



Research Article

# A new approach to model the counts of earthquakes: INARPQX(1) process

Emrah Altun<sup>1</sup> · Deepesh Bhati<sup>2</sup> · Naushad Mamode Khan<sup>3</sup>

Received: 26 October 2020 / Accepted: 29 December 2020  
© The Author(s) 2021 **OPEN**

## Abstract

This paper introduces a first-order integer-valued autoregressive process with a new innovation distribution, shortly INARPQX(1) process. A new innovation distribution is obtained by mixing Poisson distribution with quasi-xgamma distribution. The statistical properties and estimation procedure of a new distribution are studied in detail. The parameter estimation of INARPQX(1) process is discussed with two estimation methods: conditional maximum likelihood and Yule-Walker. The proposed INARPQX(1) process is applied to time series of the monthly counts of earthquakes. The empirical results show that INARPQX(1) process is an important process to model over-dispersed time series of counts and can be used to predict the number of earthquakes with a magnitude greater than four.

**Keywords** Discrete distribution · Earthquake · Over-dispersion · INAR(1) process

**Mathematics Subject Classification** 62E15

## 1 Introduction

Destructive earthquakes are one of the biggest problems of all humanity. Earthquakes are not only natural disasters that threaten the people's lives, but also affect the economies of the countries negatively. Therefore, statistical modeling of the counts of earthquakes and taking measures according to the model results are of great importance. Many different statistical models have been used in the literature to model earthquake events. Turkan and Ozel [26] used the linear regression, beta regression, and semi-parametric additive regression to model the casualty rate of earthquakes and related covariates. The relation between the magnitude of an earthquake and the number of deaths was investigated in several piece of researches such as Samardjieva and Badal [22], Gutiérrez et al. [10], Wyss [30], Koshimura et al. [15], Zhao et al.

[31], and Wu et al. [29]. Besides these regression models, stochastic models are also powerful tools to model earthquakes. Stochastic point process models for earthquakes were discussed by Brillinger [7], Vere-Jones [27], Ogata [19, 20] and Zhuang et al. [32, 33]. Gospodinov and Rotondi [9] and Aktas et al. [1] used the compound Poisson process to model the cumulative energy release of mainshocks and the expected value of the earthquake, respectively. Ozel [21] introduced a bivariate compound Poisson model to predict the number of foreshocks and aftershocks.

The first-order integer-valued autoregressive process, shortly INAR(1), is an alternative model to predict the number of the time of series of counts such as monthly deaths from corona-virus, monthly counts of passengers for a specific airline company or monthly counts of destructive earthquakes for a specific region or a country. The INAR(1) process with Poisson innovations, known as

✉ Emrah Altun, emrahaltun123@gmail.com; Deepesh Bhati, deepesh.bhati@curaj.ac.in; Naushad Mamode Khan, n.mamodekhan@uom.ac.mu | <sup>1</sup>Department of Mathematics, Bartin University, 74100 Bartin, Turkey. <sup>2</sup>Department of Statistics, Central University of Rajasthan, Ajmer, India. <sup>3</sup>Department of Economics and Statistics, University of Mauritius, Reduit, Mauritius.



INARP(1), was developed by McKenzie [17, 18] and Al-Osh and Alzaid [2] and increased its popularity in the last decade. The time series of counts generally display over-dispersion. The INARP(1) process can be used to model equi-dispersed time series of counts which means that the ratio of sample variance to sample mean should be one. The over-dispersion occurs when the sample variance is greater than the sample mean, the opposite indicates the under-dispersion. After the important researches of McKenzie [17, 18] and Al-Osh and Alzaid [2], the researches have focussed on the distribution of an innovation process of INAR(1) to develop new models for over-dispersed or under-dispersed time series of counts. In what follows, we list some recent contributions on overdispersed INAR(1) processes: INAR(1) process with geometric innovations (INARG(1)) by Jazi et al. [13], INAR(1) with Poisson–Lindley innovations (INARPL(1)) by Lívio et al. [16], compound Poisson INAR(1) by Schweer and Weiß [23], processes INAR(1) process with Katz family innovations by Kim and Lee [14], INAR(1) process with generalized Poisson and double Poisson innovations by Bourguignon et al. [6], INAR(1) process with geometric marginals by Borges et al. [5] and INAR(1) process with Skellam innovations by Andersson and Karlis [4], INAR(1) process with a new Poisson-weighted exponential innovations by Altun [3].

In this study, we first introduce a new two-parameter discrete distribution as an alternative distribution for the case of over-dispersion. The proposed distribution is obtained by mixing the Poisson distribution with a quasi-xgamma distribution of Sen and Chandra [24]. The resulting distribution is called as Poisson-quasi-xgamma (PQX) distribution. The statistical properties of the PQX distribution are studied in detail as well as its parameter estimation with different methods of estimations. Then, we introduce a new INAR(1) model by using the PQX distribution as an innovation process and called this process as a INARPQX(1). The goal of the presented study is to open a new opportunity to model over-dispersed time series of counts with a more flexible innovation distribution than existing ones. To demonstrate the effectiveness of the proposed process, we use the earthquake data set (magnitude 4 and above) of Turkey. The earthquake data set is regularized as a monthly basis to predict the monthly counts of the earthquakes with magnitude 4 and above occurred in Turkey. We model the monthly counts of earthquakes by INAR(1) processes with several innovation distributions as well as PQX distribution.

The ongoing sections of the study are organized as follows. In Sect. 2, the PQX distribution is presented with its statistical properties. In Sect. 3, we discuss the parameter estimation procedure of the PQX distribution with maximum likelihood and method of moments estimation methods. In Sect. 4, Expectation–Maximization (EM) algorithm for the PQX distribution is given. Section 5 deals with the simulation study to compare the finite sample performance of the estimation methods. In Sect. 6, we introduce a new INAR(1) process with PQX innovations. Section 7 contains an application to the earthquake data of Turkey. We summarize the findings of the study in Sect. 8.

## 2 Poisson-quasi-xgamma distribution

Let the random variable (rv)  $X$  follows a Poisson distribution. The probability mass function (pmf) of  $X$  is

$$\Pr(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots, \quad (1)$$

where  $\lambda > 0$ . The Poisson distribution is an attractive distribution and is widely used in many fields because of its tractable properties and software support. However, many count data sets exhibit over-dispersion or under-dispersion in which case the Poisson distribution does not work well. When the empirical variance is greater than the empirical mean, the over-dispersion occurs, and opposite case represents the under-dispersion. The negative-binomial (NB) distribution is a common choice for over-dispersed count data sets. However, in the last decade, researchers have introduced more sophisticated discrete probability distributions to open a new opportunity to model fat-tailed over-dispersed count data sets. Here, we introduce an alternative distribution to model over-dispersed time series of counts. To do this, the quasi-xgamma distribution, introduced by Sen and Chandra [24], is used as a mixing distribution for Poisson parameter  $\lambda$ . The probability density function (pdf) of quasi-xgamma distribution is

$$f(x; \alpha, \theta) = \frac{\theta}{1 + \alpha} \left( \alpha + \frac{\theta^2}{2} x^2 \right) e^{-\theta x}, x > 0 \quad (2)$$

where  $\alpha > 0$  and  $\theta > 0$ . The quasi-xgamma distribution is a special mixture of exponential and gamma densities and contains Gamma( $\theta, 3$ ) (for  $\alpha = 0$ ) and xgamma( $\theta$ ) (for  $\alpha = \theta$ ), introduced by Sen et al. [25], distributions as its submodels. Now, we introduce a new two-parameter

discrete distribution by mixing Poisson distribution with quasi-xgamma distribution.

**Proposition 1** Assume that the parameter  $\lambda$  of the Poisson distribution follows a quasi-xgamma distribution. Then, the resulting distribution is

$$\Pr(X = x) = \frac{2\alpha\theta(\theta + 1)^2 + \theta^3(x + 1)(x + 2)}{2(\alpha + 1)(\theta + 1)^{x+3}}, \quad (3)$$

$x = 0, 1, 2, \dots,$

where  $\alpha > 0$  and  $\theta > 0$ . Hereafter, we refer to (3) as PQX( $\alpha, \theta$ ).

**Proof** The proof is straightforward. □

**Remark 1** The PQX distribution contains the below distributions as its sub-models or limiting case.

1. The PQX distribution reduces to Poisson-xgamma distribution for  $\alpha = \theta$ ,
2. The PQX distribution reduces to NB(3,  $\theta/(\theta + 1)$ ) for  $\alpha = 0$ ,
3. The PQX distribution reduces to Geometric( $\theta/(\theta + 1)$ ) for  $\alpha \rightarrow \infty$ .

The corresponding cumulative distribution function (cdf) to (3) is

$$F(x) = \frac{2\alpha + (\theta + 1)^{-x-3}(-2\alpha(\theta + 1)^2 - \theta(x + 3)(\theta(x + 2) + 2) - 2) + 2}{2(\alpha + 1)}. \quad (4)$$

The rest of the section is devoted to the inference of the statistical properties of PQX distribution and its parameter estimation.

