



# A new INAR(1) process with bounded support for counts showing equidispersion, underdispersion and overdispersion

Yao Kang<sup>1</sup> · Dehui Wang<sup>1</sup> · Kai Yang<sup>2</sup>

Received: 11 July 2018 / Revised: 6 April 2019 / Published online: 16 April 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

The present work introduces a mixture INAR(1) model based on the mixing Pogram and binomial thinning operators with a finite range  $\{0, 1, \dots, n\}$ . The new model can be used to handle equidispersion, underdispersion, overdispersion, zero-inflation and multimodality. Several probabilistic and statistical properties are explored. Estimators of the model parameters are derived by the conditional maximum likelihood method. The asymptotic properties and numerical results of the estimators are also studied. In addition, the forecasting problem is addressed. Applications to real data sets are given to show the application of the new model.

**Keywords** Binomial AR(1) processes · Pogram operator · Binomial thinning operator · Parameter estimation · Forecasting

## 1 Introduction

During the past few decades, there has been an upsurge of interest in the field of count data time series analysis. In particular, INAR(1) models have attracted the attention of a great number of researchers due to the strong application value, see Weiß (2008). The most common way to construct INAR(1) processes is based on different types of thinning operators. The binomial thinning operator “ $\circ$ ” is the most popular one which was proposed by Steutel and Van Harn (1979). It is defined as

---

✉ Dehui Wang  
wangdh@jlu.edu.cn

<sup>1</sup> School of Mathematics, Jilin University, Changchun 130012, China

<sup>2</sup> School of Mathematics and Statistics, Changchun University of Technology, Changchun 130012, China

$$\alpha \circ X = \sum_{i=1}^X B_i, \quad (1)$$

where  $\{B_i\}$  is an independent identically distributed (iid) Bernoulli( $\alpha$ ) random sequence independent of  $X$  and  $\alpha \in [0, 1]$ . Based on the binomial thinning operator, some INAR(1) processes have been proposed, see, for example, Kim and Lee (2017), Bourguignon and Vasconcellos (2015) and Jazi et al. (2012) and Bourguignon et al. (2019), and the references therein. The models cited above are suitable to model count data having an infinite range  $\{0, 1, \dots\}$ .

The use of mixture models is another approach for constructing INAR(1) processes. As pointed out by Shirozhan and Mohammadpour (2018a) and Khoo et al. (2017), the mixture models provide an appealing tool for time series modeling, having been used in modeling population heterogeneity. It is well-known that finite mixture models provide more flexibility in empirical modeling and are able to cater for multimodality in the data. For these reasons, some mixture INAR(1) models based on Peggam's operator have been proposed. The Peggam's operator was proposed by Peggam (1980) and is defined below.

**Definition 1** (*Peggam's Operator*) Consider two independent discrete random variables  $U$  and  $V$ . Peggam operator mixes  $U$  and  $V$  with the weights  $\phi$  and  $1 - \phi$  as

$$Z = (\phi, U) * (1 - \phi, V) \quad (2)$$

with the corresponding marginal probability function

$$P(Z = j) = \phi P(U = j) + (1 - \phi)P(V = j), \quad j = 0, 1, 2, \dots$$

Based on Peggam and different types of thinning operators, three new INAR(1) models have been proposed. Khoo et al. (2017) introduced a novel INAR(1) process to provide more flexibility in empirical modeling. Shirozhan and Mohammadpour (2018a, b) proposed INAR(1) models with Poisson marginal distribution and serially dependent innovation, respectively. However, the above three models focus on count data having an infinite range  $\{0, 1, \dots\}$ . The binomial AR(1) (BAR(1)) process, proposed by McKenzie (1985), is the most common way to model count data having a finite range  $\{0, 1, \dots, n\}$ . The definition of the BAR(1) model is given below.

**Definition 2** (*BAR(1) Model*) Let  $p \in (0; 1)$  and  $\rho \in (\max\{-\frac{p}{1-p}, -\frac{1-p}{p}\}; 1)$ . Defined  $\beta := p(1 - \rho)$  and  $\alpha := \beta + \rho$ . Fix  $n \in \mathbb{N}$ . A BAR(1) process  $\{X_t\}$  is defined by the recursion

$$X_t = \alpha \circ X_{t-1} + \beta \circ (n - X_{t-1}) \quad \text{with } X_0 \sim B(n, p), \quad (3)$$

where all thinnings are performed independently of each other, and where the thinnings at time  $t$  are independent of  $(X_s)_{s < t}$ .

The BAR(1) model has been widely used due to its strong application value. Some authors have studied this model during the past ten years. Cui and Lund (2010) studied

inference methods for the BAR(1) model. Weiß (2009a) investigated marginal and serial properties of jumps in the BAR(1) process. Weiß (2009b) proposed several approaches to monitor a binomial AR(1) process. Weiß and Kim (2013) presented four approaches for estimating the parameters of the BAR(1) model. Weiß (2013) considered the moments, cumulants, and estimation of the BAR(1) model. Scott et al. (2014) introduced new classes of bivariate time series models being useful to fit count data time series with a finite range of counts. Yang et al. (2018) presented a new approach for estimating the parameters of the self-exciting threshold BAR(1) model.

The binomial index of dispersion, BID, is a useful metric which is used to quantify the dispersion behavior of a count data random variable  $X$  with a finite range  $\{0, 1, \dots, n\}$ . It is defined as

$$\text{BID} = \frac{n\sigma^2}{\mu(n - \mu)},$$

where  $\mu$  and  $\sigma^2$  are the mean and variance of the random variable  $X$ , respectively. A finite range count data random variable is said to have overdispersion if  $\text{BID} > 1$ , it is equidispersed if  $\text{BID} = 1$ , and it is underdispersed if  $\text{BID} < 1$ . The BID of binomial distribution equals one, which leads to the result that the BAR(1) model can not explain underdispersion and overdispersion. To solve this problem, some extended BAR(1) models were proposed. Weiß and Pollett (2014) considered a class of density-dependent BAR(1) process  $n_t$  with range  $\{0, 1, \dots, n\}$ , where the thinning probabilities were not constant in time but rather depend on the current density  $n_t/n$ . Kim and Weiß (2015) considered the modeling of count data time series with a finite range having extra-binomial variation. Möller et al. (2016) proposed types of self-exciting threshold BAR(1) models for integer-valued time series with a finite range, which are based on the BAR(1) model. Möller et al. (2018) developed four extensions of the BAR(1) processes, which can accommodate a broad variety of zero patterns.

To better model count data with bounded support, in this paper, we introduce a mixture INAR(1) model with bounded support based on the mixing Pogram and binomial thinning operators. One main advantage of the mixture model is that multimodality can be well described. Furthermore, the new model not only has the ability to handle equidispersion, underdispersion and overdispersion, but also gives good performances in explaining the zero-inflated phenomenon.

The contents of this paper are organized as follows. In Sect. 2, we introduce the new model. Some probabilistic and statistical properties are investigated. In Sect. 3, the CML method is used to estimate the model parameters. Section 4 presents some simulation studies for the proposed estimation method. In Sect. 5, the forecasting problem is addressed. Section 6 gives two real data examples and the forecasting methods discussed in Sect. 5 are applied. The paper ends with a discussion section.

## 2 Definition and properties of the new process

In this section, a mixture INAR(1) model with a finite range  $\{0, 1, \dots, n\}$  based on the mixing Pogram and binomial thinning operators is presented. Note that the BAR(1)

process defined in (3) comprises two thinned elements:  $\alpha \circ X_{t-1}$  and  $\beta \circ (n - X_{t-1})$ . Pegram's mixing operator ( $*$ ) with mixing weight  $\phi$  on the two integer-valued random variables  $\alpha \circ X_{t-1}$  and  $\beta \circ (n - X_{t-1})$  yields the proposed model given below.

**Definition 3** (*Mixture of Pegram-BAR(1) Model*) Let  $\phi, \alpha, \beta \in (0; 1)$ . Fix  $n \in \mathbb{N}$  and the initial value of the process  $X_0 \in \{0, 1, \dots, n\}$ . Then the mixture of Pegram-BAR(1) model  $\{X_t\}$  is defined by the recursion

$$X_t = (\phi, \alpha \circ X_{t-1}) * (1 - \phi, \beta \circ (n - X_{t-1})), \quad (4)$$

where  $\circ$  and  $*$  are the binomial and mixing Pegram thinning operators, respectively. The random variables  $\alpha \circ X_{t-1}$  and  $\beta \circ (n - X_{t-1})$  are independent of each other when  $X_{t-1}$  is given. All thinnings are performed independently of each other and the thinnings at time  $t$  are independent of  $(X_s)_{s < t}$ . For convenience, we denote the new model by MPTBAR(1) (BAR(1) model with the mixture of Pegram and binomial thinning operators) model.

Transition probabilities are very important in determining the process since the MPTBAR(1) model is Markovian. The transition probabilities of the MPTBAR(1) model are given in the following proposition.

