Miroslav M. Ristić*, Christian H. Weiß and Ana D. Janjić

A Binomial Integer-Valued ARCH Model

DOI 10.1515/ijb-2015-0051

Abstract: We present an integer-valued ARCH model which can be used for modeling time series of counts with under-, equi-, or overdispersion. The introduced model has a conditional binomial distribution, and it is shown to be strictly stationary and ergodic. The unknown parameters are estimated by three methods: conditional maximum likelihood, conditional least squares and maximum likelihood type penalty function estimation. The asymptotic distributions of the estimators are derived. A real application of the novel model to epidemic surveillance is briefly discussed. Finally, a generalization of the introduced model is considered by introducing an integer-valued GARCH model.

Keywords: binomial INARCH (p) model, binomial INGARCH (p, q) model, overdispersion, parameter estimation, epidemic surveillance

JEL Classification: 2000 MSC: 62M10

1 Introduction

In this paper, we introduce a model for integer-valued time series with finite range $\{0, 1, \dots, n\}$, where $n \in \mathbb{N}$ denotes the (known) upper limit. In recent years, integer-valued time series with such a finite range have been widely reported in diverse real-life applications, such as the monitoring of computer pools with n workstations [1, 2], the number of transactions of n companies [3], the number of metapopulations with n patches [4], etc. Integer-valued time series with binomial marginals have been studied by many authors and different approaches have been used to construct them. The first approach is based on the binomial thinning operator “ $\circ$ ” as introduced by Steutel and van Harn [5]. McKenzie [6] defined the binomial AR(1) model as

$$X_t = \alpha \circ X_{t-1} + \beta \circ (n - X_{t-1}), \quad t \geq 1,$$

where X_0 has the binomial $\text{Bin}(n, p)$ distribution, $\rho \in (\max(-p/(1-p), -(1-p)/p), 1)$, $\alpha = \beta + \rho$, $\beta = p(1-\rho)$ and $p \in (0, 1)$. All the counting series in “ $\alpha \circ$ ” and “ $\beta \circ$ ” are mutually independent sequences of independent Bernoulli distributed random variables with parameters α and β , respectively, and the counting series at time t are independent of the random variables $\{X_s\}$ for all $s < t$. The binomial AR(1) model is a stationary and ergodic Markov chain with binomial $\text{Bin}(n, p)$ marginal distribution. Its autocorrelation function is of the same form as the autocorrelation function of the usual AR(1) process, and it is given by $\rho(k) = \rho^k$, $k \geq 0$. More properties of the binomial AR(1) model and some estimation issues can be found in Cui and Lund [7], Weiß and Pollett [4] and Weiß and Kim [2, 3]. Weiß [1] extended the binomial AR(1) model to the high-order binomial AR(p) model and defined it as

$$X_t = \sum_{i=1}^p D_{t,i} \cdot [\alpha \circ_{t+i} X_{t-i} + \beta \circ_{t+i} (n - X_{t-i})] \equiv \sum_{i=1}^p D_{t,i} \cdot f_i(X_{t-i}),$$

where $\{\mathbf{D}_t = (D_{t,1}, \dots, D_{t,p})\}$ is a sequence of independent random vectors with the multinomial distribution $\text{MULT}(1, \phi_1, \dots, \phi_k)$, $\mathbf{D}_t$ are independent of X_s and $f_i(X_s)$ for all $s < t$ and $i = 1, 2, \dots, p$, and the conditional probabilities $P(f_1(X_t) = i_1, \dots, f_p(X_t) = i_p | X_t = x_t, \mathcal{H}_{t-1})$ and $P(f_1(X_t) = i_1, \dots, f_p(X_t) = i_p | X_t = x_t)$ are equal, where

*Corresponding author: Miroslav M. Ristić, Department of Mathematics, University of Niš, Niš, Serbia,

E-mail: miristic72@gmail.com

Christian H. Weiß, Department of Mathematics and Statistics, Helmut Schmidt University Hamburg, Hamburg, Germany,

E-mail: weissc@hsu-hh.de

Ana D. Janjić, Department of Mathematics, University of Niš, Niš, Serbia

$\mathcal{H}_{t-1}$ is the process history of all random variables X_s and $f_j(X_s)$, $j=1, 2, \dots, p$. A bivariate extension of the binomial AR(1) model can be found in Scotto et al. [8].

