

Article

Poisson Extended Exponential Distribution with Associated INAR(1) Process and Applications

Radhakumari Maya ¹ , Christophe Chesneau ^{2,*}, Anuresha Krishna ³ and Muhammed Rasheed Irshad ³ 

¹ Department of Statistics, Government College for Women, Trivandrum 695 014, Kerala, India

² Department of Mathematics, Université de Caen Basse-Normandie, F-14032 Caen, France

³ Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India

* Correspondence: christophe.chesneau@unicaen.fr

Abstract: The significance of count data modeling and its applications to real-world phenomena have been highlighted in several research studies. The present study focuses on a two-parameter discrete distribution that can be obtained by compounding the Poisson and extended exponential distributions. It has tractable and explicit forms for its statistical properties. The maximum likelihood estimation method is used to estimate the unknown parameters. An extensive simulation study was also performed. In this paper, the significance of the proposed distribution is demonstrated in a count regression model and in a first-order integer-valued autoregressive process, referred to as the INAR(1) process. In addition to this, the empirical importance of the proposed model is proved through three real-data applications, and the empirical findings indicate that the proposed INAR(1) model provides better results than other competitive models for time series of counts that display overdispersion.

Keywords: extended exponential distribution; overdispersion; simulation; regression model; INAR(1) process

MSC: 60E05; 62E10; 62F10



Citation: Maya, R.; Chesneau, C.; Krishna, A.; Irshad, M.R. Poisson Extended Exponential Distribution with Associated INAR(1) Process and Applications. *Stats* **2022**, *5*, 755–772. <https://doi.org/10.3390/stats5030044>

Academic Editor: Wei Zhu

Received: 8 July 2022

Accepted: 3 August 2022

Published: 5 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In many fields of applied sciences, such as engineering, medicine, insurance, economics, and marketing, studying and analyzing count data play a significant role. Count data sets are often modeled using a Poisson distribution. However, the Poisson distribution cannot handle overdispersed data sets. Overdispersion occurs when the variance exceeds the mean. As a consequence, many researchers have developed mixed-Poisson distributions to provide alternative models for overdispersed count data, including [1–4]. Recent studies in this area are [5–7], among others. When using count data as a response variable, Poisson regression is a popular model. It is assumed that the dependent variable's mean and variance are both identical in the Poisson regression model. There is a lot of evidence to support the overdispersion that the count data sets exhibit. Thus, the Poisson regression's theoretical premise is practically violated. In the beginning, negative binomial regression (NB) was employed to model overdispersion in the context of count regression. The Poisson-transmuted exponential linear model was introduced by [2] and applied to healthcare data sets. The generalized Poisson–Lindley linear model was introduced by [8], who showed that generalized Poisson–Lindley linear models provide better modeling abilities than Poisson and NB regression models when there is an overdispersion of data.

There are many instances of integer-valued time series in the real world, such as the number of births at a hospital in successive months, the number of accidents, the number of patients, the number of chromosome exchanges in cells, and so on. As an inaugural approach, refs. [9–11] proposed a stochastic model for integer-valued time series called INAR(1)P for a first-order non-negative integer-valued autoregressive process

with Poisson innovations. As time series of counts mostly exhibit overdispersion, the Poisson distribution is no longer applicable to the INAR(1) process. To overcome this issue, researchers have proposed different INAR(1) processes with flexible innovation distributions. Consequently, Aghababaei Jazi et al. [12] proposed an INAR(1) process with geometric innovations (INAR(1)G), Altun E. [5] presented an INAR(1) process with new Poisson weighted exponential innovation distribution (INAR(1)PWE), Altun et al. [13] introduced an INAR(1) process with Poisson quasi-xgamma innovations (INAR(1)PQX), and so on. Although these methods are excellent for overdispersed time series count data sets, they have significant drawbacks in real-world applications. By discovering more INAR(1) models, more opportunities will be available for optimally fitting real data sets by choosing those models that are most appropriate for each situation.

Therefore, this paper provide new facts on what we call a two-parameter mixed-Poisson distribution, namely the Poisson extended exponential (PEE) distribution, obtained by compounding the Poisson distribution with the extended exponential (EE) distribution proposed by [14]. The EE distribution is obtained by mixing exponential and gamma distributions. The probability density function (pdf) of the EE distribution is given by

$$f(x) = \frac{\alpha^2(1 + \beta x)e^{-\alpha x}}{\alpha + \beta}, \quad x > 0, \alpha > 0, \beta \geq 0. \quad (1)$$

It is sometimes denoted as $EE(\alpha, \beta)$ to specify the parameters. This distribution also appears in a different form in [15], presented as a two-parameter Lindley distribution. Recent statistical literature has paid a lot of attention to the EE distribution. As a result of this, an EE regression model was proposed by [16] in which the reparameterization of the EE model based on the mean is performed. In addition, de Andrade et al. [17] proposed the exponentiated generalized EE distribution. Refs. [18,19] also showed the novelty and possibility of EE distribution through their study of different generalizations of the EE model. The PEE distribution appears in [20] under a discrete two-parameter Poisson–Lindley distribution version. However, to the best of our knowledge, some of these aspects are understudied, and the goal of this research is to rehabilitate them from applied perspectives. In particular, the appealing applicability and competence of the EE regression model inspired us to present a two-parameter mixed-Poisson distribution created by compounding Poisson with the EE distribution and elucidating its regression characteristics and associated INAR(1) process.

In the rest of the paper, the sections are arranged as follows. Section 2 presents the PEE distribution and explores some of its statistical properties. The finite sample performance of the estimation method is examined in Section 3 with a simulation study for the maximum likelihood estimation of the model parameters. A regression model is discussed in Section 4. The INAR(1)PEE process is developed in Section 5 using PEE innovations. An empirical analysis of three real data sets is conducted in Section 6 to prove that the proposed model is useful when compared to some existing models. In Section 7, a few concluding remarks are presented.

2. The Poisson Extended Exponential Distribution

In the new formulation, Poisson distribution is compounded with EE distribution to produce a mixed-Poisson distribution, which is known as the PEE distribution. Let the random variable X follow the PEE distribution which holds the following stochastic representation: $X|\lambda \sim P(\lambda)$ and $\lambda|\alpha, \beta \sim EE(\alpha, \beta)$, where $\lambda > 0$, $\alpha > 0$ and $\beta \geq 0$. Then the unconditional probability mass function (pmf) of X has the following form:

$$P(x; \alpha, \beta) = \frac{\alpha^2(1 + \alpha + \beta + \beta x)}{(\alpha + \beta)(\alpha + 1)^{x+2}}, \quad x = 0, 1, 2, 3, \dots \quad (2)$$

In fact, by construction, the random variable X has the Poisson distribution with a parameter λ , and we assume that the parameter λ represents a random variable with the

EE(α, β) distribution. Then, the unconditional distribution of X is obtained by the classical method of compounding, which gives

$$\begin{aligned} P(x; \alpha, \beta) &= \int_0^\infty \frac{e^{-\lambda} \lambda^x}{x!} \frac{\alpha^2 (1 + \beta \lambda) e^{-\alpha \lambda}}{\alpha + \beta} d\lambda \\ &= \frac{\alpha^2}{(\alpha + \beta) x!} \left[\int_0^\infty e^{-\lambda(\alpha+1)} \lambda^x d\lambda + \beta \int_0^\infty e^{-\lambda(\alpha+1)} \lambda^{x+1} d\lambda \right] \\ &= \frac{\alpha^2}{(\alpha + \beta) x!} \left[\frac{\Gamma(x+1)}{(\alpha+1)^{x+1}} + \beta \frac{\Gamma(x+2)}{(\alpha+1)^{x+2}} \right] \\ &= \frac{\alpha^2 (1 + \alpha + \beta + \beta x)}{(\alpha + \beta)(\alpha + 1)^{x+2}}. \end{aligned}$$

The gamma function $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du$ was used here and the relation $\Gamma(m+1) = m!$, for any positive integer m .