### 2.1 Statistical properties

**Proposition 2** Let the rv  $X$  follow a PQX distribution. The factorial moments of  $X$  are given by

$$\mu_{[r]} = \frac{\Gamma(r + 3) + 2\alpha\Gamma(r + 1)}{2\theta^r(\alpha + 1)}. \quad (5)$$

**Proof** The factorial moments of  $X$  can be obtained by using mixing formula, as follows

$$\begin{aligned} \mu_{[r]} &= \int_0^\infty \frac{\theta\lambda^r}{1 + \alpha} \left( \alpha + \frac{\theta^2}{2}\lambda^2 \right) e^{-\theta\lambda} d\lambda \\ &= \frac{\theta}{1 + \alpha} \int_0^\infty \lambda^r \left( \alpha + \frac{\theta^2}{2}\lambda^2 \right) \exp(-\theta\lambda) d\lambda \\ &= \frac{\theta}{1 + \alpha} \left\{ \int_0^\infty \alpha\lambda^r \exp(-\theta\lambda) d\lambda \right. \\ &\quad \left. + \frac{\theta^2}{2} \int_0^\infty \lambda^{r+2} \exp(-\theta\lambda) d\lambda \right\} \\ &= \frac{\theta}{1 + \alpha} \left\{ \frac{\alpha\Gamma(r + 1)}{\theta^{r+1}} + \frac{\theta^2\Gamma(r + 3)}{2\theta^{r+3}} \right\} \\ &= \frac{\theta}{1 + \alpha} \left\{ \frac{\alpha\Gamma(r + 1)}{\theta^{r+1}} + \frac{\Gamma(r + 3)}{2\theta^{r+1}} \right\} \\ &= \frac{\Gamma(r + 3) + 2\alpha\Gamma(r + 1)}{2\theta^r(\alpha + 1)} \end{aligned}$$

The proof is completed. □

Using (5), the mean and variance of PQX distribution are given, respectively, by

$$E(X) = \frac{2(\alpha + 3)}{\theta(2\alpha + 2)}, \quad (6)$$

$$\text{Var}(X) = \frac{\alpha^2 + (\alpha + 1)(\alpha + 3)\theta + 8\alpha + 3}{(\alpha + 1)^2\theta^2}. \quad (7)$$

The dispersion index (Var( $X$ )/E( $X$ )) of PQX is given by

$$\text{DI} = 1 + \frac{\alpha(\alpha + 8) + 3}{(\alpha + 1)(\alpha + 3)\theta}. \quad (8)$$

As seen from (8), since the parameter  $\alpha$  and  $\theta$  are greater than zero, the dispersion index of PQX distribution is always greater than 1 which ensures that the PQX distribution is over-dispersed.

**Proposition 3** The probability generating function (pgf) of PQX distribution is given by

$$G(s) = \frac{\theta}{1 + \alpha} \left[ \frac{\theta^2}{(\theta - s + 1)^3} + \frac{\alpha}{(\theta - s + 1)} \right]. \quad (9)$$

**Proof** The pgf of  $X$  can be written in the form

$$G(s) = \int_0^\infty \exp[\lambda(s-1)]f(\lambda;\beta\theta)d\lambda. \tag{10}$$

Using (10), the pgf of  $X$  can be obtained as follows

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

□

The moment generating function of PQX distribution can be easily obtained by substituting  $s$  with  $\exp(t)$  in (9). Then, we have

$$M(t) = \frac{\theta}{1+\alpha} \left[ \frac{\theta^2}{(\theta - \exp(t) + 1)^3} + \frac{\alpha}{(\theta - \exp(t) + 1)} \right]. \tag{11}$$

The quasi-xgamma distribution is a special mixture of exponential( $\theta$ ) and gamma(3,  $\theta$ ) distributions with mixing proportion,  $p = \alpha/(\alpha + 1)$ . This property of the quasi-xgamma distribution can be used to generate rvs from PQX distribution. The below algorithm is given based on the mixture property of the quasi-xgamma.

---

**Algorithm 1**

---

1. Generate  $u_i$  from  $U(0, 1)$  distribution,
  2. If  $U_i \leq \frac{\alpha}{\alpha+1}$  generate  $\lambda_i \sim \text{exponential}(\theta)$ , otherwise, generate  $\lambda_i \sim \text{gamma}(3, \theta)$ ,
  3. Generate  $Y_i \sim \text{Poisson}(\lambda_i)$
  4. Repeat steps 1,2 and 3  $n$  times.
- 

The skewness and kurtosis measures of PQX distribution can be easily obtained based on the first four raw moments by using

$$S = \frac{\alpha^3 \theta^2 + 3 \alpha^3 \theta + 2 \alpha^3 + 5 \alpha^2 \theta^2 + 27 \alpha^2 \theta + 30 \alpha^2 + 7 \alpha \theta^2 + 33 \alpha \theta + 18 \alpha + 3 \theta^2 + 9 \theta + 6}{(8 \alpha + 3 \theta + 4 \alpha \theta + \alpha^2 \theta + \alpha^2 + 3)^{\frac{3}{2}}}, \tag{12}$$

$$K = \frac{\alpha^4 \theta^3 + 10 \alpha^4 \theta^2 + 18 \alpha^4 \theta + 9 \alpha^4 + 6 \alpha^3 \theta^3 + 94 \alpha^3 \theta^2 + 264 \alpha^3 \theta + 192 \alpha^3 + 12 \alpha^2 \theta^3 + 206 \alpha^2 \theta^2 + 516 \alpha^2 \theta + 306 \alpha^2 + 10 \alpha \theta^3 + 170 \alpha \theta^2 + 360 \alpha \theta + 216 \alpha + 3 \theta^3 + 48 \theta^2 + 90 \theta + 45}{(8 \alpha + 3 \theta + 4 \alpha \theta + \alpha^2 \theta + \alpha^2 + 3)^2}. \tag{13}$$

The skewness and kurtosis plots of the PQX distribution against the mean and dispersion values are displayed in Fig. 1. The parameter  $\alpha$  is taken 1 and parameter  $\theta$  is increased in the interval (0.5, 15). From these figures, we conclude that the skewness and kurtosis decrease when the mean and dispersion increase.

The pmf of the NB distribution is

$$p(y;\alpha, p) = \frac{\Gamma(y + \theta)}{\Gamma(y + 1)\Gamma(\theta)}(1 - p)^\theta p^y, \tag{14}$$

where  $\theta > 0, p \in (0, 1)$ . The mean, variance, skewness, and kurtosis values of the NB distribution are given, respectively, by

$$\begin{aligned} E(Y) &= \frac{\theta p}{1 - p}, \\ \text{Var}(Y) &= \frac{\theta p}{(1 - p)^2}, \\ S(Y) &= \frac{1 + p}{\sqrt{\theta p}}, \\ K(Y) &= \frac{6}{\theta} + \frac{(1 - p)^2}{\theta p}. \end{aligned}$$

The skewness and kurtosis plots of the NB distribution against the mean and dispersion values are displayed in Fig. 2. The parameter  $\theta$  is taken 0.5, and parameter  $p$  is increased in the interval (0.1, 0.9). From these figures, we conclude that the skewness and kurtosis decrease when the mean and dispersion increase. So, the NB and PQX distributions have the similar behaviors. However, to understand the differences between these distributions,

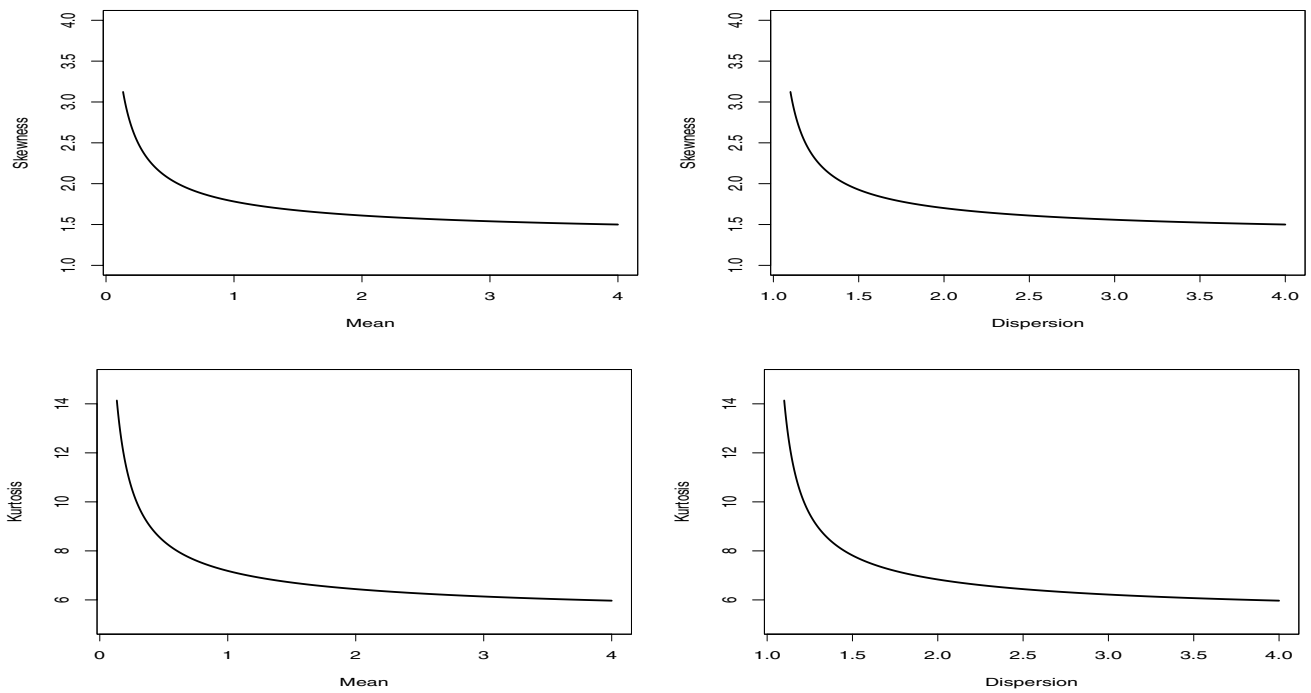


Fig. 1 The skewness and kurtosis plots of the PQX distribution against the mean and dispersion values