The second approach is based on the use of the hypergeometric thinning operator as introduced by Al-Osh and Alzaid [9]. Al-Osh and Alzaid [9] defined a binomial AR(1) model as

$$X_t = S(X_{t-1}) + \varepsilon_t, \quad t \geq 1,$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables, independent of the initial state X_0 , and the random variable $S(X)$ for given $X=x$ has the hypergeometric distribution with parameters n , x and m , i.e.

$$P(S(X) = k | X = x) = \frac{\binom{x}{k} \binom{n-x}{m-k}}{\binom{n}{m}}, \quad \max(0, m-n+x) \leq k \leq \min(x, m).$$

From the definition of the model, it follows that the random variable $S(X)$ has the binomial distribution with parameters m and p , and the random variable ε_t has the binomial distribution with parameters $n-m$ and p . Al-Osh and Alzaid [9] showed that their binomial AR(1) model is a stationary Markov chain with binomial $\text{Bin}(n, p)$ marginal distribution and autocorrelation function given by $\rho(k) = (m/n)^k$, $k \geq 0$.

A third approach for count data time series with a finite range $\{0, 1, \dots, n\}$ is related to the so-called INGARCH models Ferland et al. [10], i.e. the integer-valued GARCH models. Weiß and Pollett [11] introduced the INARCH(1) model with binomial marginals as a boundary case of the binomial AR(1) processes with density dependent thinning. They considered an integer-valued time series model $\{X_t\}_{t \in \mathbb{Z}}$ such that $X_t | \mathcal{F}_{t-1} : \text{Bin}(n, \alpha_t)$, $t \in \mathbb{Z}$, where $\mathcal{F}_{t-1}$ is the σ -field generated by the random variables $\{X_{t-k}\}_{k \geq 1}$, and α_t is generated as $\alpha_t = a_0 + \frac{a_1}{n} X_{t-1}$, $t \in \mathbb{Z}$, where $a_0 > 0$ and $a_1 \geq 0$. This model is referred to as the binomial INARCH(1) model. Weiß and Pollett [11] derived some properties of the binomial INARCH(1) model and compared it with the (infinite-range) Poisson INARCH(1) model introduced by Ferland et al. [10]. Such INGARCH models were further investigated and generalized by several authors including Zhu [12, 13]; Xu et al. [14]; Gonçalves et al. [15]. In contrast to the binomial INARCH(1) model, however, all these models are designed for processes with the infinite range $\mathbb{N}_0$.

In this paper, we follow the third approach and extend the binomial INARCH(1) model to the binomial INARCH(p) model. In Section 2, we introduce the binomial INARCH(p) model, and we prove strict stationarity and ergodicity of this model. In Section 3, we estimate the unknown parameters by three different estimation methods: conditional maximum likelihood, conditional least squares and maximum likelihood type penalty function estimation. The strong consistency and asymptotic normality of the obtained estimators are derived and discussed. In Section 4, we provide some simulation results to check the performance of the three estimation methods. Section 5 demonstrates that our new model is particularly beneficial in biostatistics. There, we discuss a possible application of the introduced model to a real data set from the field of epidemiology. It might be utilized for epidemic surveillance systems, where approaches based on (infinite-range) count data time series already have been successfully applied [16, 17]. Finally, in Section 6, we introduce the full binomial INGARCH(p, q) model as a generalization of the binomial INARCH(p) model.

2 The binomial INARCH model

In this section, we extend the binomial INARCH(1) model and consider an integer-valued time series model for $\{X_t\}_{t \in \mathbb{Z}}$ given as

$$X_t | \mathcal{F}_{t-1} : \text{Bin}(n, \alpha_t), \quad \forall t \in \mathbb{Z}, \quad (1)$$

where $\mathcal{F}_{t-1}$ is the σ -field generated by the random variables $\{X_{t-k}\}_{k \geq 1}$, n is a positive integer and α_t is generated as

$$\alpha_t = a_0 + \frac{1}{n} \sum_{i=1}^p a_i X_{t-i}, \quad \forall t \in \mathbb{Z}, \quad (2)$$

where $a_0 > 0$, $a_i \geq 0$, for $i = 1, 2, \dots, p$, and $p \in \{1, 2, \dots\}$. We suppose that the parameters a_i , $i = 0, 1, \dots, p$, satisfy the inequality $a_0 + \sum_{i=1}^p a_i < 1$, which implies that $\{\alpha_t\}$ given by (2) is well-defined, i.e. that α_t belongs to the interval $(0, 1)$ for all $t \in \mathbb{Z}$. We will say that the time series $\{X_t\}_{t \in \mathbb{Z}}$ given by (1) and (2) is the *binomial integer-valued ARCH model* and we will denote it as $\text{BINARCH}(p)$.