**Proposition 1** For fixed  $n \in \mathbb{N}$ , the transition probabilities of the MPTBAR(1) model are given by

$$\begin{aligned} P(X_t = i | X_{t-1} = j) &= I_{\{i \leq j, i \leq n-j\}} \left( \phi \binom{j}{i} \alpha^i (1 - \alpha)^{j-i} \right. \\ &\quad \left. + (1 - \phi) \binom{n-j}{i} \beta^i (1 - \beta)^{n-j-i} \right) \\ &\quad + I_{\{i \leq j, n-j < i\}} \phi \binom{j}{i} \alpha^i (1 - \alpha)^{j-i} \\ &\quad + I_{\{j < i, i \leq n-j\}} (1 - \phi) \binom{n-j}{i} \beta^i (1 - \beta)^{n-j-i}. \quad (5) \end{aligned}$$

**Remark 1** As pointed out by a referee: from (5), the MPTBAR(1) model encounters the problem that the impossible one-step transitions exist. To be specific, for  $i, j \in \{0, 1, \dots, n\}$ , we have  $P(X_t = i | X_{t-1} = j) = 0$  if  $j < i$  and  $n - j < i$ . For example, without loss of generality, suppose that  $n = 6$ , we have  $P(X_t = 6 | X_{t-1} = 1) = 0$ ,  $P(X_t = 6 | X_{t-1} = 2) = P(X_t = 5 | X_{t-1} = 2) = 0$ ,  $P(X_t = 6 | X_{t-1} = 3) = P(X_t = 5 | X_{t-1} = 3) = P(X_t = 4 | X_{t-1} = 3) = 0$ ,  $P(X_t = 6 | X_{t-1} = 4) = P(X_t = 5 | X_{t-1} = 4) = 0$  and  $P(X_t = 6 | X_{t-1} = 5) = 0$ . With the referee's help, this problem will be fixed in Sect. 7.

The strict stationarity and ergodicity of the MPTBAR(1) model are very important to establish the asymptotic properties of the parameter estimates. Then we state the following theorem.

**Theorem 1** *The process  $\{X_t\}$  is an irreducible, aperiodic and positive recurrent (and thus ergodic) Markov chain. Hence, there exists a strictly stationary process satisfying Eq. (4).*

**Proof** Suppose that  $I$  is the state space of  $\{X_t\}$ . From Definition 3, we know that  $I = \{0, 1, \dots, n\}$  is finite. Firstly, from (5), for  $\forall i, j \in I$ , we have

$$P(X_{t+1} = 0|X_t = j) = \phi(1 - \alpha)^j + (1 - \phi)(1 - \beta)^{n-j} > 0$$

and

$$P(X_{t+1} = i|X_t = 0) = I_{\{i=0\}}[\phi + (1 - \phi)(1 - \beta)^n] + I_{\{i \neq 0\}}\left[(1 - \phi)\binom{n}{i}\beta^i(1 - \beta)^{n-i}\right] > 0.$$

For  $\forall i, j \in I$ , we have

$$\begin{aligned} P(X_{t+2} = i|X_t = j) &= \sum_{k=0}^n P(X_{t+2} = i, X_{t+1} = k|X_t = j) \\ &= \sum_{k=0}^n P(X_{t+1} = k|X_t = j)P(X_{t+2} = i|X_t = j, X_{t+1} = k) \\ &= \sum_{k=0}^n P(X_{t+1} = k|X_t = j)P(X_{t+2} = i|X_{t+1} = k) \\ &\geq P(X_{t+1} = 0|X_t = j)P(X_{t+2} = i|X_{t+1} = 0) > 0. \end{aligned}$$

For  $\forall i, j \in I$ , based on  $P(X_{t+2} = i|X_t = j) > 0$ , we have

$$\begin{aligned} P(X_{t+3} = i|X_t = j) &= \sum_{k=0}^n P(X_{t+3} = i, X_{t+2} = k|X_t = j) \\ &= \sum_{k=0}^n P(X_{t+2} = k|X_t = j)P(X_{t+3} = i|X_t = j, X_{t+2} = k) \\ &= \sum_{k=0}^n P(X_{t+2} = k|X_t = j)P(X_{t+3} = i|X_{t+2} = k) \\ &\geq P(X_{t+2} = 0|X_t = j)P(X_{t+3} = i|X_{t+2} = 0) > 0. \end{aligned}$$

By the similar way, for  $\forall i, j \in I$  and  $s \geq 4$ , we can prove that

$$P(X_{t+s} = i|X_t = j) > 0.$$

Thus, we can conclude that for  $\forall i, j \in I$  and  $m \geq 2$ , we have

$$P(X_{t+m} = i | X_t = j) > 0.$$

This implies primitivity and thus the process is ergodic with uniquely determined stationary marginal distribution since we have a finite Markov Chain.  $\square$

Next, some statistical properties, such as the autocorrelation function and binomial dispersion index BID, are studied. Since the one-step conditional moments are the most important regression properties, we give the following proposition now.

**Proposition 2** *Suppose  $\{X_t\}$  is the stationary process defined in (4), the one-step conditional moments of  $\{X_t\}$  are given by*

$$\begin{aligned} E(X_t | X_{t-1}) &= [\phi\alpha - (1 - \phi)\beta]X_{t-1} + (1 - \phi)n\beta, \\ E(X_t^2 | X_{t-1}) &= (n^2 - n)(1 - \phi)\beta^2 + n(1 - \phi)\beta + [\phi\alpha^2 + (1 - \phi)\beta^2]X_{t-1}^2 \\ &\quad + [\phi\alpha - \phi\alpha^2 + (1 - \phi)\beta^2 - 2n(1 - \phi)\beta^2 + (1 - \phi)\beta]X_{t-1}. \end{aligned}$$

**Proof** It is easy to obtain

$$\begin{aligned} E[X_t | X_{t-1}] &= \phi\alpha X_{t-1} + (1 - \phi)\beta(n - X_{t-1}) \\ &= [\phi\alpha - (1 - \phi)\beta]X_{t-1} + (1 - \phi)n\beta \end{aligned}$$

and

$$\begin{aligned} E[X_t^2 | X_{t-1}] &= \phi E[(\alpha \circ X_{t-1})^2 | X_{t-1}] + (1 - \phi) E[(\beta \circ (n - X_{t-1}))^2 | X_{t-1}] \\ &= \phi[\alpha X_{t-1} + \alpha^2 X_{t-1}(X_{t-1} - 1)] \\ &\quad + (1 - \phi)[\beta(n - X_{t-1}) + \beta^2(n - X_{t-1})(n - X_{t-1} - 1)] \\ &= (1 - \phi)(n\beta + n^2\beta^2 - n\beta^2) + [\phi\alpha^2 + (1 - \phi)\beta^2]X_{t-1}^2 \\ &\quad + [\phi(\alpha - \alpha^2) + (1 - \phi)(\beta^2 - 2n\beta^2 - \beta)]X_{t-1}. \end{aligned}$$

$\square$

Based on the one-step conditional moments of the MPTBAR(1) process in Proposition 2, the mean and variance of our model can be given by the following proposition.

**Proposition 3** *Suppose  $\{X_t\}$  is the stationary process defined in (4), the mean and variance of  $\{X_t\}$  are given by*

$$\begin{aligned} E(X_t) &= \frac{(1 - \phi)n\beta}{1 - [\phi\alpha - (1 - \phi)\beta]}, \\ \text{Var}(X_t) &= \frac{\phi(1 - \phi)n\beta(\alpha - \alpha^2) + (1 - \phi)^2 n\beta(\beta^2 - 2n\beta^2 - \beta)}{[1 - (\phi\alpha^2 + (1 - \phi)\beta^2)][1 - (\phi\alpha + (1 - \phi)\beta)]} \end{aligned}$$

$$+ \frac{(1 - \phi)(n^2\beta^2 - n\beta^2 + n\beta)}{1 - [\phi\alpha^2 + (1 - \phi)\beta^2]} - \frac{(1 - \phi)^2n^2\beta^2}{[1 - (\phi\alpha + (1 - \phi)\beta)]^2}.$$

**Proof** Based on Proposition 2, we have

$$\begin{aligned} E(X_t) &= E[E(X_t|X_{t-1})] = E[(\phi\alpha - (1 - \phi)\beta)X_{t-1} + (1 - \phi)n\beta] \\ &= \frac{(1 - \phi)n\beta}{1 - [\phi\alpha - (1 - \phi)\beta]}, \\ E(X_t^2) &= E[E(X_t^2|X_{t-1})] = \frac{(1 - \phi)(n^2\beta^2 - n\beta^2 + n\beta)}{1 - [\phi\alpha^2 + (1 - \phi)\beta^2]} \\ &\quad + \frac{\phi(\alpha - \alpha^2) + (1 - \phi)(\beta^2 - 2n\beta^2 - \beta)}{1 - [\phi\alpha^2 + (1 - \phi)\beta^2]}E(X_t) \\ &= \frac{\phi(1 - \phi)n\beta(\alpha - \alpha^2) + (1 - \phi)^2n\beta(\beta^2 - 2n\beta^2 - \beta)}{[1 - (\phi\alpha^2 + (1 - \phi)\beta^2)][1 - (\phi\alpha - (1 - \phi)\beta)]} \\ &\quad + \frac{(1 - \phi)(n^2\beta^2 - n\beta^2 + n\beta)}{1 - [\phi\alpha^2 + (1 - \phi)\beta^2]}, \\ \text{Var}(X_t) &= E(X_t^2) - [E(X_t)]^2 \\ &= \frac{\phi(1 - \phi)n\beta(\alpha - \alpha^2) + (1 - \phi)^2n\beta(\beta^2 - 2n\beta^2 - \beta)}{[1 - (\phi\alpha^2 + (1 - \phi)\beta^2)][1 - (\phi\alpha + (1 - \phi)\beta)]} \\ &\quad + \frac{(1 - \phi)(n^2\beta^2 - n\beta^2 + n\beta)}{1 - [\phi\alpha^2 + (1 - \phi)\beta^2]} - \frac{(1 - \phi)^2n^2\beta^2}{[1 - (\phi\alpha + (1 - \phi)\beta)]^2}. \end{aligned}$$

□

Next, we give the expression for the binomial dispersion index of the MPTBAR(1) process.