The discrete two-parameter Poisson–Lindley distribution proposed by [20] has the same pmf but had a different support for the parameters, i.e., $\alpha + \beta > 0$, and merely explored its various distributional characteristics. In contrast to [20], our applied work is more focused on the count regression model and the accompanying INAR(1) process, which are of current interest. Our theoretical work adds more aspects to the aforementioned study. Different pmf shapes are presented in Figure 1 for several parameter combinations of PEE distribution. The figure unequivocally demonstrates that the PEE distribution is right skewed.

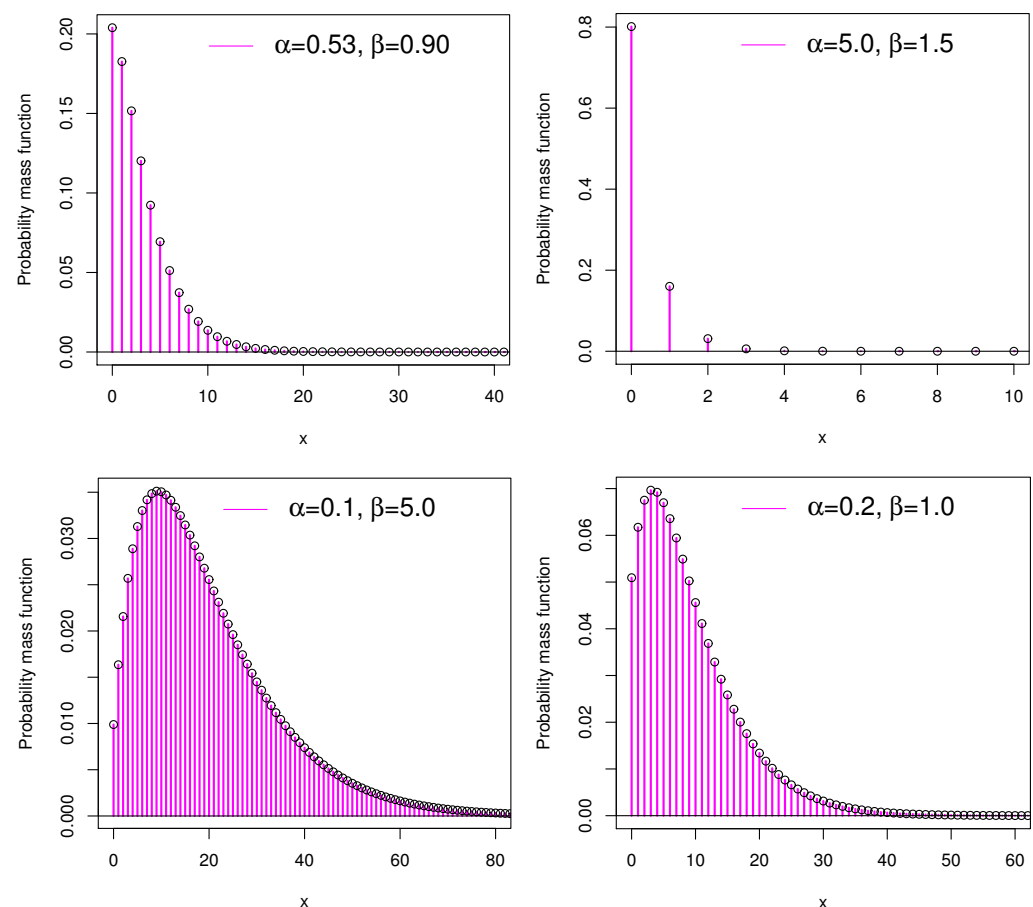


Figure 1. Various shapes of the pmfs of the PEE distribution for the varying values of the parameters.

2.1. Moments, Skewness and Kurtosis

Some results that can be derived from [20] are now presented in this portion. The probability-generating function for a random variable X with the PEE distribution is provided by

$$G(s; \alpha, \beta) = \frac{\alpha^2(1-s+\alpha+\beta)}{(\alpha+\beta)(1+\alpha-s)^2}, \quad (3)$$

for $|s| < \alpha + 1$. Correspondingly, the moment-generating function of X is given by

$$M(t; \alpha, \beta) = \frac{\alpha^2(1-e^t+\alpha+\beta)}{(\alpha+\beta)(1+\alpha-e^t)^2}, \quad (4)$$

for $t \leq \log(\alpha + 1)$. Let r be a positive integer. The r th factorial moment of a random variable X with the PEE distribution is given by

$$\mu_{[r]} = \frac{r!(\alpha + \beta + \beta r)}{\alpha^r(\alpha + \beta)}. \quad (5)$$

That is, in accordance with the definition of the r th factorial moment, we have

$$\begin{aligned} \mu_{[r]} &= \int_0^\infty \frac{\lambda^r \alpha^2 (1 + \beta \lambda) e^{-\alpha \lambda}}{\alpha + \beta} d\lambda = \frac{\alpha^2}{\alpha + \beta} \int_0^\infty \lambda^r (1 + \beta \lambda) e^{-\alpha \lambda} d\lambda \\ &= \frac{\Gamma(r+1)(\alpha + \beta + \beta r)}{\alpha^r(\alpha + \beta)}. \end{aligned}$$

From the last equality, (5) is determined by applying the relation, $\Gamma(m+1) = m!$, r being a positive integer. The first four non-central moments are derived as

$$E(X) = \frac{\alpha + 2\beta}{\alpha(\alpha + \beta)},$$

$$E(X^2) = \frac{\alpha^2 + 6\beta + 2\alpha(1 + \beta)}{\alpha^2(\alpha + \beta)},$$

$$E(X^3) = \frac{\alpha^3 + 24\beta + 2\alpha^2(3 + \beta) + 6\alpha(1 + 3\beta)}{\alpha^3(\alpha + \beta)}$$

and

$$E(X^4) = \frac{\alpha^4 + 120\beta + 2\alpha^3(7 + \beta) + 24\alpha(1 + 6\beta) + 6\alpha^2(6 + 7\beta)}{\alpha^4(\alpha + \beta)}.$$

The variance of X is given by

$$Var(X) = \frac{\alpha^3 + \alpha^2 + 4\alpha\beta + 3\alpha^2\beta + 2\beta^2 + 2\alpha\beta^2}{\alpha^2(\alpha + \beta)^2}.$$

The explicit versions of measures such as skewness and kurtosis of X can be found using the following formulas:

$$S(X) = \frac{E(X^3) - 3E(X^2)E(X) + 2[E(X)]^3}{[Var(X)]^{\frac{3}{2}}}$$

and

$$K(X) = \frac{E(X^4) - 4E(X^2)E(X) + 6E(X^2)[E(X)]^2 - 3[E(X)]^4}{[Var(X)]^2},$$

respectively.

2.2. Dispersion Index and Coefficient of Variation

The dispersion index (DI) of the PEE distribution is given by

$$DI = \frac{Var(X)}{E(X)} = \frac{\alpha^3 + 3\alpha^2\beta + \alpha^2 + 2\alpha\beta^2 + 4\alpha\beta + 2\beta^2}{\alpha(\alpha + \beta)(\alpha + 2\beta)}.$$

As a complementary measure, the coefficient of variation (CV) of the PEE distribution is given by

$$CV = \frac{\sqrt{Var(X)}}{E(X)} = \frac{\alpha(\alpha + \beta)}{\alpha + 2\beta} \sqrt{\frac{\alpha^3 + \alpha^2 + 4\alpha\beta + 3\alpha^2\beta + 2\beta^2 + 2\alpha\beta^2}{\alpha(\alpha + \beta)(\alpha + 2\beta)}}.$$

Now, Tables 1 and 2 provide some numerical values for the PEE distribution's mean, variance, and DI for a variety of parameter configurations. For the values considered, we check the mean, variance, and DI of the PEE distribution, and it is inferred that the DI of the PEE distribution is always greater than one, clearly showing overdispersion.