Under the assumption (1), the conditional probability of the random variable X_t for given $\mathcal{F}_{t-1}$ equals

$$P(X_t = x | \mathcal{F}_{t-1}) = \binom{n}{x} \alpha_t^x (1 - \alpha_t)^{n-x}, \quad x = 0, 1, 2, \dots, n.$$

Then the conditional mean and conditional variance are given as $\mathbf{E}(X_t | \mathcal{F}_{t-1}) = n\alpha_t$ and $\text{Var}(X_t | \mathcal{F}_{t-1}) = n\alpha_t(1 - \alpha_t)$. Thus, we have that $\text{Var}(X_t | \mathcal{F}_{t-1}) < \mathbf{E}(X_t | \mathcal{F}_{t-1})$. Let us now consider the unconditional mean and unconditional variance. The unconditional mean of the random variable X_t is given by $\mathbf{E}(X_t) = n\mathbf{E}(\alpha_t)$. On the other hand, the unconditional variance of the random variable X_t is given by

$$\begin{aligned} \text{Var}(X_t) &= \mathbf{E}[\text{Var}(X_t | \mathcal{F}_{t-1})] + \text{Var}[\mathbf{E}(X_t | \mathcal{F}_{t-1})] = n\mathbf{E}(\alpha_t) - n\mathbf{E}(\alpha_t^2) + n^2 \text{Var}(\alpha_t) \\ &= n(n-1) \text{Var}(\alpha_t) + n\mathbf{E}(\alpha_t)(1 - \mathbf{E}(\alpha_t)). \end{aligned} \quad (3)$$

We shall further investigate the dispersion behaviour of the $\text{BINARCH}(p)$ in Example 1.

In the next theorem, we derive some properties of the model given by (1) and (2).

Theorem 1 *The $\text{BINARCH}(p)$ process given by (1) and (2) is an ergodic, strict and second-order stationary process.*

The proof of Theorem 1 is provided by Appendix A.1.

Remark 1. If $\{X_t\}_{t \in \mathbb{Z}}$ is the $\text{BINARCH}(p)$ process given by (1) and (2), then

$$\mu \equiv \mathbf{E}(X_t) = \frac{na_0}{1 - \sum_{i=1}^p a_i}.$$

In the rest of this section, we will derive and discuss the autocovariance structure of the $\text{BINARCH}(p)$ model given by (1) and (2).

Theorem 2 *Let $\{X_t\}_{t \in \mathbb{Z}}$ be the $\text{BINARCH}(p)$ process given by (1) and (2). The autocovariance function $\gamma_X(k) = \text{Cov}(X_t, X_{t-k})$, $k \geq 0$, satisfies the equations*

$$\begin{aligned} \gamma_X(0) &= \mu(1 - \mu/n) + (1 - 1/n) \sum_{i=1}^p a_i \gamma_X(i), \\ \gamma_X(k) &= \sum_{i=1}^p a_i \gamma_X(|k-i|), \quad k \geq 1. \end{aligned}$$

The proof of Theorem 2 is provided in Appendix A.2.

Thus, we can see that the autocovariances of the $\text{BINARCH}(p)$ model form equations similar to the Yule-Walker equations of the standard $\text{AR}(p)$ model. Also, from the above theorem, we obtain that the autocorrelation function $\rho_X(k) \equiv \text{Corr}(X_t, X_{t-k})$ satisfies the equations $\rho_X(k) = \sum_{i=1}^p a_i \rho_X(|k-i|)$, $k \geq 1$, and is

Figure 1: Examples of autocorrelation functions of BINARCH(p) model for different values of parameters a_i and p .

independent of n and a_0 . In Figure 1, we present some examples of the autocorrelation function of the BINARCH(p) model, with different values of the parameters a_i , $i = 1, 2, \dots, p$, and $p \in \{2, 3, 4\}$.

Example 1 Let us consider the special case $p = 1$ in some more detail, i.e. the BINARCH(1) model as in Weiß and Pollett [11]. From Remark 1, we have that $\mu = na_0/(1 - a_1)$, while Theorem 2 implies that the autocovariance structure of the time series $\{X_t\}_{t \in \mathbb{Z}}$ is given by $\gamma_X(k) = a_1^k \gamma_X(0)$ for $k \geq 0$ (exponentially decaying), where

$$\gamma_X(0) = \frac{n^2 a_0 (1 - a_0 - a_1)}{(1 - a_1)^2 [a_1^2 + n(1 - a_1^2)]}.$$

To further investigate the dispersion behaviour of the BINARCH(1) model, let us define the Poisson index of dispersion by $I_{\text{Pois}} := \text{Var}(X_t)/\mathbf{E}(X_t)$. Then we say that we have underdispersion if $I_{\text{Pois}} < 1$, equidispersion if $I_{\text{Pois}} = 1$, and overdispersion if $I_{\text{Pois}} > 1$ (all with respect to a Poisson distribution). For the BINARCH(1) model, we have that

$$I_{\text{Pois}} = \frac{n(1 - a_0 - a_1)}{(1 - a_1)[a_1^2 + n(1 - a_1^2)]}.$$

Thus, when $a_0 = \frac{n-1}{n} a_1^2 (1 - a_1)$, we obtain equidispersion, when $a_0 < \frac{n-1}{n} a_1^2 (1 - a_1)$, we obtain overdispersion, and when $a_0 > \frac{n-1}{n} a_1^2 (1 - a_1)$, we have underdispersion. The attainable range of I_{Pois} is bounded by 0 (if $a_0 + a_1$ is close to 1) and by

