0), \(x_{i2}\) - the sex of the patients (male = 1, female = 0), \(x_{i3}\) - the type of the admission (urgent = 1, elective = 0), and \(x_{i4}\) - the age of the patients ((age > 75) = 1, (age \leq 75) = 0). The fitted nonlinear regression model is given by

\[
\mu_i = \exp(t_0 + t_1x_{i1} + t_2x_{i2} + t_3x_{i3} + t_4x_{i4}).
\]

In Table 8, we compare the performance of the PMiD regression model with the Poisson regression model, the NPWE regression model, and the PXGD regression model and also display the summaries due to the real dataset, which include the SEs, p-value, negative log-likelihood, AIC, and BIC values. Table 8 points out that the PMiD regression model has the lowest values across all the model selection criteria, indicating that it is the better count regression model than all the Poisson, NPWE, and PXGD regression models.

| Covariates | Poisson | NPWE | PXGD | PMiD |
|------------|---------|------|------|------|
| \(t_0\)    | 1.4560  | 1.3871 | 1.3996 | 2.1007 |
| \(t_1\)    | 0.9604  | 0.9982 | 0.9721 | 0.3693 |
| \(t_2\)    | -0.1240 | -0.1276 | -0.1269 | -0.0553 |
| \(t_3\)    | 0.3266  | 0.4047 | 0.1201 | 0.2087 |
| \(t_4\)    | 0.1222  | 0.1189 | 0.3732 | 0.0309 |
| \(\theta\) | -       | 0.5311 | -     | 0.3296 |
| - log L    | 11,189.90 | 11,194.01 | 10,569.80 | 10,382.38 |
| AIC        | 22,389.80 | 22,400.02 | 21,149.60 | 20,776.76 |
| BIC        | 22,420.72 | 22,437.13 | 21,180.60 | 20,813.88 |

8.3. Japan Earthquake Data

We use data from the ETAS package of the R software to calculate annual counts of earthquakes with a magnitude of 4.5 or higher that occurred in Japan between the years 1926 and 2007. The data comprise 82 observations. For more details on this package, one can go through [23]. In addition, Figure 7 depicts the Japan earthquake catalog. The spatial distribution of earthquakes in the study area is depicted in the top-left picture. The three figures in the right half of Figure 7 depict variations in the latitude, longitude, and magnitude of the earthquakes over time. The two figures in the bottom-left corner of Figure 7 depict the earthquake catalog's completeness and time stationarity. Here, the number of earthquakes having a magnitude higher than or equal to \(m\) is represented by \(N_m\). If the plot of \(\log(N_m)\) versus \(m\) exhibits a linear trend, it reflects the completeness of the earthquake catalog. Furthermore, the time stationarity of the catalog can be determined by looking at the plot of \(N_t\) versus \(t\), where \(N_t\) is the total number of occurrences in the catalog up to time \(t\). The earthquake time series is stationary if the plotted points of \(N_t\) versus \(t\) have a functional form in such a way that \(N_t = \lambda t\), where \(\lambda > 0\).

We calculated the CML estimates of parameters, as well as the negative log-likelihood \((- \log L\), AIC, and BIC of the PMiD-INAR(1) as well as the INAR(1) with the innovations, new Poisson-weighted exponential (INAR-NPWE(1)), Poisson-transmuted exponential (INAR-PTE(1)), discrete Poisson–Lindley (INAR-DPLi(1)), and Poisson (PINAR(1)), for the comparison. For more details of these innovations, see Table 4. Now, Table 9 displays the results corresponding to the earthquake data. We see that the \(- \log L\), AIC, and BIC values of the PMiD-INAR(1) model is smaller than those of the other compared INAR(1) models,
and we conclude that the PMiD-INAR(1) is the most suitable model for the earthquake data compared to that of the other models.

Figure 7. The plots of the earthquake catalog of Japan.