Table 1. Moment measure values for the PEE distribution for $\alpha = 0.5$ and various β values.

Measures	β					
	0.1	0.5	0.9	2.6	5	8
Mean	2.3333	3.0	3.2857	3.6774	3.8182	3.8824
Variance	7.5556	10.0	10.7755	11.5734	11.7851	11.8685
DI	3.2381	3.3333	3.2795	3.1471	3.0866	3.0570

Table 2. Moment measure values for the PEE distribution for $\beta = 1.5$ and various α values.

Measures	α				
	0.1	0.9	5.0	9.0	11.0
Mean	19.3750	1.8056	0.2462	0.1270	0.1018
Variance	218.9844	4.1011	0.3025	0.1426	0.1119
DI	11.3024	2.2714	1.2288	1.1230	1.0995

3. Parameter Estimation

3.1. Maximum Likelihood Estimation

Let X_1, X_2, \dots, X_n be a random sample of size n from the PEE distribution with unknown parameters α and β , and x_1, x_2, \dots, x_n be the related observations of the variables of this sample. Then the likelihood function is given by the following finite product:

$$L = \prod_{i=1}^n \frac{\alpha^2(1 + \alpha + \beta + \beta x_i)}{(\alpha + \beta)(\alpha + 1)^{x_i+2}}.$$

The maximum likelihood estimates (MLEs) of the parameters α and β , say $\hat{\alpha}$ and $\hat{\beta}$, are obtained by $(\hat{\alpha}, \hat{\beta}) = \operatorname{argmax}_{(\alpha, \beta)} L$ or, in an equivalent manner in our setting, $(\hat{\alpha}, \hat{\beta}) = \operatorname{argmax}_{(\alpha, \beta)} \log L$. To provide more practical facts, the normal equations are given by

$$\frac{\partial \log L}{\partial \alpha} = \frac{2n}{\alpha} - \frac{n}{\alpha + \beta} - \frac{2n}{\alpha + 1} + \sum_{i=1}^n \frac{1}{1 + \alpha + \beta + \beta x_i} - \sum_{i=1}^n \frac{x_i}{\alpha + 1}$$

and

$$\frac{\partial \log L}{\partial \beta} = \sum_{i=1}^n \frac{1+x_i}{1+\alpha+\beta+\beta x_i} - \frac{n}{\alpha+\beta}.$$

Then $\hat{\alpha}$ and $\hat{\beta}$ are obtained by solving the equation $\frac{\partial \log L}{\partial \alpha} = 0$ and $\frac{\partial \log L}{\partial \beta} = 0$, provided they reach a maximum well. This can only be achieved by a numerical optimization technique by using mathematical packages such as R, Mathematica and Python.

3.2. Simulation Study

The Monte Carlo simulation was performed to demonstrate the model's efficiency using the maximum likelihood method. The estimates were calculated for true values of parameters for $N = 1000$ samples of sizes 50, 75, 200, 500, 750, and 1000. The following formulas are also used to calculate indices such as MLE, bias, mean square errors (MSEs), and coverage probabilities (CPs) and average lengths (ALs) of confidence intervals (CIs).

- (i) Mean value of MLEs: $\text{MLE}(\hat{h}) = \frac{1}{N} \sum_{j=1}^N \hat{h}_j$.
- (ii) Average bias: $\text{Bias}(\hat{h}) = \frac{1}{N} \sum_{j=1}^N (\hat{h}_j - h)$.
- (iii) MSE: $\text{MSE}(\hat{h}) = \frac{1}{N} \sum_{j=1}^N (\hat{h}_j - h)^2$.
- (iv) CP of CI: $\text{CP}(\hat{h}) = \frac{1}{N} \sum_{j=1}^N I\left\{\hat{h}_j - 1.959964 \times s_{j,\hat{h}} < h < \hat{h}_j + 1.959964 \times s_{j,\hat{h}}\right\}$.
- (v) AL of CI: $\text{AL}(\hat{h}) = \frac{2 \times 1.959964}{N} \sum_{j=1}^N s_{j,\hat{h}}$.

Here, $h = \alpha$ or β , and $s_{j,h}$ and $I\{\cdot\}$ denote the standard errors (SEs) of the MLEs and indicator function, respectively. Tables 3 and 4 show the simulation results for two sets of parameter values. It has been found that MSEs and ALs of the CIs decrease with increasing sample size. The CPs of the CIs for each parameter are relatively close to the nominal 95% level.

Table 3. Simulation results for $\alpha = 0.5$ and $\beta = 0.9$.

$\alpha = 0.5, \beta = 0.9$						
Parameter	n	MLE	Bias	MSE	CP	AL
α	50	0.4839	−0.0161	0.0090	0.9970	0.5243
	75	0.4867	−0.0133	0.0071	0.9920	0.4322
	200	0.4894	−0.0106	0.0032	0.9910	0.2640
	500	0.4974	−0.0026	0.0013	0.9790	0.1560
	750	0.4996	−0.0004	0.0009	0.9720	0.1274
	1000	0.4997	−0.0003	0.0007	0.9770	0.1103
β	50	0.9605	0.0605	0.3220	0.8790	5.3376
	75	0.9548	0.0548	0.2933	0.8840	4.6352
	200	0.9478	0.0478	0.2134	0.9060	3.0661
	500	0.9405	0.0405	0.1475	0.9220	2.0805
	750	0.9297	0.0297	0.1215	0.9310	1.7183
	1000	0.9208	0.0208	0.0978	0.9380	1.4755

Table 4. Simulation results for $\alpha = 1.2$ and $\beta = 0.8$.

$\alpha = 1.2, \beta = 0.8$						
Parameter	n	MLE	Bias	MSE	CP	AL
α	50	1.1537	−0.0463	0.0832	0.9950	2.4724
	75	1.1704	−0.0296	0.0654	0.9850	1.9964
	200	1.1707	−0.0293	0.0432	0.9940	1.4188
	500	1.1739	−0.0261	0.0282	0.9840	0.9127
	750	1.1806	−0.0194	0.0222	0.9920	0.7404
	1000	1.1815	−0.0185	0.0185	0.9940	0.6468
β	50	0.8703	0.0703	0.4569	0.8950	8.5865
	75	0.8641	0.0641	0.4554	0.8770	7.3207
	200	0.8524	0.0524	0.3843	0.8770	4.9990
	500	0.8456	0.0456	0.3089	0.9220	3.5077
	750	0.8407	0.0407	0.2611	0.9440	2.8943
	1000	0.8213	0.0213	0.2365	0.9570	2.5458

4. PEE Regression Model

According to the previous section, the PEE model can model overdispersed data sets, which is critical since the majority of data in real life displays overdispersion. As a count regression model, this section uses the PEE distribution to model overdispersed data sets.