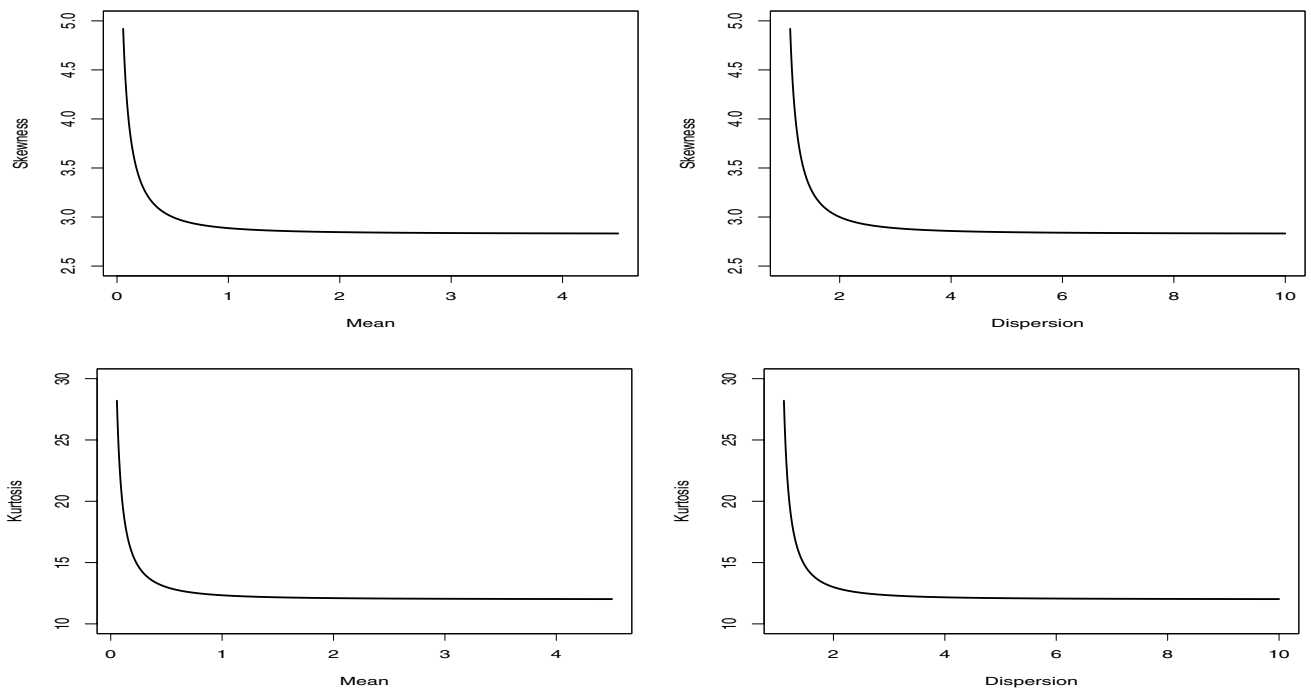
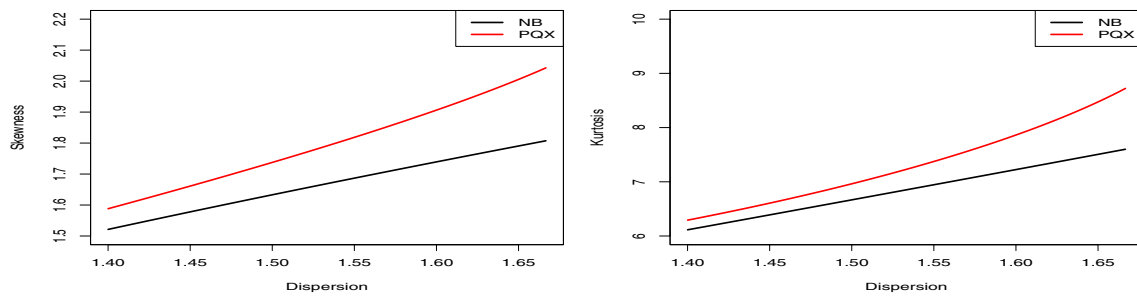


Fig. 2 The skewness and kurtosis plots of the NB distribution against the mean and dispersion values



**Fig. 3** The skewness and kurtosis comparison of the mean-parametrized NB and PQX distributions

we compare the mean-parametrized versions of both distributions. Substituting  $p = \mu / (\mu + \theta)$  in (14), the mean-parametrized NB distribution is given by

$$p(y; \alpha, \mu) = \frac{\Gamma(y + \theta)}{\Gamma(\theta)\Gamma(y + 1)} \left(\frac{\theta}{\mu + \theta}\right)^\theta \left(\frac{\mu}{\mu + \theta}\right)^y, \quad (15)$$

where  $\theta > 0$ ,  $\mu > 0$  and  $E(Y) = \mu$ . Substituting  $\alpha = (3 - \theta\mu) / (\theta\mu - 1)$  in (3), we have the pmf of the mean-parametrized PQX, given by

$$Pr(X = x) = \frac{2(3 - \theta\mu) / (\theta\mu - 1) \theta (\theta + 1)^2 + \theta^3 (x + 1)(x + 2)}{2((3 - \theta\mu) / (\theta\mu - 1) + 1) (\theta + 1)^{x+3}}, \quad (16)$$

$x = 0, 1, 2, \dots$

where  $\theta > 0$ ,  $\mu > 0$ ,  $1 < \theta\mu < 3$  and  $E(X) = \mu$ . Figure 3 displays the skewness and kurtosis values of the mean-parametrized NB and PQX distributions against their dispersion values. The mean parameter is determined as 1 for both distributions. The parameter  $\theta$  and  $\alpha$  are increased in the interval (1.5, 2.5) since we have condition on the parameter space of the mean-parametrized PQX distributions, such as  $1 < \theta\mu < 3$ . As seen from these figures, when the dispersion increases, the skewness and kurtosis increase

for both distributions. However, under these re-parametrization, the PQX distribution is able to model wider range of skewness and kurtosis than the NB distribution.

We also compare the right-tail probabilities of the mean-parametrized NB and PQX distributions. Table 1 lists the tail probabilities of the NB and PQX distributions for some values of the mean parameter  $\mu$  and  $\theta$ . The results given in Table 1 indicate that the PQX has fatter right-tail than the NB distribution.

### 2.2 Shape of PQX( $\alpha, \theta$ ) distribution

Using (3), the ratio of successive probabilities  $P(X = x + 1) / P(X = x)$  for PQX( $\alpha, \theta$ ) is given as

$$\frac{P(X = x + 1)}{P(X = x)} = \frac{(2 + x)(3 + x)\theta^2 + 2\alpha(1 + \theta)^2}{(1 + \theta)((1 + x)(2 + x)\theta^2 + 2\alpha(1 + \theta)^2)}, \quad (17)$$

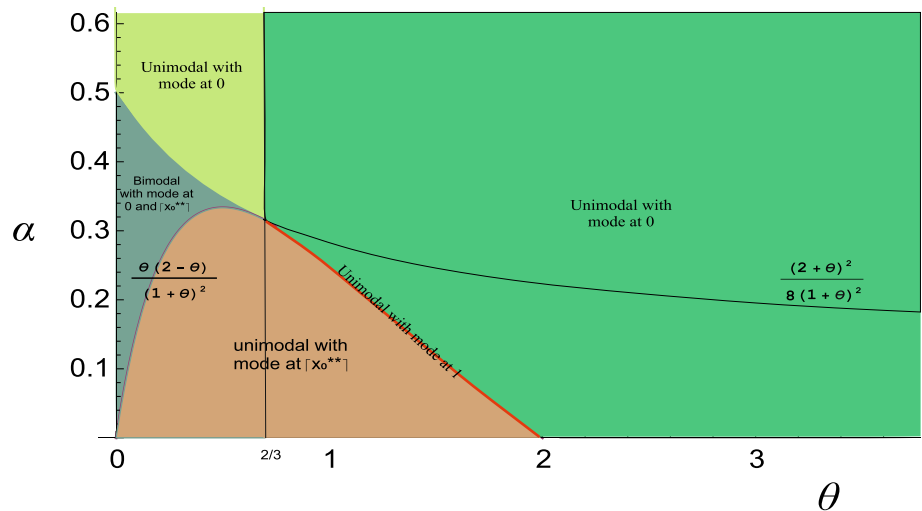
$x = 0, 1, 2, \dots$

Further, the ratio  $\frac{P(X=x+1)}{P(X=x)} < (>) 1$  implies that the pmf is decreasing (increasing). Hence, solving the equation  $\frac{P(X=x+1)}{P(X=x)} = 1$  for non-integer  $x$  (say), the roots are given as

**Table 1** The comparison of tail probabilities of the NB and PQX distributions

Tail probabilities	Models	$\mu = 1$			$\mu = 0.5$		
		$\theta = 1.5$	$\theta = 2$	$\theta = 2.5$	$\theta = 3$	$\theta = 4$	$\theta = 5$
$Pr(X > 5)$	NB	0.00973	0.00686	0.00524	0.00018	0.00012	0.00009
	PQX	0.01552	0.01052	0.00666	0.00124	0.00065	0.00034
$Pr(X > 10)$	NB	0.00013	$4.704 \times 10^{-5}$	$2.117 \times 10^{-5}$	$2.979 \times 10^{-8}$	$8.406 \times 10^{-9}$	$3.379 \times 10^{-9}$
	PQX	0.00036	0.00011	$3.329 \times 10^{-5}$	$2.943 \times 10^{-6}$	$5.431 \times 10^{-7}$	$1.164 \times 10^{-7}$
$Pr(X > 15)$	NB	$1.566 \times 10^{-6}$	$2.710 \times 10^{-7}$	$6.643 \times 10^{-8}$	$3.449 \times 10^{-12}$	$3.757 \times 10^{-13}$	$7.394 \times 10^{-14}$
	PQX	$6.717 \times 10^{-6}$	$8.492 \times 10^{-7}$	$1.215 \times 10^{-7}$	$5.384 \times 10^{-9}$	$3.337 \times 10^{-10}$	$2.901 \times 10^{-11}$
$Pr(X > 20)$	NB	$1.819 \times 10^{-8}$	$1.434 \times 10^{-9}$	$1.838 \times 10^{-10}$	$4.441 \times 10^{-16}$	$2.220 \times 10^{-16}$	$2.220 \times 10^{-16}$
	PQX	$1.097 \times 10^{-7}$	$5.672 \times 10^{-9}$	$3.778 \times 10^{-10}$	$8.508 \times 10^{-12}$	$1.747 \times 10^{-13}$	$6.106 \times 10^{-15}$

**Fig. 4** The modality of the PQX distribution



$$x_0^* = \frac{2-3\theta}{2\theta} - \frac{\theta+2}{2\theta} \sqrt{1-8\alpha\left(\frac{1+\theta}{2+\theta}\right)^2} \quad \text{and} \quad (18)$$