**Proposition 4** Suppose  $\{X_t\}$  is the stationary process defined in (4), the binomial dispersion index BID of  $\{X_t\}$  is given by

$$\begin{aligned} \text{BID} &= \frac{[1 - (\phi\alpha - (1 - \phi)\beta)]^2[\phi(\alpha - \alpha^2) + (1 - \phi)(\beta^2 - 2n\beta^2 - \beta)]}{(1 - \phi\alpha)[1 - (\phi\alpha^2 + (1 - \phi)\beta^2)][1 - (\phi\alpha + (1 - \phi)\beta)]} \\ &\quad + \frac{(n\beta - \beta + 1)[1 - (\phi\alpha - (1 - \phi)\beta)]^2}{(1 - \phi\alpha)[1 - (\phi\alpha^2 + (1 - \phi)\beta^2)]} \\ &\quad - \frac{(1 - \phi)n\beta[1 - (\phi\alpha - (1 - \phi)\beta)]^2}{(1 - \phi\alpha)[1 - (\phi\alpha + (1 - \phi)\beta)]^2}. \end{aligned}$$

We will see in Sect. 4, the MPTBAR(1) model has ability to model equidispersion, overdispersion and underdispersion according to the true values of the model parameters. In the following proposition, we consider the autocorrelation function of the MPTBAR(1) model.

**Proposition 5** Suppose  $\{X_t\}$  is the stationary process defined in (4), for any non-negative integer  $h$ , the autocovariance at lag  $h$  is given by

$$\text{Cov}(X_t, X_{t+h}) = [\phi\alpha - (1 - \phi)\beta]^h \text{Var}(X_t). \tag{6}$$

**Proof** It is easy to obtain

$$\begin{aligned} \text{Cov}(X_t, X_{t+h}) &= E(X_t X_{t+h}) - E(X_{t+h})E(X_t) \\ &= E[X_t(\phi\alpha X_{t+h-1} + (1 - \phi)\beta(n - X_{t+h-1}))] \\ &\quad - E(X_t)E(\phi\alpha X_{t+h-1} + (1 - \phi)\beta(n - X_{t+h-1})) \\ &= [\phi\alpha - (1 - \phi)\beta]E(X_t X_{t+h-1}) \\ &\quad - [\phi\alpha - (1 - \phi)\beta]E(X_t)E(X_{t+h-1}) \\ &= [\phi\alpha - (1 - \phi)\beta]\text{Cov}(X_t, X_{t+h-1}) \\ &= \dots = [\phi\alpha - (1 - \phi)\beta]^h \text{Var}(X_t). \end{aligned}$$

□

**Remark 2** From (6), it follows that the autocorrelation function of the MPTBAR(1) model can be given by  $\text{Corr}(X_t, X_{t+h}) = [\phi\alpha - (1 - \phi)\beta]^h$  for  $h \geq 0$ .

To obtain the unique stationary distribution of the MPTBAR(1) process, a Markov-Chain approach, proposed by Weiß (2010), is applied. Let  $\mathbf{P}$  denote the transition matrix of the MPTBAR(1) process, i.e.,

$$\mathbf{P} = (P(j|i))_{i,j=0,1,2,\dots} \tag{7}$$

with  $P(j|i) = P(X_t = j | X_{t-1} = i)$  is the transition probability given in (5). Let  $\mathbf{\Pi}$  denote the stationary marginal distribution of our process. Then, by solving the invariance equation  $\mathbf{\Pi} = \mathbf{\Pi P}$ , the marginal distribution will be obtained. We will show the the marginal distribution plots of our model in Sect. 4.

### 3 Conditional maximum likelihood estimation

Suppose that  $\{X_t\}$  is a strictly stationary and ergodic solution of the MPTBAR(1) model. Our aim is to estimate the parameter  $\boldsymbol{\eta} = (\alpha, \beta, \phi)^\top$  from a sample  $(X_1, X_2, \dots, X_T)$ . We assume that  $n$  is known. The CML method is used to estimate the model parameters. Furthermore, some analytical and asymptotic results for estimators are derived. Assume that  $x_1$  is fixed. The CML estimates can be obtained by maximizing the conditional log-likelihood function for the MPTBAR(1) model

$$\ell(\boldsymbol{\eta}) = \log \left( \prod_{t=2}^T P(X_t | X_{t-1}) \right) = \sum_{t=2}^T \log (P(X_t | X_{t-1})), \tag{8}$$



where  $P(X_t|X_{t-1})$  is defined in (5). We use numerical optimization methods to obtain the CML estimates because the CML estimates do not have closed form.

**Theorem 2** *There exists a consistent CML estimator of  $\eta$ , maximizing (8), that is also asymptotically normally distributed,*

$$\sqrt{T-1} \begin{pmatrix} \widehat{\alpha}_{CML} - \alpha \\ \widehat{\beta}_{CML} - \beta \\ \widehat{\phi}_{CML} - \phi \end{pmatrix} \xrightarrow{d} N(\mathbf{0}, \mathbf{I}^{-1}(\eta)),$$

$\mathbf{I}(\eta)$  is the Fisher information matrix.

**Proof** To prove Theorem 2, we have to check if Condition 5.1 in Billingsley (1961) is satisfied. Let  $I$  be the state space of  $\{X_t\}$  and  $P_\eta(j|i) = P(X_t = j|X_{t-1} = i)$ . For  $\forall i, j \in I$ , the set  $D$  of  $(i, j)$  such that  $P_\eta(j|i) > 0$ . First,  $D$  is independent of  $\eta$ . Furthermore, for  $\forall (i, j) \in D$ ,  $P_\eta(j|i)$  are polynomials in  $\alpha, \beta, \phi$ . So partial derivatives in  $\alpha, \beta, \phi$  up to any order exist and are continuous functions in  $\alpha, \beta, \phi$ .

Without loss of generality, we will assume  $n > 2$  in the proof below. The gradients of

$$\begin{aligned} P_\eta(0|0) &= \phi + 2(1 - \phi)(1 - \beta)^n, \\ P_\eta(0|1) &= \phi(1 - \alpha) + 2(1 - \phi)(1 - \beta)^{n-1}, \\ P_\eta(0|2) &= \phi(1 - \alpha)^2 + 2(1 - \phi)(1 - \beta)^{n-2}, \end{aligned}$$

are shown to be linearly independent, i.e. the matrix

$$\begin{pmatrix} \partial P_\eta(0|0)/\partial\alpha & \partial P_\eta(0|0)/\partial\beta & \partial P_\eta(0|0)/\partial\phi \\ \partial P_\eta(0|1)/\partial\alpha & \partial P_\eta(0|1)/\partial\beta & \partial P_\eta(0|1)/\partial\phi \\ \partial P_\eta(0|2)/\partial\alpha & \partial P_\eta(0|2)/\partial\beta & \partial P_\eta(0|2)/\partial\phi \end{pmatrix}$$

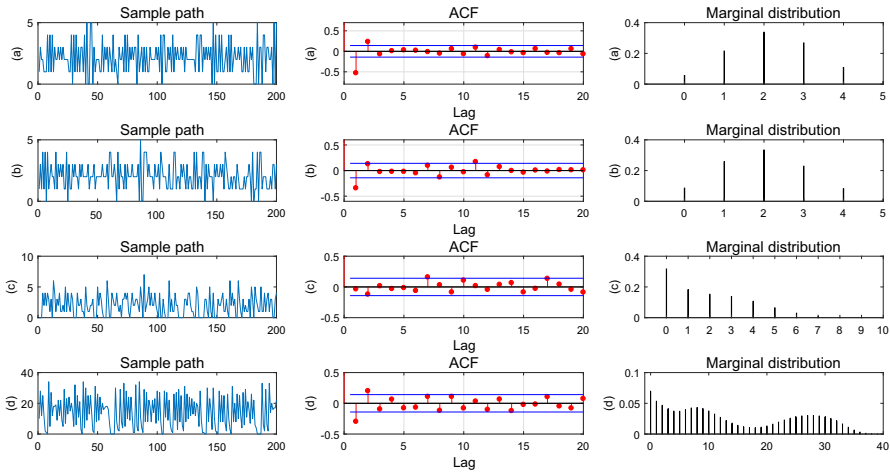
has rank 3.