4.1. Model Construction

Let Y be a random variable representing the response variable and the number of occurrences of an event that follows the PEE distribution as well. To begin, let us consider the following reparametrization: $\beta = \frac{\alpha - \alpha^2 \mu}{\alpha \mu - 2}$. With this configuration, we obtain the pmf of the PEE distribution in terms of the mean $E(Y) = \mu > 0$ and $\alpha > 0$. Then the corresponding pmf is obtained as

$$P(Y = y; \alpha, \mu) = \frac{\alpha^2 \left[1 + \alpha + \left(\frac{\alpha - \alpha^2 \mu}{\alpha \mu - 2} \right) + \left(\frac{\alpha - \alpha^2 \mu}{\alpha \mu - 2} \right) y \right]}{\left[\alpha + \left(\frac{\alpha - \alpha^2 \mu}{\alpha \mu - 2} \right) \right] (\alpha + 1)^{y+2}}, \quad (6)$$

where $y = 0, 1, 2, \dots$. With the appropriate link functions, explanatory variables can be used to model the mean of the random variable Y . Covariates and the mean of the dependent variable can be linked using the log-link function. Let us consider Y_1, Y_2, \dots, Y_n a random sample of size n from Y . Using the log-link function, the mean of Y_i is linked to the covariate vector $x_i^T = (x_{i1}, x_{i2}, \dots, x_{ik})^T$ by the following equation:

$$\mu_i = E(Y_i) = e^{x_i^T \gamma}, \quad i = 1, 2, \dots, n, \quad (7)$$

where $\gamma = (\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_k)$ is the unknown regression coefficients. Based on (7), a linear form for the pmf of $Y_i | X_i^T = x_i^T$ which follows the PEE distribution with parameter μ_i , and α is obtained as

$$P(y_i; \alpha, e^{x_i^T \gamma}) = \frac{\alpha^2 \left[1 + \alpha + \left(\frac{\alpha - \alpha^2 e^{x_i^T \gamma}}{\alpha e^{x_i^T \gamma} - 2} \right) + \left(\frac{\alpha - \alpha^2 e^{x_i^T \gamma}}{\alpha e^{x_i^T \gamma} - 2} \right) y_i \right]}{\left[\alpha + \left(\frac{\alpha - \alpha^2 e^{x_i^T \gamma}}{\alpha e^{x_i^T \gamma} - 2} \right) \right] (\alpha + 1)^{y_i+2}}, \quad (8)$$

where y_i is the i th observations of Y .

4.2. Estimation of the Model

To estimate the regression coefficients γ , the maximum likelihood method is used. The logarithmic transformation of the likelihood function of the PEE count regression model is given by

$$\begin{aligned} \log U = 2n \log \alpha + \sum_{i=1}^n \log \left[1 + \alpha + \left(\frac{\alpha - \alpha^2 e^{x_i^T \gamma}}{\alpha e^{x_i^T \gamma} - 2} \right) + \left(\frac{\alpha - \alpha^2 e^{x_i^T \gamma}}{\alpha e^{x_i^T \gamma} - 2} \right) y_i \right] \\ - \sum_{i=1}^n \log \left[\alpha + \left(\frac{\alpha - \alpha^2 e^{x_i^T \gamma}}{\alpha e^{x_i^T \gamma} - 2} \right) \right] - \log(\alpha + 1) \sum_{i=1}^n (y_i + 2). \end{aligned} \quad (9)$$

Now the unknown parameter vector γ is obtained by maximizing (9). To accomplish this, we employ the *optim* function of R software. In addition, the SEs of these estimates are calculated using the *fdHess* function in R software.

4.3. Simulation of the PEE Regression Model

In this part, the maximum likelihood method used to estimate the unknown regression parameters is analysed using a simulation study. The parametric combinations ($\alpha = 1.5$, $\gamma_0 = 0.6$, $\gamma_1 = 0.2$, $\gamma_2 = 0.3$) and ($\alpha = 1.2$, $\gamma_0 = 0.7$, $\gamma_1 = 0.3$, $\gamma_2 = 0.4$) are used to generate $N = 1000$ samples of sizes $n = 50, 100, 200$, and 500 from the following model: $\log(\mu_i) = \gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2}$. We assume that x_{i1} and x_{i2} are generated from the uniform distribution with parameters 0 and 1, which is denoted by $U(0, 1)$. Here, indices such as estimates, bias, and MSEs are used to prove the asymptotic property of the MLEs. Table 5 reports the simulation results.

Table 5. Simulation results for the PEE regression model.

$\alpha = 1.5, \gamma_0 = 0.6, \gamma_1 = 0.2, \gamma_2 = 0.3$					$\alpha = 1.2, \gamma_0 = 0.7, \gamma_1 = 0.3, \gamma_2 = 0.4$				
n	Parameters	Estimates	Bias	MSE	n	Parameters	Estimates	Bias	MSE
50	α	0.6963	0.8037	0.6459	50	α	0.6522	0.5478	0.3001
	γ_0	0.4224	0.1776	0.7988		γ_0	0.5156	0.1844	1.1900
	γ_1	0.1745	0.0255	0.2899		γ_1	0.3329	0.0329	0.5317
	γ_2	0.1406	0.1594	0.2775		γ_2	0.1578	0.2422	0.3585
100	α	0.7289	0.7711	0.5946	100	α	0.6732	0.5268	0.2775
	γ_0	0.5077	0.0923	0.7280		γ_0	0.7876	0.0876	1.0121
	γ_1	0.2117	0.0117	0.2397		γ_1	0.2711	0.0289	0.3790
	γ_2	0.1560	0.1440	0.2486		γ_2	0.1637	0.2363	0.2430
200	α	0.7857	0.7143	0.5103	200	α	0.7135	0.4865	0.2367
	γ_0	0.6645	0.0645	0.5838		γ_0	0.6243	0.0757	0.8479
	γ_1	0.2074	0.0074	0.2277		γ_1	0.3217	0.0217	0.3546
	γ_2	0.1710	0.1290	0.2241		γ_2	0.1852	0.2148	0.2408
500	α	0.8031	0.6969	0.4856	500	α	0.7201	0.4799	0.2303
	γ_0	0.6019	0.0019	0.5448		γ_0	0.7273	0.0273	0.7893
	γ_1	0.2058	0.0058	0.2093		γ_1	0.3171	0.0171	0.3164
	γ_2	0.1712	0.1288	0.1559		γ_2	0.1963	0.2037	0.2028

From Table 5, it is clear that as sample size increases, the bias and MSEs are decreasing, implying the consistency property of the MLEs for estimating the regression parameters.

5. INAR(1) Model with PEE Innovations

The INAR(1) process is widely used in the modeling of time series of counts in several scientific disciplines, including actuarial, finance, and medical. By applying the binomial

thinning operator, INAR(1) differs from the first-order autoregressive process (AR(1)). The INAR(1) process is given by

$$X_t = p \circ X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z}, \quad (10)$$

where $0 \leq p < 1$, and the innovation process is denoted by $\{\epsilon_t\}_{t \in \mathbb{Z}}$ which are independent and identically distributed (iid) integer-valued random variables having mean, $E(\epsilon_t) = \mu_\epsilon$ and variance, $Var(\epsilon_t) = \sigma_\epsilon^2$. The binomial thinning operator is denoted by the symbol \circ and is defined as

$$p \circ X_{t-1} := \sum_{j=1}^{X_{t-1}} G_j, \quad (11)$$

where $\{G_j\}_{j \geq 1}$ is the sequence of Bernoulli random variables with probability

$$p = Pr(G_j = 1) = 1 - Pr(G_j = 0).$$

For the INAR(1) process, the one-step transition probability matrix is given by

$$Pr(X_t = k | X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} Pr(\epsilon_t = k-i), \quad k, l \geq 0, \quad (12)$$

where $0 < p < 1$. There are many examples in real life where these types of stochastic processes play a role, including the number of passengers each year, the growth of bacteria each day, the number of scientific books cited, and many more. Here, a new INAR(1) process is introduced by assuming that the $\{\epsilon_t\}$ innovations follow a PEE distribution. The one-step transition probability of the INAR(1)PEE model is given by

$$Pr(X_t = k | X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} \frac{\alpha^2 (1 + \alpha + \beta + \beta(k-i))}{(\alpha + \beta)(\alpha + 1)^{(k-i)+2}}. \quad (13)$$

So, hereafter, the described process will be called the INAR(1)PEE process.