Table 9. The CML estimates of the fitted INAR(1) models and corresponding GOF statistics.

| Model          | Parameters | Estimate | SE   | − log L  | AIC      | BIC     |
|----------------|------------|----------|------|----------|----------|---------|
| INAR-PMiD(1)   | \( p \)    | 0.2813   | 0.0293| 446.0982 | 898.1965 | 905.4166 |
|                | \( \alpha \) | 0.6869   | 2.8522|          |          |         |
|                | \( \theta \) | 0.0247   | 0.0019|          |          |         |
| INAR-NPWE(1)   | \( p \)    | 0.3503   | 0.0204| 465.8715 | 937.7431 | 944.9632 |
|                | \( \alpha \) | 0.0068   | 0.0043|          |          |         |
|                | \( \theta \) | 0.3408   | 0.8622|          |          |         |
| INAR-PTE(1)    | \( p \)    | 0.0320   | 0.0293| 451.1948 | 908.3896 | 915.6098 |
|                | \( \alpha \) | −0.9999  | 0.2543|          |          |         |
|                | \( \theta \) | 0.0130   | 0.00124|          |          |         |
| INAR-DPLi(1)   | \( p \)    | 0.3179   | 0.0238| 450.902  | 905.804  | 910.6174 |
|                | \( \alpha \) | 0.0172   | 0.0015|          |          |         |
| PINAR(1)       | \( p \)    | 0.0592   | 0.0145| 1418.918 | 2841.836 | 2846.65  |
|                | \( \alpha \) | 158.6002 | 2.7801|          |          |         |
The residual analysis is done to make sure the fitted model is accurate for the earthquake data. To do so, the Pearson residuals for the PMiD-INAR(1) process are computed using the following formula:

\[ r_t = \frac{x_t - E(X_t|X_{t-1} = x_{t-1})}{\sqrt{\text{Var}(X_t|X_{t-1} = x_{t-1})}} \]

where \( E(X_t|X_{t-1} = x_{t-1}) \) and \( \text{Var}(X_t|X_{t-1} = x_{t-1}) \) can be found in Equations (21) and (22), respectively. In general, the mean and variance of Pearson residuals should be near zero and one, respectively, and the computed Pearson residuals should not have any autocorrelation issues if the fitted INAR(1) process is statistically accurate. Here, the obtained Pearson residuals of the PMiD-INAR(1) process have a mean and variance of 0.0012 and 1.1612, respectively, which were quite close to the anticipated values. Therefore, the PMiD-INAR(1) process is well judged for the given earthquake data. Now, the fitted PMiD-INAR(1) process is obtained as follows:

\[ X_t = 0.2813 \cdot x_{t-1} + \epsilon_t, \]

where the innovation process \( \epsilon_t \sim \text{PMiD}(0.0247, 0.0247) \). Now, we can calculate the predicted number of earthquakes in Japan via

\[
\hat{X}_1 = \frac{\hat{\theta}}{(\hat{\theta}^2 + \hat{\alpha})(1 - \hat{\rho})} \left( 1 + \frac{3\hat{\alpha}}{\hat{\theta}^2} \right) = 168.8961 \\
\hat{X}_t = \hat{\rho} x_{t-1} + \frac{\hat{\theta}}{\hat{\theta}^2 + \hat{\alpha}} \left( 1 + \frac{3\hat{\alpha}}{\hat{\theta}^2} \right) = 0.2813x_{t-1} + 121.3856.
\]

9. Discussion

9.1. Context

Discovering new count models is inevitable in the scenario of overdispersion, which will provide more possibilities for superiorly fitting the real datasets by choosing the right models according to the situations. In that sense, we proposed a new overdispersed count model, discussed its regression aspects, and constructed an INAR(1) process based on them. The main motivation behind the construction of this model is also discussed. In all the aspects, we found that the proposed model outperforms the compared models.

9.2. This Work

The Poisson–Mirra distribution (PMiD), a novel two-parameter discrete distribution, is introduced and thoroughly investigated. We delivered specific expressions for the factorial moments, mean, variance, dispersion index, skewness, kurtosis, mode, probability generating function, moment generating function, characteristic function, and the entropy measures. The distribution parameters were estimated by using the classical maximum likelihood and Bayesian estimation methods and were also studied in their real-data analysis. A simulation study on MLE was also conducted. A new regression model for count data based on the PMiD is proposed and compared with its competitive regression models based on a real dataset. More importantly, we introduced the integer-valued autoregressive model based on the PMiD, known as the PMiD-INAR(1) process. The parameter estimation problems, which include the conditional maximum likelihood, conditional least squares, and the Yule–Walker estimation procedures for the PMiD-INAR(1) process, are discussed, and simulation studies based on these three estimation procedures are also conducted. In total, three real-world datasets were used to demonstrate the use of the novel model, the first regarding COVID-19 data, the second regarding the length of hospital stay, and the third concerning earthquake data.