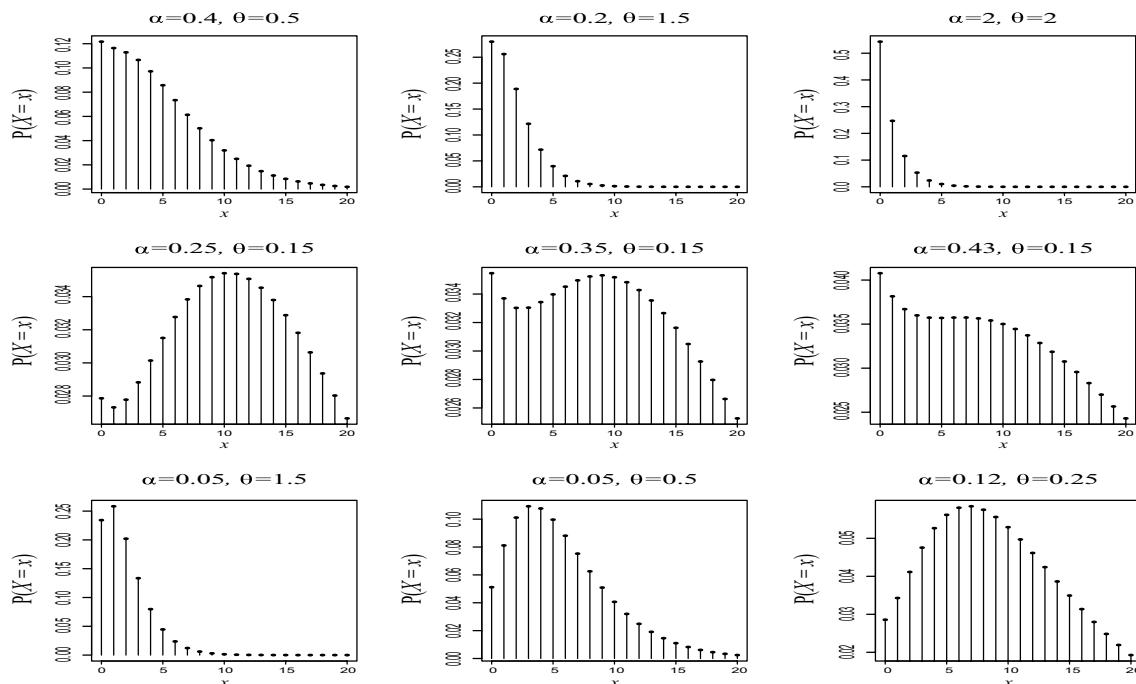
$$x_0^{**} = \frac{2-3\theta}{2\theta} + \frac{\theta+2}{2\theta} \sqrt{1-8\alpha\left(\frac{1+\theta}{2+\theta}\right)^2}$$

Hence, it can be easily verified as that

- i. If  $\left(0 < \theta < \frac{2}{3} \cap \alpha > \frac{(2+\theta)^2}{8(1+\theta)^2}\right) \cup \left(\frac{2}{3} \leq \theta < 2 \cap \alpha > \frac{\theta(2-\theta)}{(1+\theta)^2}\right) \cup (\theta \geq 2 \cap \alpha > 0)$ , the ratio is less than 1; hence, the  $PQX(\alpha, \theta)$  is unimodal with mode at 0.

- ii. If  $\frac{2}{3} \leq \theta < 2 \cap \alpha = \frac{\theta(2-\theta)}{(1+\theta)^2}$ , then mode of  $PQX(\alpha, \theta)$  is at 1.
- iii. If  $0 < \theta < 2 \cap 0 < \alpha < \frac{\theta(2-\theta)}{(1+\theta)^2}$ , then the  $PQX(\alpha, \theta)$  is unimodal and the mode is at  $[x_0^{**}]$ , where  $[.]$  denote the ceiling function.
- iv. If  $\left(0 < \theta < \frac{2}{3} \cap \frac{\theta(2-\theta)}{(1+\theta)^2} < \alpha \leq \frac{(2+\theta)^2}{8(1+\theta)^2}\right)$ ,  $PQX(\alpha, \theta)$  is bimodal with the mode at 0 and  $[x_0^{**}]$ .

The results given above are graphically summarized in Fig. 4 which displays the modality regions of the PQX



**Fig. 5** Probability mass function of  $PQX(\alpha, \theta)$  distribution for different combinations of the parameter values



distribution with respect to the values of the parameters  $\alpha$  and  $\theta$ .

In Fig. 5, we depict different shapes of  $PQX(\alpha, \theta)$  distribution for different parameters values. As seen from these figures, PQX distribution has bimodal and unimodal shapes with right skewness and symmetric. Moreover, for large value for  $\alpha$  and  $\theta$  the PQX distribution can be used to zero inflated count data sets.

### 3 Estimation

#### 3.1 Maximum likelihood

Let  $x_1, x_2, \dots, x_n$  be a random observations from the PQX distribution. The log-likelihood function is

$$\begin{aligned} \ell(\alpha, \theta) = \sum_{i=1}^n \log(2\alpha\theta(\theta + 1)^2 + \theta^3(x_i + 1)(x_i + 2)) \\ - n \log(2(\alpha + 1)) - \log(\theta + 1) \sum_{i=1}^n (x_i + 3) \end{aligned} \tag{19}$$

The partial derivatives of (19) with respect to  $\alpha$  and  $\theta$ , the following score vectors are obtained.

$$\begin{aligned} \frac{\partial \ell}{\partial \theta} = \sum_{i=1}^n \frac{4\alpha(\theta + 1)\theta + 2\alpha(\theta + 1)^2 + 3\theta^2(x_i + 1)(x_i + 2)}{2\alpha\theta(\theta + 1)^2 + \theta^3(x_i + 1)(x_i + 2)} \\ + \sum_{i=1}^n \frac{x_i + 3}{\theta + 1}, \end{aligned} \tag{20}$$

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^n \frac{2\theta(\theta + 1)^2}{2\alpha\theta(\theta + 1)^2 + \theta^3(x_i + 1)(x_i + 2)} - \frac{n}{\alpha + 1}. \tag{21}$$

The simultaneous solution of (20) and (21) gives the maximum likelihood estimates (MLEs) of  $\alpha$  and  $\theta$ . However, since the likelihood equations contain nonlinear functions, there is no explicit form for the MLEs of the parameters of PQX distribution. Therefore, the log-likelihood function in (19) should be maximized by using statistical software such as R, S-PLUS, Mathematica or MATLAB. Here, **nlm** function of R software is used to minimize the minus of log-likelihood function which is equivalent to maximization of log-likelihood. The inverse of observed information matrix evaluated at MLEs of the parameters is used to obtain corresponding standard errors. To do this, **fdHess** function of R software is used. The asymptotic confidence intervals of the parameters are obtained by

$$\hat{\alpha} \pm z_{p/2} \sqrt{\text{Var}(\hat{\alpha})}, \quad \hat{\theta} \pm z_{p/2} \sqrt{\text{Var}(\hat{\theta})},$$

where  $z_{p/2}$  is the upper  $p/2$  quantile of the standard normal distribution.

#### 3.2 Method of moments

The method of moments estimators of the parameters  $\alpha$  and  $\theta$  can be obtained by equating theoretical moments of PQX distribution to sample moments, given as follows

$$\mu'_1 = \frac{(\alpha + 3)}{\theta(\alpha + 1)} = m_1, \tag{22}$$

$$\mu'_2 = \frac{\alpha\theta + 2\alpha + 3\theta + 12}{\theta^2(\alpha + 1)} = m_2, \tag{23}$$

where  $m_1$  and  $m_2$  are the first and second sample moments, respectively. Simultaneous solution of (22) and (23) yields to following results

$$\begin{aligned} \hat{\theta}_{MM} = \frac{\hat{\alpha}_{MM} + 3}{m_1(1 + \hat{\alpha}_{MM})}, \\ \hat{\alpha}_{MM} = \frac{-7m_1^2 + \sqrt{25m_1^4 + 12m_1^3 - 12m_1^2m_2} - 3(m_1 - m_2)}{2m_1^2 + m_1 - m_2}. \end{aligned} \tag{24}$$

**Proposition 4** For fixed  $\alpha$ , the MM estimator  $\hat{\theta}$  of  $\theta$  is positively biased.

**Proof** Let  $\hat{\theta}_{MM} = g(\bar{x})$  where  $g(t) = (\alpha + 3)/(t(1 + \alpha))$  for  $t > 0$ . Taking the second partial derivative of  $g(t)$ , we have

$$g''(t) = \frac{2(\alpha + 3)}{(\alpha + 1)t^3} > 0 \tag{25}$$

which ensures that the  $g(t)$  is strictly convex. Using the Jensen's inequality,

$$E(g(\bar{X})) > g(E(\bar{X})) \tag{26}$$

So, since  $g(E(\bar{X})) = g(\mu) = g[2(\alpha + 3)/(\theta(2\alpha + 2))] = \theta$ , we obtain  $E(\hat{\theta}_{MM}) > \theta$ .  $\square$

**Proposition 5** For fixed  $\alpha$ , the MM estimator  $\hat{\theta}$  of  $\theta$  is consistent and asymptotically normal

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, v^2(\theta)) \tag{27}$$

where

$$v^2(\theta) = \frac{\theta^2[\alpha(8 + \alpha) + (\alpha + 1)(\alpha + 3)\theta + 3]}{(\alpha + 3)^2}. \tag{28}$$

**Proof** Since  $g'(\mu)$  exists and is nonzero valued, by the delta method, we have

$$\sqrt{n}(g(\bar{x}) - g(\mu)) \xrightarrow{d} N(0, [g'(\mu)]^2 \sigma^2) \tag{29}$$



where  $g(\bar{x}) = \hat{\theta}$ ,  $g(\mu) = \theta$  and

$$g'(\mu) = -\frac{(\alpha + 1)\theta^2}{\alpha + 3} \quad (31)$$

The proof is completed.  $\square$

#### 4 EM algorithm for PQX( $\alpha, \theta$ ) distribution

In this section, we deal in the estimation of parameter  $\theta$  and  $\alpha$  of PQX distribution by another estimation method known as Expectation–Maximization (EM) (see Dempster et al., [8]). The EM algorithm consists of two steps: the E-step and the M-step. E-Step computes the expectation of the unobservable part given the current values of the parameters and M-step maximizes the complete data likelihood and updates the parameters using the conditional expectations obtained in E-step. This procedure can be useful when there are no closed-form expressions for estimating the parameters and the derivatives of the likelihood are complicated.