Weiss C.H. [21] provide the mean, variance, and DI of $\{X_t\}_{t \in \mathbb{Z}}$ by using the mean, variance, and DI of the innovation distribution. For the INAR(1)PEE process, they are

$$E(X_t) = \frac{\alpha + 2\beta}{\alpha(\alpha + \beta)(1-p)}, \quad (14)$$

$$Var(X_t) = \frac{\alpha^2(\alpha + \alpha p + 1) + 2\beta^2(\alpha + \alpha p + 1) + \alpha\beta(3\alpha(p+1) + 4)}{\alpha^2(1-p^2)(\alpha + \beta)^2} \quad (15)$$

and

$$DI(X_t) = \left(\frac{1}{\alpha + \beta} - \frac{1}{\alpha + 2\beta} + \frac{1}{\alpha} + p + 1 \right) \frac{1}{p+1}. \quad (16)$$

According to [21,22], the conditional expectation and variance of the INAR(1)PEE process are given by

$$E(X_t | X_{t-1}) = pX_{t-1} + \frac{\alpha + 2\beta}{\alpha(\alpha + \beta)} \quad (17)$$

and

$$Var(X_t | X_{t-1}) = p(1-p)X_{t-1} + \frac{\alpha^3 + \alpha^2 + 4\alpha\beta + 3\alpha^2\beta + 2\beta^2 + 2\alpha\beta^2}{\alpha^2(\alpha + \beta)^2}, \quad (18)$$

respectively.

5.1. Estimation

The conditional maximum likelihood (CML), conditional least squares (CLS), and Yule–Walker (YW) methods are used to obtain the unknown parameters of the INAR(1) process.

5.1.1. Conditional Maximum Likelihood

The complicated form of the likelihood function resulting from the usual maximum likelihood method motivated the researchers to use the CML method instead of maximum likelihood. The knowledge of the transition probabilities is sufficient for the creation of likelihood in the CML technique since conditioning on the first observation results in a simple form of the likelihood, whereas there is no such conditioning present in the traditional maximum likelihood approach. The conditional log-likelihood function for the INAR(1)PEE process of the random sample X_1, X_2, \dots, X_T based on associated observations x_1, x_2, \dots, x_T is given by

$$\begin{aligned} l(p, \alpha, \beta) &= \log \left[\prod_{t=2}^T \Pr(X_t = x_t | X_{t-1} = x_{t-1}) \right] \\ &= \sum_{t=2}^T \log[\Pr(X_t = x_t | X_{t-1} = x_{t-1})], \end{aligned} \quad (19)$$

where X_1 is fixed, and $\Pr(X_t = x_t | X_{t-1} = x_{t-1})$ is given by (13). By the maximization of (19), the CML estimates are obtained by using the *constrOptim* function of R.

5.1.2. Conditional Least Squares

The below function is minimized to obtain the CLS estimates of the parameters of the INAR(1) process

$$S(p, \alpha, \beta) = \sum_{t=2}^T [x_t - E(X_t | X_{t-1} = x_{t-1})]^2.$$

5.1.3. Yule–Walker

As a result of the YW approach, the theoretical moments as well as the empirical ones are solved synchronously. Given that the autocorrelation function (ACF) of the INAR(1) process at lag η is $\rho_x(\eta) = p^\eta$, the YW estimate of p is given by

$$\hat{p}_{YW} = \frac{\sum_{t=2}^T (x_t - \bar{x})(x_{t-1} - \bar{x})}{\sum_{t=1}^T (x_t - \bar{x})^2}, \quad (20)$$

where $\bar{x} = \frac{1}{T} \sum_{t=1}^T x_t$. Now, the theoretical mean is solved with their empirical equivalents to derive the YW estimates of α and β . More precisely, when the theoretical mean equated with the empirical mean, we obtain

$$\hat{\beta}_{YW} = \frac{\hat{\alpha}_{YW}[\hat{\alpha}_{YW}\bar{x}(1 - \hat{p}_{YW}) - 1]}{2 - \hat{\alpha}_{YW}\bar{x}(1 - \hat{p}_{YW})}. \quad (21)$$

By substituting (21) in (16) and equating it with the sample dispersion, $\hat{\alpha}_{YW}$ is obtained.

5.2. Simulation

A simulation study was performed to check the finite sample performance of the CML, CLS, and YW estimates. In this regard, the number of replications is chosen as $N = 1001$ for different sample sizes, $n = 50, 100, 200, 300$, and 500 . The two parameter vectors used here are $(p = 0.5, \alpha = 0.7, \beta = 1)$ and $(p = 0.7, \alpha = 0.5, \beta = 0.8)$. The simulation results are interpreted based on the biases and MSEs. The R-code is given in Appendix A. Tables 6 and 7 show the results. The biases and MSEs of the CML estimates

are the smallest when the three estimation methods are compared, and the CML estimation approach outperforms the others. The CML estimation approach is then applied.

Table 6. Simulation for the INAR(1)PEE model for $p = 0.5$, $\alpha = 0.7$ and $\beta = 1$.

$p = 0.5, \alpha = 0.7, \beta = 1$							
Parameter	n	CML		CLS		YW	
		Bias	MSE	Bias	MSE	Bias	MSE
p	50	0.0041	0.0048	−0.0518	0.0197	−0.0614	0.0204
	100	0.0020	0.0024	−0.0258	0.0088	−0.0307	0.0089
	200	0.0013	0.0014	−0.0141	0.0047	−0.0164	0.0048
	300	0.0010	0.0008	−0.0062	0.0027	−0.0077	0.0028
	500	0.0009	0.0006	−0.0046	0.0018	−0.0057	0.0018
α	50	−0.0318	0.0233	−0.0110	0.0237	−0.4307	0.1912
	100	−0.0312	0.0159	−0.0078	0.0121	−0.4152	0.1762
	200	−0.0280	0.0115	−0.0044	0.0063	−0.3969	0.1605
	300	−0.0191	0.0088	0.0026	0.0040	−0.3865	0.1523
	500	−0.0128	0.0055	0.0016	0.0025	−0.3823	0.1482
β	50	−0.0336	0.3716	0.0969	0.0677	−0.9361	0.8767
	100	−0.0244	0.3649	0.0825	0.0422	−0.9360	0.8766
	200	−0.0215	0.3303	0.0675	0.0251	−0.9357	0.8756
	300	−0.0053	0.3024	0.0520	0.0149	−0.9352	0.8746
	500	−0.0021	0.2514	0.0514	0.0111	−0.9347	0.8737

Table 7. Simulation for the INAR(1)PEE model for $p = 0.7$, $\alpha = 0.5$, $\beta = 0.8$.

$p = 0.7, \alpha = 0.5, \beta = 0.8$							
Parameter	n	CML		CLS		YW	
		Bias	MSE	Bias	MSE	Bias	MSE
p	50	0.0007	0.0018	−0.0601	0.0182	−0.0751	0.0200
	100	0.0006	0.0010	−0.0347	0.0071	−0.0418	0.0076
	200	0.0004	0.0005	−0.0168	0.0032	−0.0206	0.0033
	300	0.0002	0.0003	−0.0095	0.0018	−0.0121	0.0018
	500	0.0001	0.0002	−0.0056	0.0011	−0.0070	0.0011
α	50	−0.0214	0.0122	−0.0270	0.0209	−0.3335	0.1144
	100	−0.0202	0.0081	−0.0214	0.0105	−0.3159	0.1021
	200	−0.0170	0.0053	−0.0089	0.0055	−0.3008	0.0923
	300	−0.0114	0.0039	−0.0020	0.0032	−0.2900	0.0855
	500	−0.0113	0.0024	−0.0010	0.0020	−0.2823	0.0810
β	50	0.1302	0.4455	0.1081	0.0866	−0.7603	0.5782
	100	0.1200	0.3912	0.0785	0.0368	−0.7574	0.5738
	200	0.1143	0.3544	0.0644	0.0209	−0.7554	0.5707
	300	0.1092	0.3145	0.0535	0.0140	−0.7545	0.5694
	500	0.0818	0.2597	0.0531	0.0168	−0.7544	0.5692

6. Empirical Studies

With the help of three real-life data sets, the superiority of the PEE model is illustrated.