So Condition 5.1 in Billingsley (1961) is satisfied, and there exists a consistent CML estimator of  $\eta$  that is also asymptotic normally distributed (Billingsley 1961, Theorems 2.1 and 2.2). □

### 4 Simulation study

In this section, we conduct a simulation study to the finite sample performances of the CML estimates. The initial value  $X_0$  is fixed at 3. We generate the data from the MPTBAR(1) model and set the sample sizes  $T = 50, 100, 200, 500$ . In simulations, the mean squared error (MSE), mean absolute deviation error (MADE) and standard deviation (SD) are computed over 1000 replications. The true values of the parameters are selected as:

- (a)  $(n, \alpha, \beta, \phi) = (5, 0.9, 0.8, 0.1)$  (BID = 0.9845, underdispersion);
- (b)  $(n, \alpha, \beta, \phi) = (5, 0.8, 0.7, 0.2)$  (BID = 1, equidispersion);



**Fig. 1** Sample paths, ACF and marginal distributions of the MPTBAR(1) model for: (a)  $(n, \alpha, \beta, \phi) = (5, 0.9, 0.8, 0.1)$ ; (b)  $(n, \alpha, \beta, \phi) = (5, 0.8, 0.7, 0.2)$ ; (c)  $(n, \alpha, \beta, \phi) = (10, 0.3, 0.4, 0.5)$ ; (d)  $(n, \alpha, \beta, \phi) = (40, 0.3, 0.8, 0.4)$

- (c)  $(n, \alpha, \beta, \phi) = (10, 0.3, 0.4, 0.5)$  (BID = 2.1155, overdispersion; zero-inflation);
- (d)  $(n, \alpha, \beta, \phi) = (40, 0.3, 0.8, 0.4)$  (multimodality).

Figure 1 shows the sample paths, autocorrelation function (ACF), marginal distributions of some typical MPTBAR(1) models. Note that the ACF of the MPTBAR(1) model with  $(n, \alpha, \beta, \phi) = (10, 0.3, 0.4, 0.5)$  indicates that the samples are not stemmed from a first-order autoregressive process. The explanation for this phenomenon is that the first-order autocorrelation coefficients of the MPTBAR(1) model is close to zero under this situation.

The summary of the simulation results are shown in Table 1. We observe that MSE, MADE and SD of the estimators decrease as the sample size  $n$  increases, as expected. The CML estimation method can produce reliable estimates for the model parameters. Figure 2 shows the Q–Q plots of CML estimators for the MPTBAR(1) model with  $(n, \alpha, \beta, \phi) = (10, 0.3, 0.4, 0.5)$ . Q–Q plots in Fig. 2 appear to be roughly normally distributed as expected. Similar results can be obtained from the MPTBAR(1) model with other parameter combinations. To save space, the figures are omitted here.

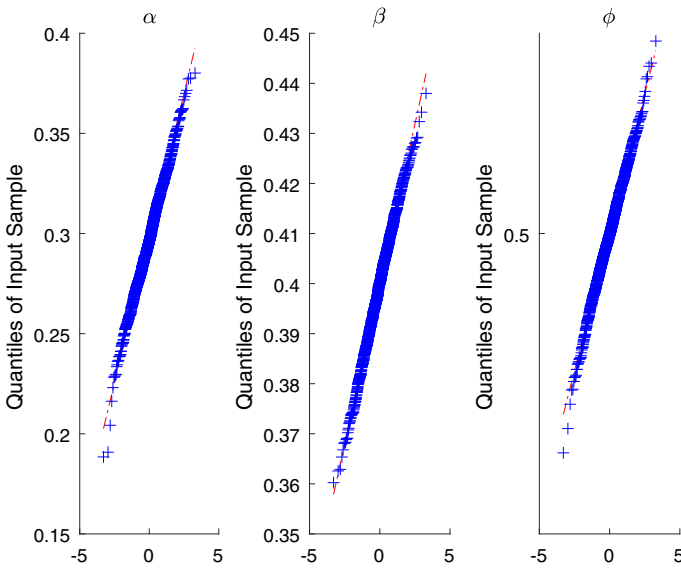
### 5 Forecasting for the MPTBAR(1) process

The forecasting problem is always an important topic in time series analysis. The most common method of forecasting is to use the conditional expectation, which yields forecasts with minimum mean square error. Based on the above reason, We use this method to forecast a MPTBAR(1) process. The  $h$ -step-ahead predictor of the MPTBAR(1) model is given by

$$\widehat{X}_{t+h} | X_t = E[X_{t+h} | X_t], \quad h = 1, 2, \dots$$

**Table 1** MSE, MADE and SD of the CML estimates

Sample size	$(\alpha, \beta, \phi)$	$\alpha$			$\beta$			$\phi$		
		MSE	MADE	SD	MSE	MADE	SD	MSE	MADE	SD
		$T = 50$	0.0475	0.1466	0.2124	0.0015	0.0306	0.0383	0.0037	0.0490
$T = 100$	0.0146	0.0895	0.1210	0.0008	0.0217	0.0274	0.0017	0.0324	0.0408	
$T = 200$	0.0061	0.0626	0.0781	0.0004	0.0159	0.0198	0.0009	0.0239	0.0297	
$T = 500$	0.0023	0.0387	0.0483	0.0002	0.0102	0.0126	0.0003	0.0148	0.0185	
$T = 50$	(0.8, 0.7, 0.2)	0.0304	0.1297	0.1745	0.0023	0.0383	0.0477	0.0075	0.0688	0.0864
$T = 100$	(0.8, 0.7, 0.2)	0.0128	0.0859	0.1130	0.0013	0.0283	0.0356	0.0040	0.0506	0.0636
$T = 200$	(0.8, 0.7, 0.2)	0.0046	0.0529	0.0678	0.0006	0.0198	0.0247	0.0018	0.0340	0.0427
$T = 500$	(0.8, 0.7, 0.2)	0.0019	0.0341	0.0433	0.0002	0.0119	0.0150	0.0008	0.0221	0.0276
$T = 50$	(0.3, 0.4, 0.5)	0.0091	0.0758	0.0952	0.0016	0.0320	0.0401	0.0086	0.0740	0.0924
$T = 100$	(0.3, 0.4, 0.5)	0.0045	0.0542	0.0670	0.0008	0.0222	0.0278	0.0040	0.0509	0.0633
$T = 200$	(0.3, 0.4, 0.5)	0.0019	0.0346	0.0436	0.0004	0.0159	0.0199	0.0021	0.0376	0.0462
$T = 500$	(0.3, 0.4, 0.5)	0.0008	0.0222	0.0279	0.0002	0.0098	0.0123	0.0008	0.0233	0.0290
$T = 50$	(0.3, 0.8, 0.4)	0.0011	0.0260	0.0326	0.0002	0.0120	0.0152	0.0061	0.0626	0.0781
$T = 100$	(0.3, 0.8, 0.4)	0.0005	0.0182	0.0223	0.0001	0.0086	0.0107	0.0028	0.0422	0.0528
$T = 200$	(0.3, 0.8, 0.4)	0.0002	0.0125	0.0158	0.0001	0.0059	0.0075	0.0013	0.0287	0.0360
$T = 500$	(0.3, 0.8, 0.4)	0.0001	0.0075	0.0095	0.0000	0.0038	0.0048	0.0006	0.0201	0.0250



**Fig. 2** Q-Q plots of the CML estimators for the MPTBAR(1) model with  $n = 10$ ,  $\alpha = 0.3$ ,  $\beta = 0.4$ ,  $\phi = 0.5$  and sample size  $T = 500$

From the properties of the binomial thinning operator, we have the following proposition.