To start with, a hypothetical complete-data distribution is defined with joint probability function

$$g(X, \lambda; \theta, \alpha) = \frac{\theta \lambda^x e^{-(1+\theta)\lambda} (2\alpha + \theta^2 \lambda^2)}{2(1 + \alpha)x!}, \quad \theta > 0, \alpha > 0. \quad (32)$$

It is straightforward to verify that the E-step of an EM cycle requires the computation of the conditional expectations of  $\left(\frac{\lambda^2}{2\alpha + \lambda^2 \theta^2} | x_i; \theta^{(h)}, \alpha^{(h)}\right)$ ,  $\left(\frac{1}{2\alpha + \lambda^2 \theta^2} | x_i; \theta^{(h)}, \alpha^{(h)}\right)$  and  $(\lambda | x_i; \theta^{(h)}, \alpha^{(h)})$ , say  $t_i^{(h)}$ ,  $s_i^{(h)}$  and  $u_i^{(h)}$ , respectively, where  $(\theta^{(h)}, \alpha^{(h)})$  is the current value of  $(\theta, \alpha)$ . Using,

$$g(\lambda | x; \theta, \alpha) = \frac{e^{-(1+\theta)\lambda} (1 + \theta)^{x+3} \lambda^x (2\alpha + \theta^2 \lambda^2)}{x! (2\alpha(1 + \theta)^2 + \theta^2(x + 1)(x + 2))}$$

$$t_i^{(h)} = \mathbb{E}\left(\frac{\lambda^2}{2\alpha + \lambda^2 \theta^2} | x_i; \theta^{(h)}, \alpha^{(h)}\right) = \frac{(x_i + 1)(x_i + 2)}{2\alpha^{(h)}(1 + \theta^{(h)})^2 + (\theta^{(h)})^2(x_i + 1)(x_i + 2)},$$

$$s_i^{(h)} = \mathbb{E}\left(\frac{1}{2\alpha + \lambda^2 \theta^2} | x_i; \theta^{(h)}, \alpha^{(h)}\right) = \frac{(1 + \theta_i^{(h)})^2}{2\alpha^{(h)}(1 + \theta^{(h)})^2 + (\theta^{(h)})^2(x_i + 1)(x_i + 2)},$$

and

$$u_i^{(h)} = \mathbb{E}(\lambda | x_i; \theta^{(h)}, \alpha^{(h)}) = \frac{(x_i + 1)(2\alpha^{(h)}(1 + \theta^{(h)})^2 + (\theta^{(h)})^2(x_i + 2)(x_i + 3))}{(1 + \theta^{(h)})(2\alpha^{(h)}(1 + \theta^{(h)})^2 + (\theta^{(h)})^2(x_i + 1)(x_i + 2))}.$$

The EM cycle completes with the M-step, involving complete data maximum likelihood over  $(\theta, \alpha)$ , with the missing  $\lambda$ 's replaced by their conditional expectations

$$\theta^{(h+1)} = \frac{\bar{u}^{(h)} + \sqrt{((\bar{u}^{(h)})^2 - 8(\bar{t}^{(h)})^2)}}{4\bar{t}^{(h)}},$$

$$\alpha^{(h+1)} = \frac{1}{2\bar{s}^{(h)} - 1}.$$

#### 5 Simulation

The finite sample performance of MM and MLE methods is compared via simulation study. The below simulation procedure is implemented for this purpose.

1. Determine the sample size  $n$  and the parameter values of PQX distribution,  $\Theta = (\alpha, \theta)$ ,
2. Generate random observations from PQX distribution for given  $n$  and parameter vector,
3. Using the random sample in step 2, estimate  $\Theta$  with MLE and MM methods,
4. Repeat the steps 2 and 3 based on the replication number,  $N$ ,
5. Using the estimated parameter vector,  $\hat{\Theta}$ , and true parameter vector,  $\Theta$ , calculate the biases, mean relative estimates (MREs) and mean square errors (MSEs) by using,

$$\text{Bias} = \sum_{j=1}^N \frac{\hat{\theta}_{ij} - \theta_i}{N}, \quad \text{MRE} = \sum_{j=1}^N \frac{\hat{\theta}_{ij}/\theta_i}{N}, \quad (33)$$

$$\text{MSE} = \sum_{j=1}^N \frac{(\hat{\theta}_{ij} - \theta_i)^2}{N}, \quad i = 1, 2.$$

The simulation is carried out with statistical software R. We generate  $n = 50, 55, 60, \dots, 300$  sample of size from PQX distribution. The simulation replication number is  $N = 1000$ . The true parameter vector is  $\Theta = (\alpha = 0.5, \theta = 1.5)$ . Figure 6 displays the simulation results. We expect that the estimated biases and MSEs should be near the zero for large  $n$  values. The results verify the expectation. The biases and MSEs approach their nominal value, zero, for  $n \rightarrow \infty$ . Additionally, the MREs are near the one. The MLE and MM methods behave very similar for the parameter  $\alpha$ . However, the MLE method approaches the nominal values of biases, MSEs, and MREs more faster than MM method for the parameter  $\theta$ . Therefore, we suggest practitioners to use MLE method for estimating the parameters of PQX distribution.

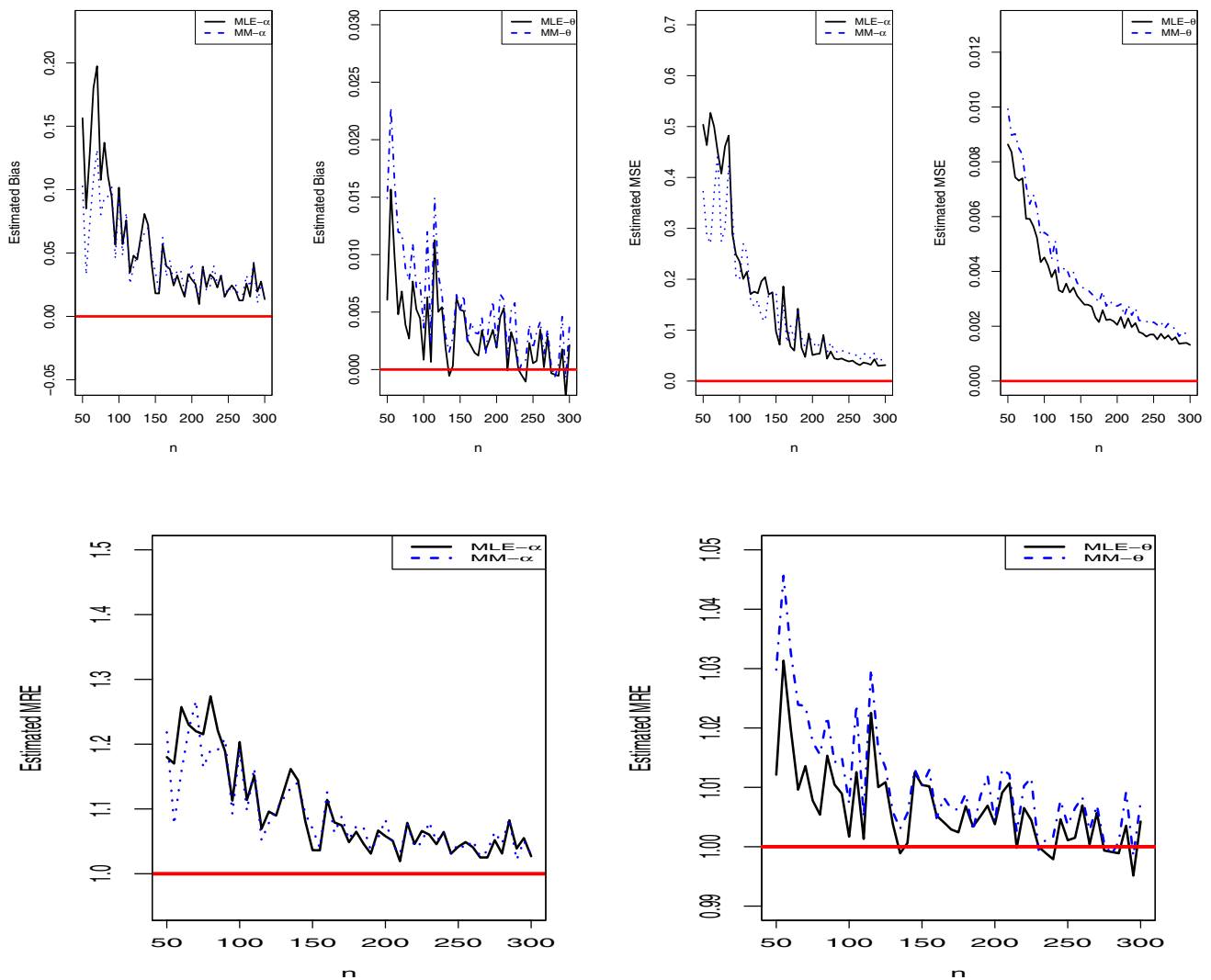


Fig. 6 Estimated biases, MSEs, and MREs of the parameters of PQX distribution based on the MLE and MM estimation methods

### 6 INAR(1) model with PQX innovations

Let the innovations  $(\varepsilon_t)_{\mathbb{N}}$  be an independent and identically distributed (iid) process with  $E(\varepsilon_t) = \mu_\varepsilon$  and  $Var(\varepsilon_t) = \sigma_\varepsilon^2$ . The process,  $(X_t)_{\mathbb{N}}$ , is called as INAR(1) if it follows a recursion

$$X_t = p \circ X_{t-1} + \varepsilon_t, \tag{34}$$

where  $0 \leq p < 1$ . The symbol,  $\circ$ , represents a binomial thinning operator and is defined as

$$p \circ X_{t-1} := \sum_{j=1}^{X_{t-1}} Z_j, \tag{35}$$

where  $\{Z_j\}_{j \geq 1}$  is referred to as a counting series and be a sequence of iid Bernoulli rvs with  $Pr(Z_j = 1) = p$ . The INAR(1) process is stationary for  $0 \leq p < 1$ . The process is non-stationary for the case  $p = 1$ . The INAR(1) process is a homogeneous Markov chain with the one-step transition probabilities given by McKenzie [17], Al-Osh and Alzaid [2]

$$Pr(X_t = k | X_{t-1} = l) = \sum_{i=1}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} Pr(\varepsilon_t = k-i), k, l \geq 0. \tag{36}$$