6.1. Corn Borer Data

The first data set is from [23]. The data are from the biological experiment, representing the number of larvae of the European corn borer (ECB) in the field (*Pyrausta*).

Several competing distributions were compared to the fit of the PEE distribution, including the discrete Burr (DB) distribution (see [24]), the discrete log-logistic (DLL) distribution (see [25]), the discrete Gumbel (DG) distribution (see [26]), the Poisson quasi-xgamma (PQX) distribution (see [13]), the exponentiated discrete Lindley (EDL) distribution

(see [27]), the discrete Bilal (DBL) distribution (see [28]), the discrete inverse Rayleigh (DIR) distribution, and the discrete Pareto (DP) distribution (see [24]).

Utilizing the *optim* function of R, the Hessian and the Fisher information matrices are assessed. Each parameter's SE is evaluated by using the fact that the SEs can be computed as the square root of the diagonal elements of the inverse of the Fisher information matrix. As shown in Table 8, the MLEs with their corresponding SEs and confidence intervals (CIs) (lower bound of CI, upper bound of CI) for the numbers of borers data set are provided. From Table 9, it is clearly evident that the PEE distribution is the best among the considered competitive models since it has the lowest AIC, BIC, and value with the highest log L and *p*-value. The fitted PEE distribution is overdispersed since the mean and variance of the PEE distribution for the corn borer data are 1.375 and 2.2131, respectively.

Table 8. Corn borer data: MLEs, SEs and CIs.

Statistic	PEE	DB	DLL	DG	PQX	EDL	DBL	DIR	DP
MLE _{α}	1.0583	2.3570	1.9429	3.1063	0.9259	0.4691	0.6566	0.3196	0.3292
SE _{α}	0.2751	0.3655	0.1879	0.3667	0.8718	0.0421	0.0186	0.0421	0.0338
95% CI	lower _{α}	0.5190	1.6407	1.5745	2.3876	0	0.3865	0.6202	0.2370
	upper _{α}	1.5976	3.0733	2.3113	3.8250	2.6346	0.5517	0.6930	0.4022
MLE _{β}	1.4022	0.5190	1.4007	0.4067	1.3743	0.9015	-	-	-
SE _{β}	2.4893	0.0508	0.1212	0.0294	0.3391	0.1707	-	-	-
95% CI	lower _{β}	0	0.4194	1.1631	0.3492	0.7097	0.5669	-	-
	upper _{β}	6.2812	0.6186	1.6383	0.4642	2.0389	1.2361	-	-

Table 9. Corn borer data: MLE, χ^2 , *p*-value, AIC and BIC for the competitive models.

X	Of	PEE	DB	DLL	DG	PQX	EDL	DBL	DIR	DP
0	43	44.6167	43.8359	41.0317	28.5533	45.4765	44.0244	32.7337	38.3520	64.4467
1	35	30.4598	39.6006	38.9381	37.8611	29.3320	30.5905	39.5856	51.8743	20.1489
2	17	19.0658	15.6218	17.7752	25.5848	18.7843	19.4565	24.2772	15.4890	9.6863
3	11	11.3361	7.2063	8.4315	12.8520	11.5226	11.5650	12.5077	6.0275	5.6474
4	5	6.5147	3.9102	4.4846	5.7001	6.7542	6.5845	5.9702	2.9050	3.6805
5	4	3.6545	2.3755	2.6300	2.4017	3.8056	3.6386	2.7375	1.6096	2.5800
6	1	2.0132	1.5625	1.6634	0.9909	2.0750	1.9668	1.2265	0.9814	1.9042
7	2	1.0936	1.0894	1.1152	0.4054	1.1012	1.0452	0.5419	0.6414	1.4605
8	2	1.2456	4.7977	3.9304	5.6506	1.1487	1.1286	0.4198	2.1198	10.4456
Total	120	120	120	120	120	120	120	120	120	120
log L		−200.4152	−204.2933	−202.6303	−213.1911	−200.6567	−200.4922	−204.6753	−208.4404	−220.6182
AIC		404.8303	412.5865	409.2606	430.3823	405.3134	404.9844	411.3505	418.8808	443.2363
BIC		410.4053	418.1615	414.8356	435.9573	410.8883	410.5593	414.1380	421.6683	446.0238
χ^2		0.9877	2.6739	1.3113	7.6151	1.4760	1.0070	6.9961	14.2949	30.5180
df		2	2	2	2	2	2	3	3	3
<i>p</i> -value		0.6103	0.2626	0.5191	0.0222	0.4781	0.6044	0.0720	0.0025	0.0001

Figure 2 presents the estimated pmfs of all the considered models from which the distribution adequacy of the PEE model is clearly seen.