**Proposition 6** *Suppose  $\{X_t\}$  is defined in (4), the  $h$ -step conditional mean of  $X_t$  can given by*

$$E(X_{t+h}|X_t) = [\phi\alpha - (1 - \phi)\beta]^h X_t + (1 - \phi)n\beta \frac{1 - [\phi\alpha - (1 - \phi)\beta]^h}{1 - [\phi\alpha - (1 - \phi)\beta]}.$$

**Proof** It is easy to obtain

$$\begin{aligned} E(X_{t+1}|X_t) &= [\phi\alpha - (1 - \phi)\beta]X_t + n(1 - \phi)\beta, \\ E(X_{t+2}|X_t) &= E[(\phi\alpha X_{t+1} + (1 - \phi)\beta(n - X_{t+1}))|X_t] \\ &= E\left[[\phi\alpha - (1 - \phi)\beta](\phi\alpha X_t + (1 - \phi)\beta(n - X_t)) + (1 - \phi)n\beta\right|X_t] \\ &= [\phi\alpha - (1 - \phi)\beta]^2 X_t + [\phi\alpha - (1 - \phi)\beta](1 - \phi)n\beta + (1 - \phi)n\beta, \\ E(X_{t+h}|X_t) &= [\phi\alpha - (1 - \phi)\beta]^h X_t + (1 - \phi)n\beta \sum_{j=0}^{h-1} [\phi\alpha - (1 - \phi)\beta]^j \\ &= [\phi\alpha - (1 - \phi)\beta]^h X_t + (1 - \phi)n\beta \frac{1 - [\phi\alpha - (1 - \phi)\beta]^h}{1 - [\phi\alpha - (1 - \phi)\beta]}. \end{aligned}$$

□

**Remark 3** It can be easily seen that  $E(X_t) = \lim_{h \rightarrow \infty} E(X_{t+h}|X_t) = \frac{(1-\phi)n\beta}{1-[\phi\alpha-(1-\phi)\beta]}.$

From Proposition 6, based on the sample  $X_1, X_2, \dots, X_t$ , the  $h$ -step-ahead predictor  $\widehat{X}_{t+h}$ ,  $h \in \mathbb{N}$ , can be given by

$$\widehat{X}_{t+h} = [\widehat{\phi}\widehat{\alpha} - (1 - \widehat{\phi})\widehat{\beta}]^h X_t + (1 - \widehat{\phi})n\widehat{\beta} \frac{1 - [\widehat{\phi}\widehat{\alpha} - (1 - \widehat{\phi})\widehat{\beta}]^h}{1 - [\widehat{\phi}\widehat{\alpha} - (1 - \widehat{\phi})\widehat{\beta}]}, \quad h \in \mathbb{N},$$

where  $\widehat{\alpha}$ ,  $\widehat{\beta}$  and  $\widehat{\phi}$  are estimators for  $\alpha$ ,  $\beta$  and  $\phi$ , respectively.

However, this procedure will seldom produce integer-valued  $\widehat{X}_{t+h}$ . For this, Freeland and McCabe (2004) proposed that using the  $h$ -step-ahead forecasting conditional distributions to forecast the future value. Freeland and McCabe (2004), Möller et al. (2016), Li et al. (2018) and Maiti and Biswas (2017) have applied this method to forecast their processes. As pointed out by Möller et al. (2016), this approach leads to forecasts being themselves counts and therefore being coherent with the sample space, and the point forecasts are easily obtained from the median or the mode of the forecasting distribution. By the Chapman–Kolmogorov equations,  $p_h(x_{t+h}|X_t)$ , the  $h$ -step-ahead conditional distribution of  $X_{t+h}$  given  $X_t$  of the MPTBAR(1) process, can be given by

$$p_h(x_{t+h}|X_t = x_t) = P(X_{t+h} = x_{t+h}|X_t = x_t) = [\mathbf{P}^h]_{\mathbf{x}_t, \mathbf{x}_{t+h}},$$

where  $\mathbf{P}$  denotes the transition matrix defined by (7).

Now, we have obtained the  $h$ -step-ahead conditional distribution, as  $p_h(x_{t+h}|X_t, \boldsymbol{\eta})$ , where  $\boldsymbol{\eta} = (\alpha, \beta, \phi)$ . Estimating  $\boldsymbol{\eta}$  before we implement the forecasting method is the problem we concern. For this, we can use CML method in practice. As pointed out by Theorem 2, the CML estimate  $\widehat{\boldsymbol{\eta}}_{CML}$  is asymptotically normally distributed around the true value  $\boldsymbol{\eta}_0$ . Following Li et al. (2018), we have the next theorem.

**Theorem 3** For a fixed  $x \in \{0, 1, \dots, n\}$ , the quantity  $p_h(x|X_t, \widehat{\boldsymbol{\eta}}_{CML})$  has an asymptotically normal distribution, i.e.,

$$\sqrt{T-1}(p_h(x|X_t, \widehat{\boldsymbol{\eta}}_{CML}) - p_h(x|X_t, \boldsymbol{\eta}_0)) \xrightarrow{d} N(0, \mathbf{D}\mathbf{I}^{-1}(\boldsymbol{\eta})\mathbf{D}^\top),$$

where  $\mathbf{D} = \left( \left. \frac{\partial p_h(x|X_t, \boldsymbol{\eta})}{\partial \alpha} \right|_{\boldsymbol{\eta}=\boldsymbol{\eta}_0}, \left. \frac{\partial p_h(x|X_t, \boldsymbol{\eta})}{\partial \beta} \right|_{\boldsymbol{\eta}=\boldsymbol{\eta}_0}, \left. \frac{\partial p_h(x|X_t, \boldsymbol{\eta})}{\partial \phi} \right|_{\boldsymbol{\eta}=\boldsymbol{\eta}_0} \right)$ ,  $\mathbf{I}(\boldsymbol{\eta})$  is the Fisher information matrix in Theorem 2.

Based on Theorem 5, the  $100(1 - \alpha)\%$  confidence interval for  $p_h(x_{t+h}|X_t, \boldsymbol{\eta})$  can be given by

$$C_\eta^\alpha = \left( p_h(x|X_t, \widehat{\boldsymbol{\eta}}_{CML}) - \frac{\sigma}{\sqrt{T-1}}u_{1-\frac{\alpha}{2}}, p_h(x|X_t, \widehat{\boldsymbol{\eta}}_{CML}) + \frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}} \right),$$

where  $\sigma = \sqrt{\mathbf{D}\mathbf{I}^{-1}(\boldsymbol{\eta})\mathbf{D}^\top}$ ,  $u_{1-\frac{\alpha}{2}}$  is the  $(1 - \frac{\alpha}{2})$ -upper quantile of  $N(0, 1)$ .

To compare the two forecasting methods, we will apply the two methods to a real data in Sect. 6.

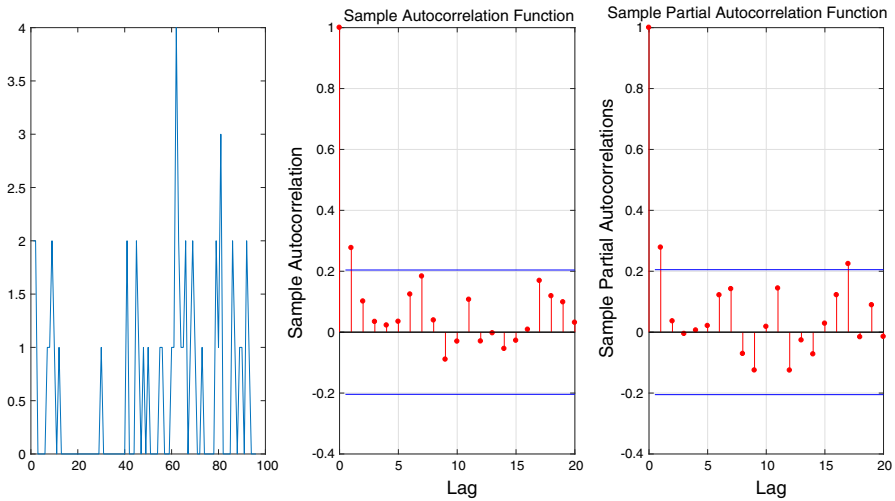


Fig. 3 Sample path, ACF and PACF of the prostitution data

Table 2 Descriptive statistics for the prostitution data

Minimum	Median	Mean	Variance	BID	$p_0$	Maximum	n
0	0	0.53	0.67	1.28	0.64	4	96

## 6 Data analysis

In this section, the first application of the MPTBAR(1) model is conducted to real data for illustrative purposes. We consider a data set which represents monthly counts of car beats in Pittsburgh (among  $n = 42$  such car beats) that had at least one police offense report of prostitution in that month. The data consist of 96 observations, starting from January 1990 and ending in December 1997. This data set has been investigated by Möller et al. (2018). Figure 3 shows the sample path, ACF and PACF of the observations. The descriptive statistics for the data are listed in Table 2. The binomial dispersion index of the data set is given by 1.28, indicating that the data set shows overdispersion.

We illustrate the competitiveness and usefulness of the MPTBAR(1) model in applications by comparing our process with the following models:

- BAR(1) model (McKenzie 1985);
- RZ-BAR(1), IZ-BAR(1), ZIB-AR(1) and ZT<sup>0</sup>-BAR(1) models (Möller et al. 2018). The unknown parameters of the fitted models are estimated by the (conditional) maximum likelihood method. Also, the following statistics are derived: Akaike information criterion (AIC), Bayesian information criterion (BIC), the binomial index of dispersion BID and zero frequency  $p_0$ .