The mean, variance, and dispersion index of INAR(1) process are given, respectively, by Weiß [28]

$$E(X_t) = \frac{\mu_\varepsilon}{1-p}, \quad (37)$$

$$\text{Var}(X_t) = \frac{p\mu_\varepsilon + \sigma_\varepsilon^2}{1-p^2}, \quad (38)$$

$$DI_X = \frac{DI_\varepsilon + p}{1+p}, \quad (39)$$

where  $\mu_\varepsilon$ ,  $\sigma_\varepsilon^2$  and  $DI_\varepsilon$  are the mean, variance, and dispersion index of innovation distribution, respectively. In real-life problems, the empirical dispersion is greater than one (over-dispersion) in general. For instance, monthly counts of passengers, yearly number of destructive earthquakes, yearly number of goals of a and among others. To model these kind of datasets, the innovation distribution of INAR(1) process should be able to model over-dispersion. Here, we propose a new INAR(1) process with PQX innovations to model such data sets. Let  $\{\varepsilon_t\}_N$  follows a PQX distribution given in (3). The one-step transition probability of INARPQX(1) model is given by

$$\Pr(X_t = k | X_{t-1} = l) = \sum_{i=1}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} \frac{2\alpha\theta(\theta+1)^2 + \theta^3(k-i+1)(k-i+2)}{2(\alpha+1)(\theta+1)^{k-i+3}} \quad (40)$$

From now on, this process is called as INARPQX(1) process. Using the properties of INAR(1) process, the mean, variance, and dispersion index of INARPQX(1) process are given, respectively, by

$$\mu_X = \frac{(\alpha+3)}{\theta(\alpha+1)(1-p)}, \quad (41)$$

$$\sigma_X^2 = \frac{\alpha^2(\theta + \theta p + 1) + 4\alpha(\theta + \theta p + 2) + 3(\theta + \theta p + 1)}{\theta^2(\alpha+1)^2(1-p^2)}, \quad (42)$$

$$DI_X = 1 + \frac{\alpha(\alpha+8)+3}{\theta(\alpha+1)(\alpha+3)(1+p)}. \quad (43)$$

According to the results given in Al-Osh and Alzaid [2], the conditional expectation and variance of INARPQX(1) process are given, respectively, by

$$E(X_t | X_{t-1}) = pX_{t-1} + \frac{2(\alpha+3)}{\theta(2\alpha+2)} \quad (44)$$

$$\text{Var}(X_t | X_{t-1}) = p(1-p)X_{t-1} + \frac{\alpha^2 + (\alpha+1)(\alpha+3)\theta + 8\alpha + 3}{(\alpha+1)^2\theta^2} \quad (45)$$

## 6.1 Estimation

The estimation procedure of INAR(1) process is discussed with three estimation methods. These are conditional maximum likelihood (CML), Yule–Walker (YL), and conditional least squares (CLS). The finite sample performance of these estimation methods is compared via extensive simulation study. The rest of this section is devoted to theoretical background of the used estimation methods.

### 6.1.1 Conditional maximum likelihood

Let  $X_1, X_2, \dots, X_T$  be a random sample from the stationary process, INARPQX(1). The conditional log-likelihood function of INARPQX(1) is

$$\begin{aligned} \ell(\Theta) &= \sum_{t=2}^T \ln [\Pr(X_t = k | X_{t-1} = l)] \\ &= \sum_{t=2}^T \ln \left[ \sum_{i=1}^{\min(X_t, X_{t-1})} \binom{X_{t-1}}{i} p^i (1-p)^{X_{t-1}-i} \frac{2\alpha\theta(\theta+1)^2 + \theta^3(X_t - i + 1)(X_t - i + 2)}{2(\alpha+1)(\theta+1)^{X_t-i+3}} \right] \end{aligned} \quad (46)$$

The CML estimators of  $(p, \alpha, \theta)$  can be obtained by direct maximization of log-likelihood function given in (46). The **nlm** function of **R** software is used to minimize the minus of log-likelihood function. The inverse of observed information matrix is used to obtain corresponding standard errors of the CML estimation of the INARPQX(1) process. The observed information matrix is obtained by **fdHess** function of **R** software.

### 6.1.2 Yule–Walker

The YW estimators of INARPQX(1) process are obtained by equating the sample moments to the theoretical moments of the process. Since the autocorrelation function (ACF) of INAR(1) process at lag  $h$  is  $\rho_X(h) = p^h$ , the YW estimator of  $p$  is given by

$$\hat{\rho}_{YW} = \frac{\sum_{t=2}^T (X_t - \bar{X})(X_{t-1} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} \tag{47}$$

The YW estimators of  $\alpha$  and  $\theta$  are obtained by equating the sample mean and sample dispersion to theoretical mean and theoretical dispersion of the process. The YW estimator of  $\theta$  is given by

$$\hat{\theta}_{YW} = \frac{\hat{\alpha}_{YW} + 3}{\bar{X}(1 + \hat{\alpha}_{YW})(1 - \hat{\rho}_{YW})} \tag{48}$$

where  $\bar{X} = \sum_{t=1}^T X_t / N$ . Substituting  $\theta$  with (48) in (43) and equating (43) to sample dispersion,  $\hat{Dl}_X$ , we have

$$\hat{\alpha}_{YW} = \frac{\left( \begin{array}{l} -3\hat{Dl}_X + 4\bar{X} + \hat{\rho}_{YW}(-3\hat{Dl}_X - 4\bar{X} + 3) \\ -\sqrt{\bar{X}(\hat{\rho}_{YW} - 1)}(12\hat{Dl}_X - 13\bar{X} + \hat{\rho}_{YW}(12\hat{Dl}_X + 13\bar{X} - 12) - 12) + 3 \end{array} \right)}{\hat{Dl}_X - \bar{X} + \hat{\rho}_{YW}(\hat{Dl}_X + \bar{X} - 1) - 1} \tag{49}$$

### 6.2 Model accuracy

The standardized Pearson residual is used to check the fitted model accuracy. The standardized Pearson residuals are given by

$$e_t = \frac{x_t - E(X_t|X_{t-1})}{\sqrt{\text{Var}(X_t|X_{t-1})}}, t = 2, 3, \dots, T, \tag{50}$$

where  $E(X_t|X_{t-1})$  and  $\text{Var}(X_t|X_{t-1})$  are given in (44) and (45), respectively. When the model is adequate for fitted data set, the standardized Pearson residuals should be uncorrelated with zero mean and unit variance. If the variance of standardized Pearson residuals is higher/lower than one, it

**Table 2** Simulation results of INARPQX(1) process for CML and YW estimation methods

Sample size		$\rho = 0.3$		$\alpha = 0.5$		$\theta = 0.5$	
		CML	YW	CML	YW	CML	YW
100	Bias	-0.0170	-0.0294	0.2888	0.0304	-0.0027	-0.0090
	MSE	0.0033	0.0085	0.4373	0.4743	0.0067	0.0075
	MRE	0.9661	0.9412	1.0578	1.0608	0.9945	0.9819
300	Bias	-0.0070	-0.0118	0.0093	0.0636	0.0014	-0.0037
	MSE	0.0012	0.0029	0.0711	0.2281	0.0016	0.0020
	MRE	0.9861	0.9764	1.0187	1.1273	1.0028	0.9926
500	Bias	-0.0020	-0.0067	0.0128	0.0259	0.0016	0.0008
	MSE	0.0006	0.0015	0.0424	0.1243	0.0011	0.0014
	MRE	0.9960	0.9866	1.0256	1.0518	1.0032	1.0017
		$\rho = 0.3$		$\alpha = 0.5$		$\theta = 2$	
		CML	YW	CML	YW	CML	YW
100	Bias	-0.0126	-0.0191	-0.0461	0.1417	0.1381	0.0375
	MSE	0.0061	0.0097	0.1575	0.7691	0.2210	0.1819
	MRE	0.9581	0.9364	0.9079	1.2834	1.0691	1.0188
300	Bias	-0.0063	-0.0079	-0.0195	0.0977	0.0563	0.0229
	MSE	0.0020	0.0034	0.1093	0.4875	0.0831	0.0926
	MRE	0.9789	0.9735	0.9611	1.1955	1.0282	1.0115
500	Bias	-0.0022	-0.0025	0.0005	0.0651	0.0357	0.0257
	MSE	0.0014	0.0021	0.0794	0.2553	0.0475	0.0634
	MRE	0.9928	0.9918	1.0009	1.1303	1.0179	1.0129

shows that there is more or less dispersion considered by fitted model (see, Harvey and Fernandes, [11]).

### 6.3 Simulation

We compare the finite sample performance of CML and YW estimators of the parameters of INARPQX(1) process with a brief simulation study. We generate  $n = 100, 300$  and  $500$  sample of sizes from INAR(1) process with PQX innovations. The simulation replication number is  $N = 1,000$ . The two parameter vectors are used:  $(p = 0.3, \alpha = 0.5, \theta = 0.5)$ ,  $(p = 0.3, \alpha = 0.5, \theta = 2)$ . The simulation is carried out with **R** software. The results are interpreted based on the estimated biases, MSEs and MREs. The required formulas of these measures are given in Sect. 4. We expect to see that

the estimated biases and MSEs approach to zero for large values of  $n$ . Beside this, the estimated MREs should be near the one. The simulation results are given in Table 2. Based on the results given in this table, it is clear that the estimated biases and MSEs are very near the their desired value, zero. Moreover, all estimated MREs are near the one. However, the CML estimators approach to the nominal values of MSEs and MREs more faster than those of YW estimators. Therefore, we suggest practitioners to use CML estimation method to obtain the unknown parameters of INARPQX(1) process. Actually, the both estimation methods work well for large sample sizes. So, if the number of time series of counts are large enough, YW estimation method is also preferable.