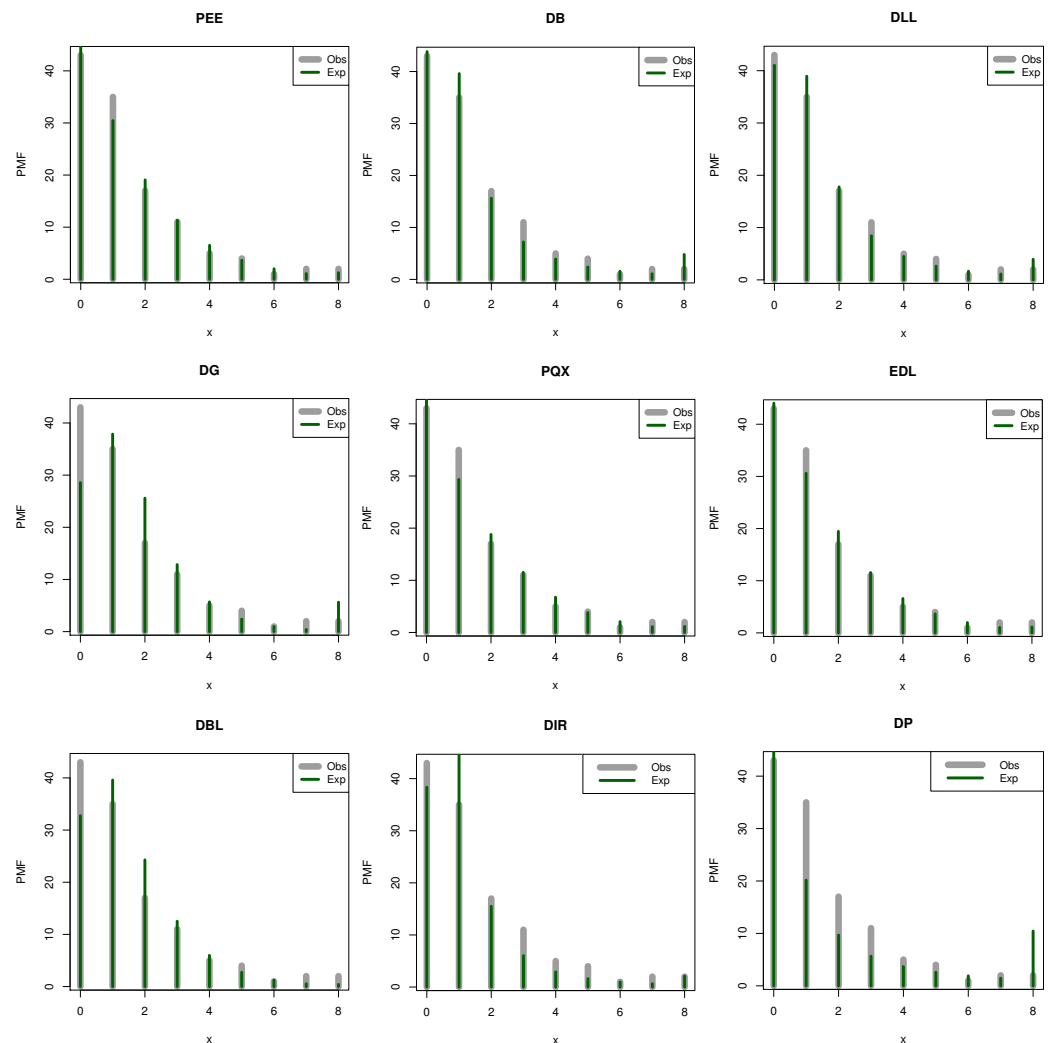


Figure 2. Pmfs of fitted models for corn borer data.

6.2. Length of Hospital Stay

The effectiveness of the count regression model under the PEE distribution is assessed using the second data set. The data consists of 3589 observations from the files of 1991 Arizona cardiovascular patients that were located in the COUNT package of the R programming language. The PEE regression model is used to model the length of stay (y_i) by using the covariates: cardiovascular procedure (x_{1i}) (1 = CABG, 0 = PTCA), sex (x_{2i}) (1 = male, 0 = female), type of admission (x_{3i}) (1 = urgent, 0 = elective), and age (x_{4i}) (1 = age > 75, 0 = age ≤ 75). Given below is the regression structure which will be fitted by the PEE distribution, the new Poisson generalized Lindley (NPGL) regression model (see [29]), the Poisson-xgamma (PX) regression model (see [7]), the Poisson–Lindley (PL) regression model and the basic Poisson regression model:

$$\mu_i = e^{\gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \gamma_3 x_{3i} + \gamma_4 x_{4i}}.$$

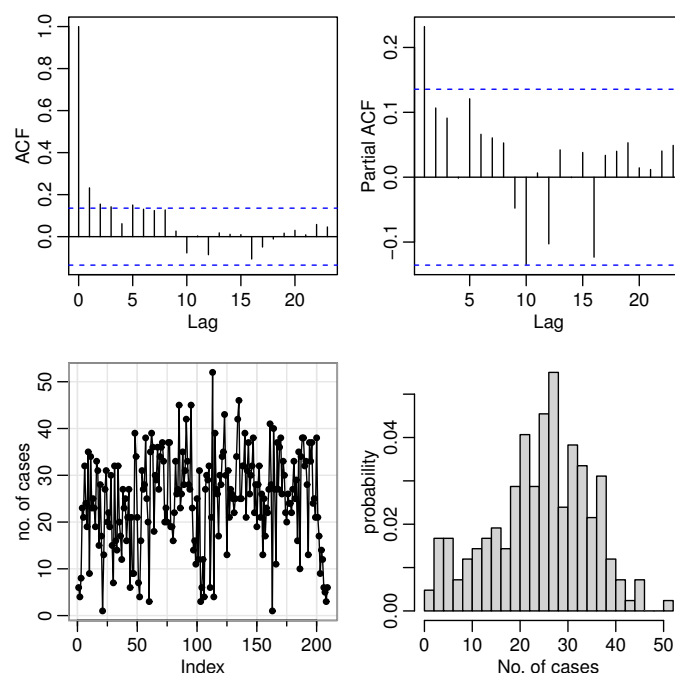
The mean and variance of the dependent variable are calculated as 8.831 and 47.973, respectively, stating the clear overdispersion. Table 10 gives the parameter estimates and results of information criterion.

Table 10. The MLE, $-\log L$, AIC and BIC of the fitted regression models for the length of stay data set.

Covariates	P		PL		PQX		NPGL		PEE			
	Estimate	SE	<i>p</i> -Value	Estimate	SE	<i>p</i> -Value	Estimate	SE	<i>p</i> -Value	Estimate	SE	<i>p</i> -Value
γ_0	1.4560	0.0158	<0.001	1.4133	0.0372	<0.001	1.3996	0.0349	<0.001	1.4044	0.0353	<0.001
γ_1	0.9606	0.0122	<0.001	0.9843	0.0291	<0.001	0.9725	0.0270	<0.001	0.9761	0.0274	<0.001
γ_2	−0.1240	0.0118	<0.001	−0.1288	0.0304	<0.001	−0.1269	0.0280	<0.001	−0.1267	0.0285	<0.001
γ_3	0.3266	0.0121	<0.001	0.3843	0.0302	<0.001	0.3732	0.0280	<0.001	0.3759	0.0284	<0.001
γ_4	0.1224	0.0124	<0.001	0.1193	0.0323	<0.001	0.1202	0.0298	<0.001	0.1198	0.0303	<0.001
−log L	−11,189.8976			−10,625.5957			−10,569.8162			−10,563.2551		
AIC	22,389.7952			21,239.1913			21,127.6324			21,114.5102		
BIC	22,420.7233			21,202.0775			21,090.5187			21,077.3964		

Table 11. Estimates and modelling adequacy statistics for the number of syphilis cases data.

Model	Parameters	Estimate	St. Error	AIC	BIC	μ_x	σ_x^2	DI _x
INAR(1)PEE	α	0.1050	0.0074	1629.8848	1639.9118	23.0431	144.3587	6.2647
	β	5.5000	7.5600					
	p	0.2365	0.0371					
INAR(1)P	λ	21.0634	0.7087	2016.5395	2023.2242	25.3493	25.3493	1
	p	0.1480	0.0261					
INAR(1)G	λ	0.0583	0.0047	1686.4277	1693.1124	23.8947	252.4312	10.5643
	p	0.3469	0.0323					
INAR(1)PWE	λ	0.0584	0.1589	1688.4277	1698.4547	24.9904	369.2114	14.7741
	α	0.0598	2.8834					
	p	0.3468	0.0323					
INAR(1)PTE	λ	−1.0000	0.0860	1637.0544	1647.0814	25.2105	278.4266	11.0441
	α	0.0788	0.0058					
	p	0.2425	0.0390					
Empirical						24.6316	105.6761	4.2903

**Figure 3.** ACF, PACF, time series, and histogram plots of weekly number of syphilis cases data.

7. Conclusions

7.1. Concluding Remarks

This paper focuses on a two-parameter discrete distribution obtained by compounding the Poisson and EE distributions and called the PEE distribution. The properties of the PEE distribution were derived and discussed. The properties, including the factorial moments, the moment-generating function, and the probability-generating functions, are evaluated, and they are in explicit forms. The article thus highlights the PEE distribution and, for the first time, its regression model and the INAR(1) model. The PEE model is found to outperform all other compared models in all aspects of the present study. In the modelling of positive integer-valued data sets from various fields of study, the proposed model is expected to increase its prevalence and have a broader variety of applications.

7.2. Future Work

This study may take a different turn if either the bivariate PEE model and its corresponding BINAR(1) model or the p th-order integer-valued auto regressive process (INAR(p)) with PEE innovations are developed. We will leave the substantial revisions, research, and software support for this effort to future studies.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would also like to thank three reviewers for their thorough comments which led to improvement in the presentation of the manuscript.