9.3. Contributions and Limitations

The article thus developed the new distribution PMiD from which we derived a new count model, its regression model, and the first-order integer-valued autoregressive model.
The proposed distribution, we believe, is ideal for researching data from COVID-19 related areas, biology, and earthquake-related fields, and we hope the PMiD applies to other fields of study also. The absence of a bimodal feature is the possible limitation of the proposed distribution.

9.4. Future Work

This research could take another route if the bivariate version of the PMiD and the associated BINAR(1) process are built. This work needs considerable modifications and studies, which we shall delegate to future research.

10. Conclusions

In this article, we fitted three real datasets and concluded that the PMiD model outperforms all the compared models in all aspects. We anticipate that the proposed model will increase its prevalence and have a wider variety of applications in the modelling of positive integer-valued real-world datasets from different study areas such as physics, astronomy, survival analysis, and so on.

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

Funding: This research received no external funding.

Acknowledgments: We appreciate the constructive feedback from the three reviewers on the first version of the manuscript.

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

Abbreviations

The following abbreviations are used in this manuscript:

INAR(1)  First-order Integer-valued Autoregressive
NB  Negative Binomial
PL  Poisson–Lindley
PMiD  Poisson–Mirra Distribution
MiD  Mirra Distribution
XGD  Xgamma Distribution
pdf  Probability Density Function
cdf  Cumulative Distribution Function
sf  Survival Function
hf  Hazard Function
pmf  Probability Mass Function
PXGD  Poisson–Xgamma Distribution
DI  Dispersion Index
pgf  Probability Generating Function
mgf  Moment Generating Function
cf  Characteristic Function
MLE  Maximum Likelihood Estimate
hC  half-Cauchy
MHA  Metropolis–Hastings Algorithm
MCMC  Markov Chain Monte Carlo
MSE  Mean Squared Error
CP  Coverage Probability
AL  Average Length
CML  Conditional Maximum Likelihood
Appendix A

R-code for generating the PMiD random samples is given by

```r
ppmid <- function(q, alpha, theta){
p <- (1/(2*(alpha+theta^2))) * ((theta+1)^(-q-3)) * 
(alpha*(2*((theta+1)^(q+3)) - theta*(q+3)*(theta*(q+2)+2)-2) + 
2*(theta^2)*((theta+1)^2)* (((theta+1)^(q+1)) - 1))
return(p)
}

rpmid <- function(n, alpha, theta)
{
U <- runif(n)
X <- rep(0,n)
for(i in 1:n)
{
if(U[i] < ppmid(0, alpha, theta))
{
X[i] <- 0
} else
{
B = FALSE
I = 0
while(B == FALSE)
{
int <- c( ppmid(I, alpha, theta), ppmid(I+1, alpha, theta) )
if( (U[i] > int[1]) & (U[i] < int[2]) )
{
X[i] <- I+1
B = TRUE
} else
{
# If not, continue the while loop and increase I by 1
I=I+1
}
}
}
return(X)
}.```
R-code for finding the MLEs and GOF statistics values of the PMiD is given by

```r
library(fitdistrplus)

dpmid <- function(x, alpha, theta){
d = (theta^3)/((theta^2)+alpha) * 1/((1+theta)^(x+1)) *
    (1 + (alpha*(x+1)*(x+2))/(2*(1+theta)^2))
return(d)
}

ppmid <- function(q, alpha, theta){
p = (1/(2*(alpha+theta^2))) * ((theta+1)^(-q-3)) *
    (alpha*2*((theta+1)^(q+3)) - theta*(q+3)*(theta*(q+2)+2)-2) +
    2*(theta^2)*((theta+1)^2)*(((theta+1)^(q+1)) - 1))
return(p)
}

pre <- prefit(x, "pmid", "mle", list(alpha=0.3, theta=0.3),
    lower=c(0, 0), upper = c(Inf, Inf))