From Table 3, we conclude that the BAR(1) model is not suitable to fit this data set because the BAR(1) model fails to capture the zero-inflated and overdispersed

**Table 3** Estimates of the parameters and statistics for the prostitution data

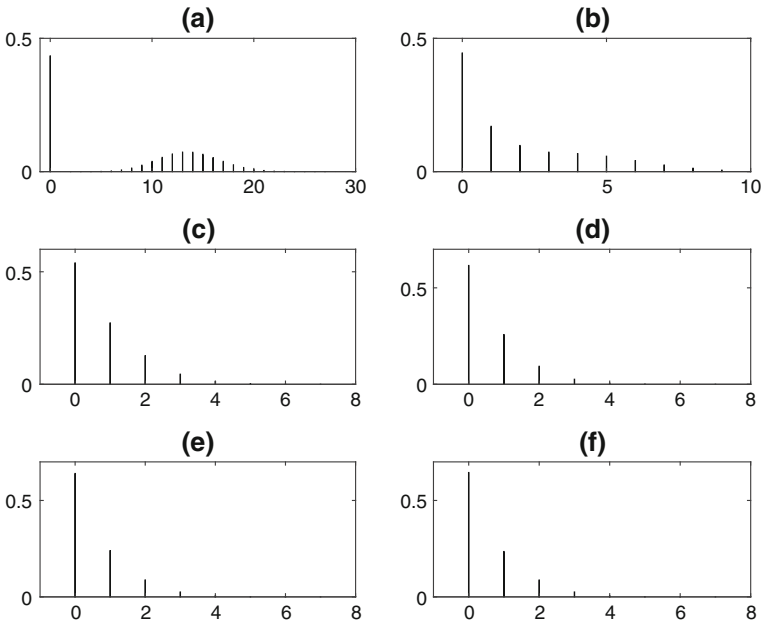
Model	Estimates	AIC	BIC	BID	$p_0$
MPTBAR(1)	$\hat{\alpha} = 0.3240$ $\hat{\beta} = 0.0233$ $\hat{\phi} = 0.5655$	181.5	189.2	1.45	0.65
BAR(1)	$\hat{\rho} = 0.255$ $\hat{\pi} = 0.013$	188.1	193.2	1	0.58
RZ-BAR(1)	$\hat{\rho} = 0.308$ $\hat{\pi} = 0.016$ $\hat{\omega} = 0.198$	189.0	196.7	1.11	0.61
IZ-BAR(1)	$\hat{\rho}_{IZ} = 0.251$ $\hat{\pi}_{IZ} = 0.013$ $\hat{\omega} = 0.251$	188.1	195.8	1.18	0.62
ZIB-AR(1)	$\hat{\rho}_{ZIB} = 0.271$ $\hat{\pi}_{ZIB} = 0.0013$ $\hat{\omega}_\alpha = 0.520$ $\hat{\omega}_\beta = 0.0091$	185.0	192.7	1.31	0.63
ZT <sup>0</sup> -BAR(1)	$\hat{\pi}_0 = 0.008$ $\hat{\pi} = 0.020$	185.7	190.9	1.10	0.59

characteristics of the data. The explanation for this phenomenon is that the marginal distribution of the BAR(1) model is the binomial distribution. While the RZ-BAR(1) and IZ-BAR(1) processes can capture the zero-inflation and overdispersion of the data, the two models perform poorly when we consider AIC and BIC. Comparing the ZIB-AR(1) model with the ZT<sup>0</sup>-BAR(1) model, we find that the ZIB-AR(1) model performs better based on each statistics (except BIC). Furthermore, we find that the MPTBAR(1) and ZIB-AR(1) models provide the most satisfactory results for this data set. To be specific, the MPTBAR(1) and ZIB-AR(1) models have good performances and the MPTBAR(1) model gives a best fit when we consider AIC and BIC. Although the MPTBAR(1) and ZIB-AR(1) models can both describe overdispersion accurately, the ZIB-AR(1) model performs a little better than the MPTBAR(1) model based on BID. Since the zero frequency of the two models are very close to the empirical zero frequency, we conclude that the MPTBAR(1) and ZIB-AR(1) models have the ability to capture the zero-inflated feature of the data precisely. Based on these considerations, the MPTBAR(1) and ZIB-AR(1) models are the most appropriate for this data set.

For the models above, we consider their corresponding Pearson residual analysis. The standardized Pearson residual are defined as

$$e_t = \frac{X_t - E(X_t|X_{t-1})}{\sqrt{\text{Var}(X_t|X_{t-1})}}, \quad t = 2, \dots, T.$$

As pointed out by Möller et al. (2018), if the model is correctly specified, then the residuals should have zero mean, unit variance, and no significant serial correlation



**Fig. 4**  $h$ -step-ahead forecasting conditional distribution of the prostitution data: (a)  $h = 1$ ; (b)  $h = 2$ ; (c)  $h = 3$ ; (d)  $h = 4$ ; (e)  $h = 5$ ; (f)  $h = 6$

**Table 4** Properties of the standardized Pearson residuals

Model	Standardized Pearson residuals			
	Mean	Variance	$\widehat{\rho}_{e_t}(1)$	$\widehat{\rho}_{e_t^2}(1)$
MPTBAR(1)	-0.002	0.94	0.07	-0.05
BAR(1)	-0.037	1.21	0.00	0.06
RZ-BAR(1)	-0.022	1.11	0.03	0.06
IZ-BAR(1)	-0.023	1.05	0.01	0.06
ZIB-AR(1)	-0.025	0.92	-0.02	0.05
ZT <sup>0</sup> -BAR(1)	-0.027	1.14	-0.03	0.03

in  $e_t$  and  $e_t^2$ . We compute the mean, variance, first-order autocorrelation coefficient of the  $e_t$  and  $e_t^2$  in Table 4. Comparing the properties of the residuals in Table 4,  $\widehat{\rho}_{e_t}(1)$  and  $\widehat{\rho}_{e_t^2}(1)$  of the six models are competitive. Based on the mean and variance of the residuals, we find that the MPTBAR(1) model gives the best performance. Thus, the MPTBAR(1) process is the most appropriate model for fitting this data set. Based on the above discussions, we conclude that the MTPBAR(1) process is an useful model to fit the count data with bounded support and suitable to capture the binomial dispersion and zero inflation characteristics of the data.

To check the predictability of the MPTBAR(1) model and compare the two forecasting methods discussed in Sect. 5, another simulation study is given to investigate the  $h$ -step-ahead forecasting for varying  $h$  of the MPTBAR(1) model. We give the



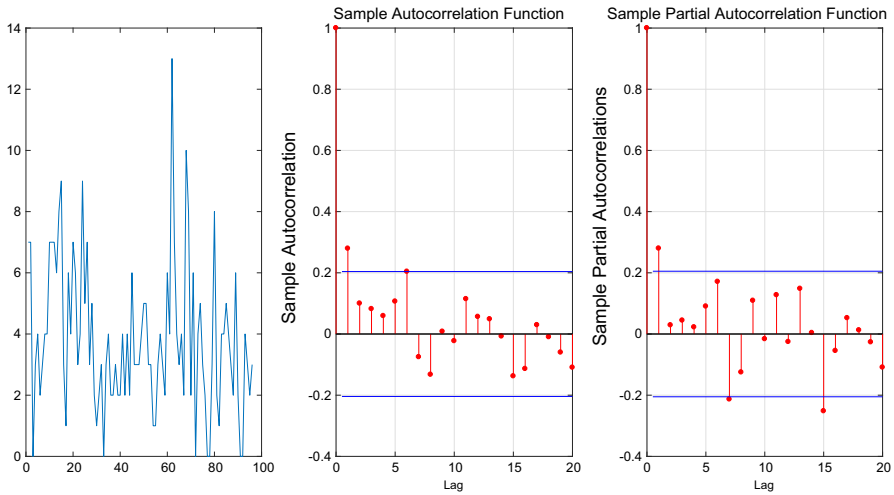


Fig. 5 Sample path, ACF and PACF of the drunken driving data

plots of the  $h$ -step-ahead conditional distribution in Fig. 4 when  $h = 1, 2, 3, 4, 5, 6$ . The median or the mode of  $h$ -step-ahead conditional distribution can be both viewed as the point prediction. For comparison, a standard descriptive measure of forecasting accuracy, namely, predicted mean absolute deviation (PMAD) is adopted. This measure can be give by

$$PMAD = \frac{1}{H} \sum_{h=1}^H |X_{t+h} - \hat{X}_{t+h}|.$$

The conditional expectation and conditional distribution point predictors of the series are presented in Table 5. From Table 5, PMAD value of the mode of  $h$ -step-ahead conditional distribution point predictors is smaller than the  $h$ -step-ahead conditional expectation point predictors. The median of  $h$ -step-ahead conditional distribution point predictors give a poor performance based on PMAD. The reason is that the median of one-step-ahead conditional distribution point predictor is ten which is much greater than the observed value zero. The explanation for this phenomenon may be that the one-step-ahead conditional distribution in Fig. 4 is heavy tail. Based on these facts, we conclude that the mode of  $h$ -step-ahead conditional distribution point predictors are more appropriate for the data set.