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

Appendix A

R-code for the generation of random numbers from the INAR(1)PEE process.

```
library(nleqslv)
datt=NULL
ppois=function(x,alpha,theta){
f=1-(((alpha+1)^(-(x+2)))*(alpha+alpha^2+theta+(x+2)*alpha*theta))
/(alpha+theta))
return(f)

ppois(2,0.5,1)
r.pois <- function(n, L,T)
{
U <- runif(n)
X <- rep(0,n)
for(i in 1:n)
{
if(U[i] < ppois(0,L,T))
{
X[i] <- 0
} else
{B = FALSE
I = 0
while(B == FALSE)
{int <- c( ppois(I, L,T), ppois(I+1,L,T) )
if( (U[i] > int[1]) & (U[i] < int[2]) )
{X[i] <- I+1
B = TRUE
} else
{I=I+1
}}}}
return(X)
```



```

}
r.pois(50, 1.5, 1.2)
r.inarnpl=function (n, alpha, lambda,theta, n.start = NA)
{length. <- n + n.start
x <- rep(NA, times = length.)
error <- r.pois(length., lambda,theta)
x[1] <- error[1]
for (t in (2):length.) {
x[t] <- 0
for (j in 1:1) {
x[t] <- x[t] + rbinom(1, x[t - 1], alpha)
}
x[t] <- x[t] + error[t]
}
ts(x[(n.start + 1):length.], frequency = 1, start = 1)
}
(x <- as.numeric(r.inarnpl(100, 0.5, 0.5, 0.2, 200)))

```

References

- Bereta, E.M.; Louzada, F.; Franco, M.A. The Poisson-Weibull distribution. *Adv. Appl. Stat.* **2011**, *22*, 107–118.
- 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. [\[CrossRef\]](#)
- Mahmoudi, E.; Zakerzadeh, H. Generalized Poisson–Lindley distribution. *Commun. Stat.—Theory Methods* **2010**, *39*, 1785–1798. [\[CrossRef\]](#)
- Miao, Y.; Kook, J.H.; Lu, Y.; Guindani, M.; Vannucci, M. Scalable Bayesian variable selection regression models for count data. In *Flexible Bayesian Regression Modelling*; Academic Press: Cambridge, MA, USA, 2020; pp. 187–219.
- Altun, E. A new generalization of geometric distribution with properties and applications. *Commun. Stat.—Simul. Comput.* **2020**, *49*, 793–807. [\[CrossRef\]](#)
- Altun, E. A new one-parameter discrete distribution with associated regression and integer-valued autoregressive models. *Math. Slovaca* **2020**, *70*, 979–994. [\[CrossRef\]](#)
- Altun, E.; Cordeiro, G.M.; Ristic, M.M. An one-parameter compounding discrete distribution. *J. Appl. Stat.* **2022**, *49*, 1935–1956. [\[CrossRef\]](#)
- Wongrin, W.; Bodhisuwan, W. Generalized Poisson–Lindley linear model for count data. *J. Appl. Stat.* **2017**, *44*, 2659–2671. [\[CrossRef\]](#)
- Al-Osh, M.A.; Alzaid, A.A. First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* **1987**, *8*, 261–275. [\[CrossRef\]](#)
- McKenzie, E. Some simple models for discrete variate time series 1. *JAWRA J. Am. Water Resour. Assoc.* **1985**, *21*, 645–650. [\[CrossRef\]](#)
- McKenzie, E. Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. *Adv. Appl. Probab.* **1986**, *18*, 679–705. [\[CrossRef\]](#)
- Aghababaei Jazi, M.; Jones, G.; Lai, C.D. Integer valued AR(1) with geometric innovations. *J. Iran. Stat. Soc.* **2012**, *11*, 173–190.
- Altun, E.; Bhati, D.; Khan, N.M. A new approach to model the counts of earthquakes: INARPQX(1) process. *SN Appl. Sci.* **2021**, *3*, 1–17. [\[CrossRef\]](#)
- Gómez, Y.M.; Bolfarine, H.; Gómez, H.W. A new extension of the exponential distribution. *Rev. Colomb. Estad.* **2014**, *37*, 25–34. [\[CrossRef\]](#)
- Shanker, R.; Sharma, S.; Shanker, R. A two-parameter Lindley distribution for modeling waiting and survival times data. *Appl. Math.* **2013**, *4*, 363–368. [\[CrossRef\]](#)
- Gómez, Y.M.; Gallardo, D.I.; Leao, J.; Gómez, H.W. Extended exponential regression model: Diagnostics and application to mineral data. *Symmetry* **2020**, *12*, 2042. [\[CrossRef\]](#)
- de Andrade, T.A.; Bourguignon, M.; Cordeiro, G.M. The exponentiated generalized extended exponential distribution. *J. Data Sci.* **2016**, *14*, 393–413. [\[CrossRef\]](#)
- Rasekhi, M.; Alizadeh, M.; Altun, E.; Hamedani, G.G.; Afify, A.Z.; Ahmad, M. The modified exponential distribution with applications. *Pak. J. Stat.* **2017**, *33*, 383–398.
- Rasekhi, M.; Chatrabgoun, O.; Daneshkhah, A. Discrete weighted exponential distribution: Properties and applications. *Filomat* **2018**, *32*, 3043–3056. [\[CrossRef\]](#)
- Shanker, R.; Sharma, S.; Shanker, R. A discrete two-parameter Poisson Lindley distribution. *J. Ethiop. Stat. Assoc.* **2012**, *21*, 15–22.
- Weiss, C.H. *An Introduction to Discrete-Valued Time Series*; John Wiley & Sons: Hoboken, NJ, USA, 2018.

22. Al-Osh, M.; Alzaid, A.A. Integer-valued moving average (INMA) process. *Stat. Pap.* **1988**, *29*, 281–300. [\[CrossRef\]](#)
23. Bodhisuwan, W.; Sangpoom, S. The discrete weighted Lindley distribution. In Proceedings of the 2016 12th International Conference on Mathematics, Statistics, and Their Applications (ICMSA), Banda Aceh, Indonesia, 4–6 October 2016; pp. 99–103.
24. Krishna, H.; Pundir, P.S. Discrete Burr and discrete Pareto distributions. *Stat. Methodol.* **2009**, *6*, 177–188. [\[CrossRef\]](#)
25. Para, B.A.; Jan, T.R. Discrete version of log-logistic distribution and its applications in genetics. *Int. J. Mod. Math. Sci.* **2016**, *14*, 407–422.
26. Chakraborty, S.; Chakravarty, D. A discrete Gumbel distribution. *arXiv* **2014**, arXiv:1410.7568.
27. El-Morshedy, M.; Eliwa, M.S.; Nagy, H. A new two-parameter exponentiated discrete Lindley distribution: Properties, estimation and applications. *J. Appl. Stat.* **2020**, *47*, 354–375. [\[CrossRef\]](#)
28. Altun, E.; El-Morshedy, M.; Eliwa, M.S. A study on discrete Bilal distribution with properties and applications on integer-valued autoregressive process. *RevStat Stat. J.* **2020**, *18*, 70–99.
29. Altun, E. A new two-parameter discrete poisson-generalized Lindley distribution with properties and applications to healthcare data sets. *Comput. Stat.* **2021**, *36*, 2841–2861. [\[CrossRef\]](#)
30. Altun, E.; Khan, N.M. Modelling with the novel INAR(1)-PTE process. *Methodol. Comput. Appl. Probab.* **2022**, *24*, 1735–1751. [\[CrossRef\]](#)
31. Schweer, S.; Weiß, C.H. Compound Poisson INAR (1) processes: stochastic properties and testing for overdispersion. *Comput. Stat. Data Anal.* **2014**, *77*, 267–284. [\[CrossRef\]](#)