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

summary(fit.pmid)
gofstat(fit.pmid).
```

References
1. Rigby, R.; Stasinopoulos, D.; Akantziliotou, C. A framework for modelling overdispersed count data, including the poisson-shifted generalized inverse gaussian distribution. *Comput. Stat. Data Anal.* 2008, 53, 381–393. [CrossRef]
2. Sellers, K.F.; Raim, A. A flexible zero-inflated model to address data dispersion. *Comput. Stat. Data Anal.* 2016, 99, 68–80. [CrossRef]
3. Contreras-Reyes, J.E. Lerch distribution based on maximum nonsymmetric entropy principle: Application to Conway’s game of life cellular automaton. *Chaos Solitons Fractals* 2021, 151, 111272. [CrossRef]
4. Al-Osh, M.A.; Alzaid, A.A. First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* 1987, 8, 261–275. [CrossRef]
5. McKenzie, E. Some simple models for discrete variate time series1. *JAWRA J. Am. Water Resour. Assoc.* 1985, 21, 645–650. [CrossRef]
6. McKenzie, E. Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. *Adv. Appl. Probab.* 1986, 18, 679–705. [CrossRef]
7. Jung, R.; Ronning, G.; Tremayne, A. Estimation in conditional first order autoregression with discrete support. *Stat. Pap.* 2005, 46, 195–224. [CrossRef]
8. Jazi, M.A.; Jones, G.; Lai, C.-D. Integer valued ar(1) with geometric innovations. *J. Iran. Stat. Soc.* 2012, 11, 173–190.
9. Livio, T.; Khan, N.M.; Bourguignon, M.; Bakouch, H. An inar(1) model with poisson-lindley innovations. *Econ. Bull.* 2018, 38, 1505–1513.
10. Altun, E. A new one-parameter discrete distribution with associated regression and integer-valued autoregressive models. *Math. Slovaca* 2020, 70, 979–994. [CrossRef]
11. Sen, S.; Ghosh, S.; Al-Mofleh, H. The Mirra distribution for modeling time-to-event data sets. In *Strategic Management, Decision Theory, and Decision Science*; Sinha, B.K., Bagchi, S.B., Eds.; Springer: Singapore, 2021; pp. 59–73.
12. Sen, S.; Maiti, S.; Chandra, N. The sgamma distribution: Statistical properties and application. *J. Mod. Appl. Stat. Methods* 2016, 15, 774–788. [CrossRef]
13. Altun, E.; Cordeiro, G.M.; Ristić, M.M. An one-parameter compounding discrete distribution. *J. Appl. Stat.* 2021, 1–22. [CrossRef]
14. Sankaran, M. The discrete poisson-lindley distribution. *Biometrics* 1970, 26, 145–149. [CrossRef]
15. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021. Available online: https://www.R-project.org/ (accessed on 6 September 2021).
16. Gelman, A.; Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*; Analytical Methods for Social Research; Cambridge University Press: Cambridge, UK, 2006.

17. Weiß, C. *An Introduction to Discrete-Valued Time Series*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2018.

18. Alzaid, A.; Alosh, M. First-order integer-valued autoregressive (inar (1)) process: Distributional and regression properties. *Stat. Neerl.* **1988**, *42*, 53–61. [CrossRef]

19. El-morshedy, M.; Altun, E.; Eliwa, M.S. A new statistical approach to model the counts of novel coronavirus cases. *Math. Sci.* **2021**, *1–14*. [CrossRef]

20. Gómez-Déniz, E.; Calderín-Ojeda, E. The discrete lindley distribution: properties and applications. *J. Stat. Comput. Simul.* **2011**, *81*, 1405–1416. [CrossRef]

21. Altun, E. A new generalization of geometric distribution with properties and applications. *Commun. Stat.-Simul. Comput.* **2020**, *49*, 793–807. [CrossRef]

22. Bhati, D.; Kumawat, P.; Gómez-Déniz, E. A new count model generated from mixed poisson transmuted exponential family with an application to health care data. *Commun. Stat. Theory Methods* **2017**, *46*, 11060–11076. [CrossRef]

23. Jalilian, A. Etas: An r package for fitting the space-time etas model to earthquake data. *J. Stat. Software Code Snippets* **2019**, *88*, 1–39. [CrossRef]