As pointed out by a referee, we should compare the results obtained by fitting the MPTBAR(1) and BAR(1) models to the data set, with other competitors capable to cope with overdispersion and underdispersion. For this, we conduct another application of the MPTBAR(1) model to a real data set for comparative purposes. We compare the MPTBAR(1) model with the BAR(1) and SETBAR(1) (Möller et al. 2016) models. The SETBAR(1) model with appropriate parameter settings has ability to capture equidispersion, underdispersion and overdispersion. The second data set is computed from the file PghCarBeat.csv, which was downloaded from The Forecasting Principles

**Table 5**  $h$ -step-ahead predictions for the real data

$h$	Observed values	$h$ -step conditional expectation	$h$ -step-ahead conditional distribution	
			Median	Mode
1	0	0.5983	10	0
2	2	0.5288	1	0
3	1	0.5167	0	0
4	0	0.5146	0	0
5	0	0.5143	0	0
6	0	0.5142	0	0
	PMAD	0.6827	2.0000	0.5000

**Table 6** Descriptive statistics for the drunken driving data

Minimum	Median	Mean	Variance	BID	Maximum	n
0	3	3.82	5.94	1.71	13	96

**Table 7** Comparison of AIC and BIC for the SETBAR(1) model with different threshold values

R	$\alpha_1$	$\beta_1$	$\alpha_2$	$\beta_2$	AIC	BIC
3	0.1724	0.0704	0.1922	0.0924	434.1	444.4
4	0.2764	0.0714	0.2808	0.0710	437.5	447.8
5	0.2794	0.0711	0.2709	0.0734	437.5	447.8
6	0.2200	0.0740	0.2461	0.0908	436.3	446.6
7	0.2624	0.0723	0.3036	0.0710	437.4	447.6

site (<http://www.forecastingprinciples.com>). The data set is given for 42 different car beats and reach from January 1990 to December 1997. For each month  $t$ , the value  $x_t$  counts the number of car beats reported at least one case of drunken driving. So our data have finite range with fixed upper limit  $n = 42$  and the series contains 96 observations.

Figure 5 shows the sample path, ACF and PACF of the observations. The descriptive statistics for the data are listed in Table 6. Table 7 shows the CML estimates, AIC and BIC for the SETBAR(1) model with different threshold values. From the sample path of the observations in Fig. 5, the threshold values  $R \in \{3, \dots, 7\}$  is a reasonable range. By comparing AIC and BIC in Table 7, the SETBAR(1) model with a threshold  $R = 3$  is the best choice. Table 8 lists the CML estimates, AIC and BIC for the MPTBAR(1) model, the BAR(1) model and the SETBAR(1) model with a threshold  $R = 3$ . From Table 8, we find that the BAR(1) model gives the worst fit based on AIC and BIC. Although the SETBAR(1) model gives the best fit when we consider AIC, the MPTBAR(1) model performs best if considering BIC. Thus, the MPTBAR(1) and SETBAR(1) models are competitive for fitting this data set. However, we have to select a suitable threshold by experiments when we decide to use the SETBAR(1) model. Selecting a suitable threshold sometimes may lead to some inconveniences in

**Table 8** Estimates of the parameters and statistics for the drunken driving data

Model	Estimates	AIC	BIC
MPTBAR(1)	$\hat{\alpha} = 0.5961$ $\hat{\beta} = 0.1120$ $\hat{\phi} = 0.2560$	435.2	442.9
BAR(1)	$\hat{p} = 0.0912$ $\hat{r} = 0.2093$	439.6	444.7
SETBAR(1)	$\hat{\alpha}_1 = 0.1724$ $\hat{\beta}_1 = 0.0704$ $\hat{\alpha}_2 = 0.1922$ $\hat{\beta}_2 = 0.0924$	434.1	444.4

practice. Based on these considerations, we recommend the use of the MPTBAR(1) model to fit this data set.

### 7 Discussion

The aim of the present work is to introduce a mixture INAR(1) process to model count data with a finite range  $\{0, 1, \dots, n\}$ . Parameter estimation, forecasting and diagnostic checking for the new model are investigated. Applications to real data sets are given to show the application of the new model. However, more research is still needed for one aspect of the MPTBAR(1) model. The issue is pointed out by a referee: from (5), the MPTBAR(1) model encounters the problem that the impossible one-step transitions exist. For example,  $P(X_t = n | X_{t-1} = 1) = 0$ ,  $P(X_t = n - 1 | X_{t-1} = 2) = P(X_t = n | X_{t-1} = 2) = 0$ . Following the referee’s suggestion, we consider another mixture INAR(1) model  $\{X_t\}$  below to fix the problem.

**Definition 4** Let  $\phi, \alpha \in (0; 1)$ . Fix  $n \in \mathbb{N}$  and the initial value of the process  $X_0 \in \{0, 1, \dots, n\}$ . The new model  $\{X_t\}$  is defined by the recursion

$$X_t = (\phi, \alpha \circ X_{t-1}) * (1 - \phi, \epsilon_t), \tag{9}$$

where  $\circ$  and  $*$  are the binomial and mixing Pegrum thinning operators, respectively.  $\{\epsilon_t\}$  is a sequence of iid discrete random variables on  $\{0, 1, \dots, n\}$ . We suppose that the mean and variance of  $\{\epsilon_t\}$  are  $\mu_\epsilon$  and  $\sigma_\epsilon^2$ .

The one-step transition probabilities of this model are given by

$$\begin{aligned} P(X_t = i | X_{t-1} = j) &= I_{\{i \leq j\}} \left( \phi \binom{j}{i} \alpha^i (1 - \alpha)^{j-i} + (1 - \phi) P(\epsilon_t = i) \right) \\ &\quad + I_{\{j < i\}} (1 - \phi) P(\epsilon_t = i) \\ &= I_{\{i \leq j\}} \phi \binom{j}{i} \alpha^i (1 - \alpha)^{j-i} + (1 - \phi) P(\epsilon_t = i). \end{aligned} \tag{10}$$

From (10), all one-step transition probabilities are positive.

Next, we consider the one-step marginal conditional moments. We obtain

$$\begin{aligned} E(X_t|X_{t-1}) &= \alpha\phi X_{t-1} + (1-\phi)\mu_\epsilon, \\ E(X_t^2|X_{t-1}) &= \alpha^2\phi X_{t-1}^2 + (\alpha\phi - \alpha^2\phi)X_{t-1} + (1-\phi)(\mu_\epsilon^2 + \sigma_\epsilon^2). \end{aligned}$$

Based on the one-step marginal conditional moments, the mean and variance are given by

$$\begin{aligned} E(X_t) &= \frac{(1-\phi)\mu_\epsilon}{1-\alpha\phi}, \\ \text{Var}(X_t) &= \frac{\mu_\epsilon(\alpha\phi - \alpha^2\phi)(1-\phi)}{(1-\alpha\phi)(1-\alpha^2\phi)} + \frac{(1-\phi)(\mu_\epsilon^2 + \sigma_\epsilon^2)}{1-\alpha^2\phi} - \frac{\mu_\epsilon^2(1-\phi)^2}{(1-\alpha\phi)^2}. \end{aligned}$$

The binomial dispersion index of the model is given by

$$\begin{aligned} \text{BID} &= \frac{n\mu_\epsilon(\alpha\phi - \alpha^2\phi)(1-\phi)(1-\alpha\phi)}{[n\mu_\epsilon(1-\phi)(1-\alpha\phi) - \mu_\epsilon^2(1-\phi)^2](1-\alpha^2\phi)} \\ &+ \frac{n(1-\phi)(1-\alpha\phi)^2(\mu_\epsilon^2 + \sigma_\epsilon^2)}{[n\mu_\epsilon(1-\phi)(1-\alpha\phi) - \mu_\epsilon^2(1-\phi)^2](1-\alpha^2\phi)} \\ &- \frac{n\mu_\epsilon^2(1-\phi)^2}{n\mu_\epsilon(1-\phi)(1-\alpha\phi) - \mu_\epsilon^2(1-\phi)^2}. \end{aligned}$$

We verify that the autocovariance function of the process defined in (9) is given by

$$\text{Cov}(X_t, X_{t+h}) = (\alpha\phi)^h \text{Var}(X_t), \quad h \in \{0, 1, \dots\}. \quad (11)$$

From (11), the autocorrelation function is given by  $\text{Corr}(X_t, X_{t+h}) = (\alpha\phi)^h$  for  $h \in \{0, 1, \dots\}$ .

## 7.1 Special cases

Now, we will consider two special cases of the process defined by (9). The first special case is that  $\{\epsilon_t\}$  in (9) is assumed to follow binomial distribution  $B(n, p)$ . Then the mean and variance of  $\{\epsilon_t\}$  are

$$E(\epsilon_t) = np, \quad \text{Var}(\epsilon_t) = np(1-p).$$

The one-step transition probabilities of the model are

$$P(X_t = i | X_{t-1} = j) = I_{\{i \leq j\}} \binom{j}{i} \alpha^i (1-\alpha)^{j-i} + (1-\phi) \binom{n}{i} p^i (1-p)^{n-i}.$$

Following the referee’s suggestion and Möller et al. (2016), since the transition matrix is primitive and the state space of the model is finite, the process is ergodic with uniquely determined stationary marginal distribution.