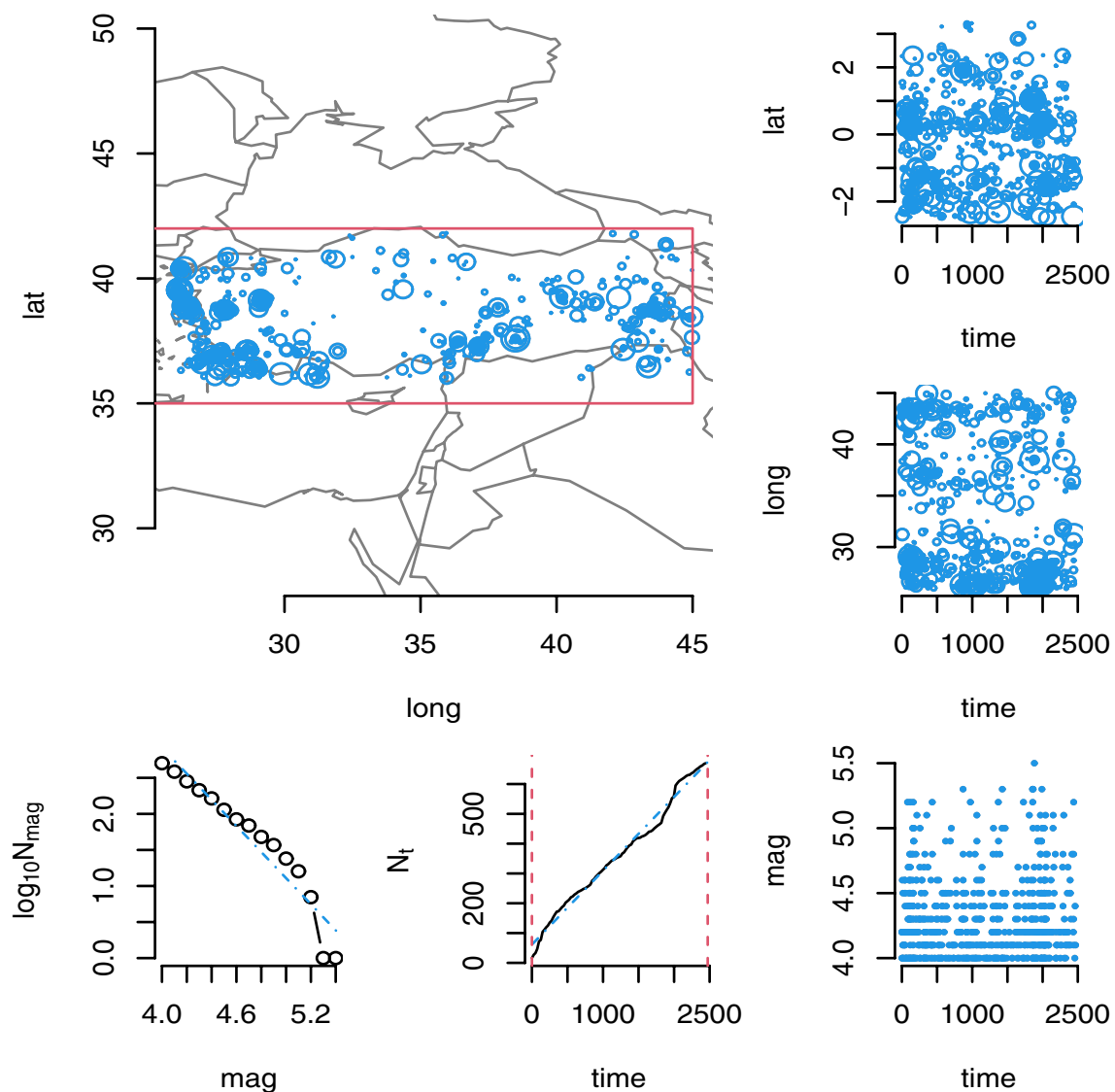


Fig. 7 The plots of the earthquake catalog of Turkey

## 7 Empirical study

In this section, we illustrate the importance of the INAR-PQX(1) by an application on the earthquake data of Turkey. Firstly, we describe the created earthquake catalog and its properties. In the second step, INAR(1) processes defined under different innovation distributions are used to model the monthly counts of the earthquake.

### 7.1 Study area and data

The earthquake data of the Turkey is obtained from the Disaster & Emergency Management Authority Presidential of Earthquake Department, Turkey. The data are available from <https://depem.afad.gov.tr/?lang=en>. The used data contain the earthquakes with magnitude 4 and above occurred in Turkey between the dates 6 January 2012 and 14 October 2018. Firstly, the earthquake catalog of the Turkey is created by using the **ETAS** package of the **R** software. The detail on this package can be found in Jalilian [12]. The earthquake catalog is displayed in Fig. 7. The top-left figure shows the spatial distribution of the earthquakes under the study area. The three figures in the right part of Fig. 7 show the changes of the latitude, longitude, and magnitude of the earthquakes over the time. The two figures in the bottom-right part of Fig. 7 show the completeness and time stationary of the earthquake catalog. Here,  $N_m$  represents the number of earthquakes with magnitude  $\geq m$ . If the plot of  $\log(N_m)$  versus  $m$  shows linear trend, it represents the completeness of the earthquake catalog. Besides, the stationary of the earthquake catalog is evaluated based on the plot of  $N_t$  versus  $t$ . Here,

$N_t$  represents the number of earthquakes up to time  $t$ . If the plotted points of  $N_t$  versus  $t$  have a functional form such as  $N_t = \lambda_0 t$  where  $\lambda > 0$ , it is an evidence that the time series of the earthquake is stationary. Therefore, we conclude that the used earthquake catalog is stationary and complete.

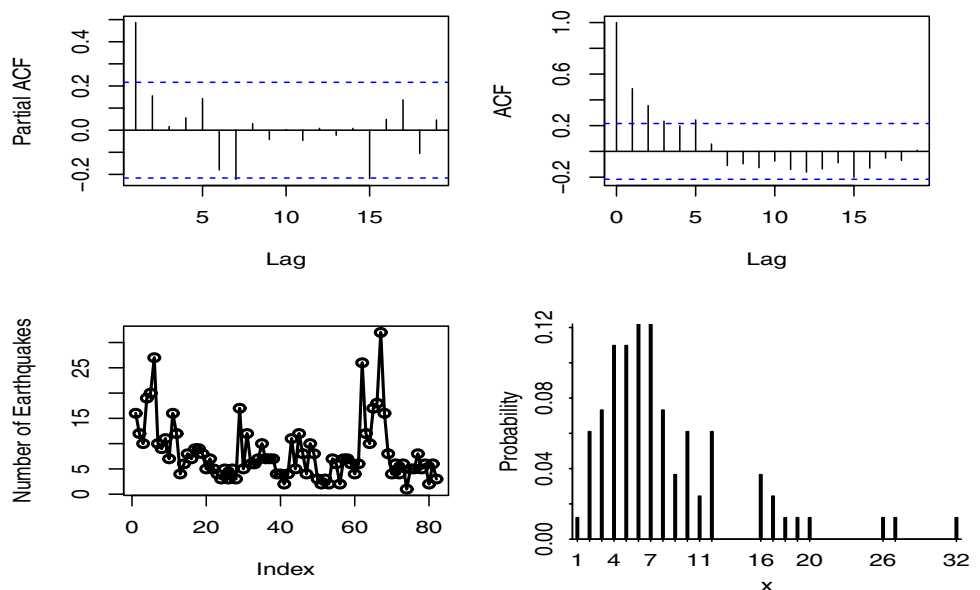
### 7.2 Modeling of the number of earthquakes

The importance of INAR-PQX(1) model is demonstrated with an application to the monthly counts of earthquakes in Turkey. The proposed model, INAR-PQX(1), is compared with well-known INAR processes, INAR(1), INARPL(1), INARG(1), and INAR(1) with NB innovations shortly INARNB(1). The one-step transition probabilities of the used models are given, respectively, by

$$\Pr(X_t = k | X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} \frac{\exp(-\lambda) \lambda^{k-i}}{(k-i)!}, \lambda > 0, \tag{51}$$

$$\Pr(X_t = k | X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} \frac{\theta^2 (k-i+\theta+2)}{(\theta+1)^{k-i+3}}, \theta > 0, \tag{52}$$

**Fig. 8** The plots of time series, autocorrelation, and partial autocorrelation functions for the monthly counts of earthquakes in Turkey



**Table 3** The estimated parameters of the fitted INAR(1) models and corresponding goodness-of-fit statistics

Model	Parameters	Estimate	SE	$-\ell$	AIC	BIC	$\mu_X$	$\sigma_X^2$	$DI_X$
INARPQX(1)	$p$	0.461	0.042	221.300	448.599	455.819	7.948	31.723	3.991
	$\alpha$	94.964	4.342						
	$\theta$	0.238	0.076						
INARNB(1)	$p$	0.427	0.058	223.887	453.774	460.994	7.966	26.711	3.353
	$\phi$	0.229	0.053						
	$n$	1.358	0.460						
INARG(1)	$p$	0.460	0.041	224.299	452.599	457.412	7.948	42.084	5.295
	$\theta$	0.189	0.022						
INARPL(1)	$p$	0.418	0.046	224.377	452.755	457.568	7.969	24.755	3.106
	$\theta$	0.373	0.043						
INARP(1)	$p$	0.307	0.047	266.499	536.998	541.812	8.015	8.015	1
	$\lambda$	5.553	0.441						
Empirical							8.183	33.978	4.152