The one-step conditional moments of the process are given by

$$E(X_t|X_{t-1}) = \alpha\phi X_{t-1} + (1 - \phi)np,$$

$$E(X_t^2|X_{t-1}) = \alpha^2\phi X_{t-1}^2 + (\alpha - \alpha^2)\phi X_{t-1} + np(np + 1 - p)(1 - \phi).$$

The mean and variance of the model are given by

$$E(X_t) = \frac{(1 - \phi)np}{1 - \alpha\phi},$$

$$\text{Var}(X_t) = \frac{np(\alpha - \alpha^2)\phi(1 - \phi)}{(1 - \alpha\phi)(1 - \alpha^2\phi)} + \frac{np(np + 1 - p)(1 - \phi)}{1 - \alpha^2\phi} - \frac{(np)^2(1 - \phi)^2}{(1 - \alpha\phi)^2}.$$

The binomial dispersion index of the model is given by

$$\text{BID} = \frac{n^2 p(\alpha - \alpha^2)\phi(1 - \phi)(1 - \alpha\phi)}{[n^2 p(1 - \phi)(1 - \alpha\phi) - (np)^2(1 - \phi)^2](1 - \alpha^2\phi)}$$

$$+ \frac{n^2 p(1 - \phi)(1 - \alpha\phi)^2(np + 1 - p)}{[n^2 p(1 - \phi)(1 - \alpha\phi) - (np)^2(1 - \phi)^2](1 - \alpha^2\phi)}$$

$$- \frac{n(np)^2(1 - \phi)^2}{n^2 p(1 - \phi)(1 - \alpha\phi) - (np)^2(1 - \phi)^2}.$$

The second special case is that we assume  $\{\epsilon_t\}$  in (9) follows zero-inflated binomial distribution ZIB( $n, p, \pi$ ). The probability mass function of ZIB( $n, p, \pi$ ) is given by

$$P(\epsilon_t = k) = \begin{cases} \pi + (1 - \pi)(1 - p)^n, & k = 0, \\ (1 - \pi)\binom{n}{k}p^k(1 - p)^{n-k}, & k = 1, 2, \dots \end{cases}$$

The mean and variance of ZIB( $n, p, \pi$ ) are

$$E(\epsilon_t) = np(1 - \pi), \quad \text{Var}(\epsilon_t) = (1 - \pi)np(1 - p) + (1 - \pi)(np)^2 - (1 - \pi)^2(np)^2.$$

The one-step transition probabilities of the model are

$$P(X_t = i|X_{t-1} = j) = I_{\{i \leq j\}}\phi \binom{j}{i} \alpha^i (1 - \alpha)^{j-i}$$

$$+ (1 - \phi) \left( \pi I_{\{i=0\}} + (1 - \pi) \binom{n}{i} p^i (1 - p)^{n-i} \right).$$

Similarly, the model is ergodic with uniquely determined stationary marginal distribution.

The one-step conditional moments of the model are given by

$$\begin{aligned} E(X_t|X_{t-1}) &= \alpha\phi X_{t-1} + np(1-\phi)(1-\pi), \\ E(X_t^2|X_{t-1}) &= \alpha^2\phi X_{t-1}^2 + (\alpha\phi - \alpha^2\phi)X_{t-1} \\ &\quad + np(1-p)(1-\phi)(1-\pi) + (np)^2(1-\phi)(1-\pi). \end{aligned}$$

The mean and variance of the model are

$$\begin{aligned} E(X_t) &= \frac{np(1-\phi)(1-\pi)}{1-\alpha\phi}, \\ \text{Var}(X_t) &= \frac{np(\alpha - \alpha^2)\phi(1-\phi)(1-\pi)}{(1-\alpha\phi)(1-\alpha^2\phi)} \\ &\quad + \frac{np(np+1-p)(1-\phi)(1-\pi)}{1-\alpha^2\phi} - \frac{(np)^2(1-\phi)^2(1-\pi)^2}{(1-\alpha\phi)^2}. \end{aligned}$$

The binomial dispersion index BID of the model is given by

$$\begin{aligned} \text{BID} &= \frac{n^2p(\alpha - \alpha^2)(1-\alpha\phi)\phi(1-\phi)(1-\pi)}{[n^2p(1-\alpha\phi)(1-\phi)(1-\pi) - (np)^2(1-\pi)^2(1-\phi)^2](1-\alpha^2\phi)} \\ &\quad + \frac{n^2p(1-\phi)(1-\alpha\phi)^2(1-\pi)(1-p+np)}{[n^2p(1-\pi)(1-\phi)(1-\alpha\phi) - (np)^2(1-\pi)^2(1-\phi)^2](1-\alpha^2\phi)} \\ &\quad - \frac{n(np)^2(1-\pi)^2(1-\phi)^2}{n^2p(1-\pi)(1-\phi)(1-\alpha\phi) - (np)^2(1-\pi)^2(1-\phi)^2}. \end{aligned}$$

Also, the marginal distributions of the two special models can be obtained by solving the invariance equation  $\Pi' = \Pi'P'$ , where  $\Pi'$  and  $P'$  are the marginal distribution and transition matrix of the process.

**Acknowledgements** We gratefully acknowledge the anonymous reviewers for their serious work and thoughtful suggestions that have helped us improve this paper substantially. This work is supported by National Natural Science Foundation of China (No. 11731015, 11571051, 11501241, 11871028), Natural Science Foundation of Jilin Province (No. 20150520053JH, 20170101057JC, 20180101216JC), Program for Changbaishan Scholars of Jilin Province (2015010), and Science and Technology Program of Jilin Educational Department during the “13th Five-Year” Plan Period (No. 2016316).

## References

- Billingsley P (1961) Statistical inference for Markov processes. University of Chicago Press, Chicago
- Bourguignon M, Vasconcellos KL (2015) First-order non-negative integer valued autoregressive processes with power series innovations. *Braz J Probab Stat* 29:71–93
- Bourguignon M, Rodrigues J, Santos-Neto M (2019) Extended Poisson INAR(1) processes with equidispersion, underdispersion and overdispersion. *J Appl Stat* 46:101–118
- Cui Y, Lund R (2010) Inference in binomial AR(1) models. *Stat Probab Lett* 80:1985–1990
- Freeland RK, McCabe BPM (2004) Forecasting discrete valued low count time series. *Int J Forecast* 20:427–434

- Jazi MA, Jones G, Lai CD (2012) First-order integer valued AR processes with zero-inflated poisson innovations. *J Time Ser Anal* 33:954–963
- Khoo WC, Ong SH, Biswas A (2017) Modeling time series of counts with a new class of INAR(1) model. *Stat Pap* 58:393–416
- Kim HY, Weiß CH (2015) Goodness-of-fit tests for binomial AR(1) processes. *Statistics* 49:291–315
- Kim H, Lee S (2017) On first-order integer-valued autoregressive process with Katz family innovations. *J Stat Comput Simul* 87:546–562
- Li H, Yang K, Zhao S, Wang D (2018) First-order random coefficients integer-valued threshold autoregressive processes. *Adv Stat Anal (ASTA)* 102:305–331
- Maiti R, Biswas A (2017) Coherent forecasting for stationary time series of discrete data. *Adv Stat Anal (ASTA)* 99:337–365
- McKenzie E (1985) Some simple models for discrete variate time series. *J Am Water Resour Assoc (JAWRA)* 21:645–650
- Möller TA, Silva ME, Weiß CH, Scotto MG, Pereira I (2016) Self-exciting threshold binomial autoregressive processes. *Adv Stat Anal (ASTA)* 100:369–400
- Möller TA, Weiß CH, Kim HY, Sirchenko A (2018) Modeling zero inflation in count data time series with bounded support. *Methodol Comput Appl Probab* 20:589–609
- Pegram GGS (1980) An autoregressive model for multilag Markov chains. *J Appl Probab* 17:350–362
- Scotto MG, Weiß CH, Silva ME, Pereira I (2014) Bivariate binomial autoregressive models. *J Multivar Anal* 125:233–251
- Shirozhan M, Mohammadpour M (2018a) A new class of INAR(1) model for count time series. *J Stat Comput Simul* 7:1348–1368
- Shirozhan M, Mohammadpour M (2018b) An INAR(1) model based on the Pegram and thinning operators with serially dependent innovation. *Commun Stat Simul Comput*. <https://doi.org/10.1080/03610918.2018.1521975>
- Steutel FW, Van Harn K (1979) Discrete analogues of self-decomposability and stability. *Ann Probab* 7:893–899
- Weiß CH (2008) Thinning operations for modeling time series of counts—a survey. *Adv Stat Anal (ASTA)* 92:319–341
- Weiß CH (2009a) Jumps in binomial AR(1) processes. *Stat Probab Lett* 79:2012–2019
- Weiß CH (2009b) Monitoring correlated processes with binomial marginals. *J Appl Stat* 36:399–414
- Weiß CH (2010) The INARCH(1) model for overdispersed time series of counts. *Commun Stat Simul Comput* 39:1269–1291
- Weiß CH (2013) Binomial AR(1) processes: moments, cumulants, and estimation. *Statistics* 47:494–510
- Weiß CH, Kim H (2013) Parameter estimation for binomial AR(1) models with applications in finance and industry. *Stat Pap* 54:563–590
- Weiß CH, Pollett PK (2014) Binomial autoregressive processes with density-dependent thinning. *J Time Ser Anal* 3:115–132
- Yang K, Wang D, Li H (2018) Threshold autoregression analysis for finite range time series of counts with an application on measles data. *J Stat Comput Simul* 88:597–614

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