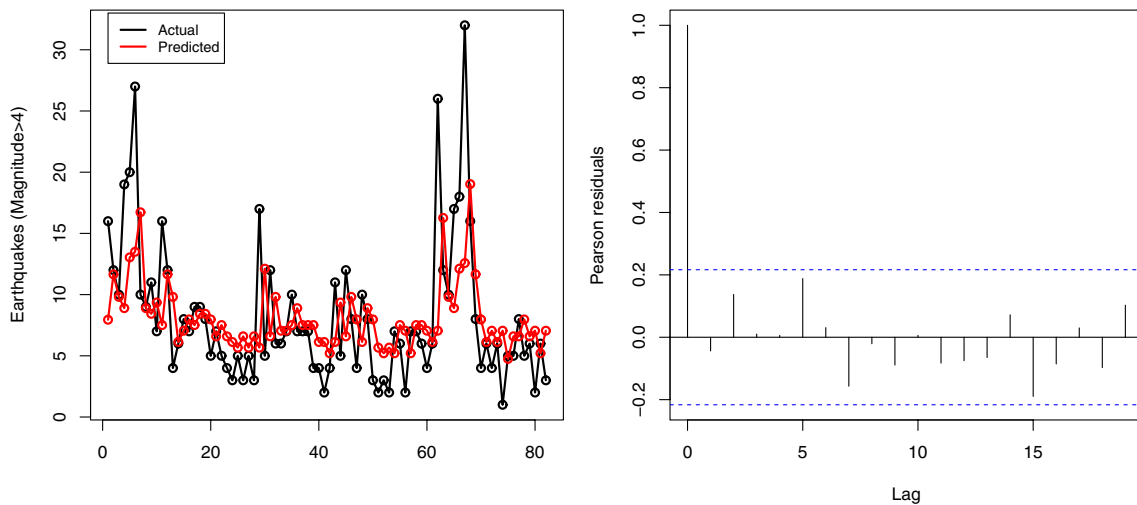
$$\Pr(X_t = k | X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} \theta (1-\theta)^{k-i}, \quad 0 < \theta < 1, \tag{53}$$

$$\Pr(X_t = k | X_{t-1} = l) = \sum_{i=0}^{\min(k,l)} \binom{l}{i} p^i (1-p)^{l-i} \binom{n+k-i-1}{k-i} (1-\pi)^{k-i} \pi^n, \quad n > 0, 0 < \pi < 1. \tag{54}$$

To analyze the reported data set in the previous section, the earthquakes are grouped monthly to predict the monthly number of the earthquakes with magnitude  $\geq 4$ . We use the  $-\ell$ , Akaike Information Criteria (AIC), and Bayesian Information Criteria (BIC) statistics to select the most appropriate model for the used data set. The lowest

values of these criteria show the best fitted model. The used data sets consist of 82 monthly counts of earthquakes with magnitude greater than 4 between the date of January 2012 and March 2018. Firstly, the possible over-dispersion in the data set should be explored. For this aim, we use the hypothesis test for over-dispersion proposed by Schweer and Weiß [23]. The test statistic value of the over-dispersion hypothesis is 24.187, and its corresponding  $p$  value is  $< 0.001$  which indicates that the used data set displays over-dispersion. Therefore, the innovation distribution of the INAR(1) process should be able to capture the over-dispersion in the data set.

The some useful plots of the used data set displayed in Fig. 8. As seen from autocorrelation function (ACF) and partial ACF (PACF) plots, there is a clear cut of at first lag which ensures that the AR(1) process can be used to model this data set.



**Fig. 9** The predicted values of the monthly counts of earthquakes in Turkey



The estimated parameters of the fitted INAR(1) processes and corresponding AIC, BIC values are listed in Table 3. As seen from the results reported in Table 3, the proposed INARPQX(1) process has the lowest values of AIC and BIC values which indicates the proposed model performs better than INARP(1), INARG(1), INARPL(1), and IARNB(1) models. Moreover, we analyze the departure from error distribution by means of residual analysis. For this aim, we use the Pearson residuals. The Pearson residuals can be calculated by using

$$r_t = \frac{X_t - E(X_t|X_{t-1})}{\text{Var}(X_t|X_{t-1})^{1/2}} \quad (55)$$

where  $E(X_t|X_{t-1})$  and  $\text{Var}(X_t|X_{t-1})$  are given in (44) and (45), respectively. When the fitted model is correct, the mean and variance of the Pearson residuals should be near the zero and one, respectively. Also, there should be no autocorrelation problem for the estimated Pearson residuals. The mean and variance of the Pearson residuals are calculated 0.0009 and 0.9754, respectively, which are very near the desired values. Additionally, the ACF plot of the Pearson residuals is displayed in right side of Fig. 9 which indicates that there is no autocorrelation problem for the Pearson residuals. The fitted INARPTE(1) model is given by

$$X_t = 0.461 \circ X_{t-1} + \varepsilon_t.$$

where  $\varepsilon_t \sim \text{PQX}(94.964, 4.342)$ . The predicted values of criminal mischief in 14th police car beat can be obtained by using (56).

$$\begin{aligned} \hat{X}_1 &= \frac{2(\hat{\alpha} + 3)}{2\hat{\theta}(\hat{\alpha} + 1)(1 - \hat{\rho})} = 7.945 \\ \hat{X}_t &= \hat{\rho}X_{t-1} + \frac{2(\hat{\alpha} + 3)}{2\hat{\theta}(\hat{\alpha} + 1)} = 0.432X_{t-1} + 4.287 \end{aligned} \quad (56)$$

The predicted values of monthly counts of earthquakes are displayed in the left side of Fig. 9.

## 8 Conclusion and future work

This study introduces a new two-parameter discrete distribution, shortly PQX, for modeling the over-dispersed counts. The statistical properties of the PQX distribution are derived and estimation of the unknown parameters of the proposed distribution is discussed in detail. INARPQX(1) processes are introduced based on the PQX distribution to predict the monthly counts of the earthquakes with magnitude 4 and above in Turkey. Empirical results show that the INARPQX(1) processes perform better than INARP(1), INARP(1), INARG(1), and INARNB(1) processes for the data used. As a future work of this study, we will

try to extend the INARPQX(1) to multivariate case for joint modeling of the more than one region in predicting the monthly counts of earthquakes. We believe that the PQX distribution will increase its popularity find a wider application area in different sciences such medical, finance, and actuarial.

## Compliance with ethical standards

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Aktas S, Konsuk H, Yigiter A (2009) Estimation of change point and compound Poisson process parameters for the earthquake data in Turkey. *Environmetrics* 20(4):416–427
2. Al-Osh MA, Alzaid AA (1987) First-order integer-valued autoregressive (INAR (1)) process. *J Time Ser Anal* 8(3):261–275
3. Altun E (2020) A new generalization of geometric distribution with properties and applications. *Commun Stat Simul Comput* 49(3):793–807
4. Andersson J, Karlis D (2014) A parametric time series model with covariates for integers in  $\mathbb{Z}$ . *Stat Model* 14(2):135–156
5. Borges P, Bourguignon M, Molinares FF (2017) A generalised NGINAR (1) process with inflated-parameter geometric counting series. *Aust N Z J Stat* 59:137–150
6. Bourguignon M, Rodrigues J, Santos-Neto M (2019) Extended Poisson INAR (1) processes with equidispersion, underdispersion and overdispersion. *J Appl Stat* 46:101–118
7. Brillinger DR (1993) Earthquake risk and insurance. *Environmetrics* 4(1):1–21
8. Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Ser B (Methodol)* 39(1):1–22
9. Gospodinov D, Rotondi R (2001) Exploratory analysis of marked Poisson processes applied to Balkan earthquake sequences. *J Balkan Geophys Soc* 4(3):61–68
10. Gutiérrez E, Taucer F, De Groeve T, Al-Khudhairi DHA, Zaldivar JM (2005) Analysis of worldwide earthquake mortality using multivariate demographic and seismic data. *Am J Epidemiol* 161(12):1151–1158
11. Harvey AC, Fernandes C (1989) Time series models for count or qualitative observations. *J Bus Econ Stat* 7(4):407–417
12. Jalilian A (2019) ETAS: an R package for fitting the space-time ETAS model to earthquake data. *J Stat Softw* 88(1):1–39

13. Jazi MA, Jones G, Lai CD (2012) Integer valued AR (1) with geometric innovations. *J Iran Stat Soc* 11(2):173–190
14. Kim H, Lee S (2017) On first-order integer-valued autoregressive process with Katz family innovations. *J Stat Comput Simul* 87:546–562
15. Koshimura S, Katada T, Mofjeld HO, Kawata Y (2006) A method for estimating casualties due to the tsunami inundation flow. *Nat Hazards* 39(2):265–274
16. Lívio T, Mamode Khan N, Bourguignon M, Bakouch H (2018) An INAR (1) model with Poisson–Lindley innovations. *Econ Bull* 38:1505–1513
17. McKenzie E (1985) Some simple models for discrete variate time series 1. *JAWRA J Am Water Resour Assoc* 21:645–650
18. McKenzie E (1986) Autoregressive moving-average processes with negative binomial and geometric marginal distributions. *Adv Appl Probab* 18:679–705
19. Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. *J Am Stat Assoc* 83(401):9–27
20. Ogata Y (1998) Space-time point-process models for earthquake occurrences. *Ann Inst Stat Math* 50(2):379–402
21. Ozel G (2011) A bivariate compound Poisson model for the occurrence of foreshock and aftershock sequences in Turkey. *Environmetrics* 22(7):847–856
22. Samardjieva E, Badal J (2002) Estimation of the expected number of casualties caused by strong earthquakes. *Bull Seismol Soc Am* 92(6):2310–2322
23. Schweer S, Weiß CH (2014) Compound Poisson INAR (1) processes: stochastic properties and testing for overdispersion. *Comput Stat Data Anal* 77:267–284
24. Sen S, Chandra N (2017) The quasi xgamma distribution with application in bladder cancer data. *J Data Sci* 15:61–76
25. Sen S, Maiti SS, Chandra N (2016) The xgamma distribution: statistical properties and application. *J Mod Appl Stat Methods* 15(1):38
26. Turkan S, Ozel G (2014) Modeling destructive earthquake casualties based on a comparative study for Turkey. *Nat Hazards* 72(2):1093–1110
27. Vere-Jones D (1995) Forecasting earthquakes and earthquake risk. *Int J Forecast* 11(4):503–538
28. Weißs CH (2018) An introduction to discrete-valued time series. Wiley, New York
29. Wu XY, Gu JH, Wu HY (2009) A modified exponential model for reported casualties during earthquakes. *Acta Seismol Sin* 31(4):457–463
30. Wyss M (2005) Human losses expected in Himalayan earthquakes. *Nat Hazards* 34(3):305–314
31. Zhao YZ, Wu ZL, Li YT (2008) Casualty in earthquake and tsunami disasters: internet-based monitoring and early estimation of the final death toll. In: Proceedings of the 14th world conference on earthquake engineering, Beijing
32. Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. *J Am Stat Assoc* 97(458):369–380
33. Zhuang J, Ogata Y, Vere-Jones D (2004) Analyzing earthquake clustering features by using stochastic reconstruction. *J Geophys Res* 109(B5):B05301. <https://doi.org/10.1029/2003JB002879>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.