$$I_{\text{Pois}} < \frac{n}{a_1^2 + n(1 - a_1^2)} < n.$$

Instead of considering over- and underdispersion with respect to a Poisson distribution as before, we might also compare the unconditional variance-mean behavior with that of a binomial distribution with population parameter n . For this purpose, let us investigate the so-called binomial index of dispersion, defined by

$$I_{\text{Bin}} := I_{\text{Bin}}(n, \mu, \sigma^2) := \frac{\sigma^2}{\mu(1 - \mu/n)}$$

for a random variable X with range $\{0, \dots, n\}$, mean μ and variance σ^2 . For the case of the BINARCH(p) model with $\mu := \mathbf{E}(X_t)$ and $\sigma^2 := \text{Var}(X_t)$, it follows from eq. (3) that

$$I_{\text{Bin}} = \frac{n(n-1)\text{Var}(\alpha_t) + \mu(1-\mu/n)}{\mu(1-\mu/n)} = 1 + \frac{n(n-1)\text{Var}(\alpha_t)}{\mu(1-\mu/n)} = 1 + \frac{(1-1/n) \sum_{i=1}^p a_i \gamma_X(i)}{\mu(1-\mu/n)},$$

i.e. for any BINARCH(p) model, we have $I_{\text{Bin}} > 1$. While the conditional distribution of $X_t|F_{t-1}$ satisfies $I_{\text{Bin}} = 1$ (conditional binomial distribution), the unconditional distribution always exhibits extra-binomial variation (overdispersion with respect to a binomial distribution).

For the example of the BINARCH(1) model, also see Weiß and Pollett [11], we have

$$I_{\text{Bin}} = 1 + \frac{(n-1)a_1^2}{n(1-a_1^2) + a_1^2}.$$

3 Estimation of parameters

In this section, we consider the estimation of the unknown parameters $\boldsymbol{\theta} = (a_0, a_1, \dots, a_p)^\top$ of the BINARCH(p) model, while we suppose that the parameter n of the conditional binomial distribution is known. We consider three estimation methods: conditional maximum likelihood estimation, conditional least squares estimation and maximum likelihood type penalty function estimation from Tjøstheim [18]. The second and the third estimation approach are based on the minimization of an objective function, so both approaches can be understood as penalty function approaches [18]. The penalty function for the conditional least squares estimation as given by (5) accumulates squared deviations, while the penalty function (7) for the maximum likelihood type penalty function estimation is motivated by the Gaussian log-likelihood function. Also, these penalty functions are chosen in a way to provide some asymptotic properties of the resulting estimators.

We suppose that $X_1, X_2, \dots, X_N$ are the observations generated by the BINARCH(p) process $\{X_t\}_{t \in \mathbb{Z}}$, where $N \in \mathbb{N}$ represents the size of the sample. Here, the parameter n (upper limit of the range) is considered as a known quantity.

3.1 Conditional maximum likelihood estimation

From the definition of the BINARCH(p) model, we obtain that the conditional log-likelihood function is given by

$$l(\boldsymbol{\theta}) = \sum_{t=p+1}^N \left\{ \log \binom{n}{X_t} + X_t \log \alpha_t + (n - X_t) \log(1 - \alpha_t) \right\}. \quad (4)$$

Since the BINARCH(p) process $\{X_t\}_{t \in \mathbb{Z}}$ can be represented equivalently as the finite Markov chain $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ (see Section 2), it is possible to apply the results in Billingsley [19] to investigate the properties of the conditional maximum likelihood (CML) estimators of the parameter vector $\boldsymbol{\theta} := (a_0, a_1, \dots, a_p)^\top$.

Theorem 3 There exists a consistent CML estimator of $\boldsymbol{\theta}$, maximizing (4), that is also asymptotically normally distributed.

The proof of Theorem 3 is provided by Appendix A.3.

3.2 Conditional least squares estimation

While the CML approach discussed in the previous section makes use of the complete conditional distribution, we shall now derive the semiparametric conditional least squares (CLS) estimators of the BINARCH(p)

model $\{X_t\}_{t \in \mathbb{Z}}$ and discuss their asymptotic properties. Let $\boldsymbol{\theta} := (a_0, a_1, \dots, a_p)^\top$ be the vector of the unknown parameters, and let $\mathbf{Z}_t := (1, X_{t-1}, \dots, X_{t-p})^\top$ and $\mathbf{W}_t := (n, X_{t-1}, \dots, X_{t-p})^\top$ be the vectors of the observations. The CLS estimates of the vector $\boldsymbol{\theta} := (a_0, a_1, \dots, a_p)^\top$ are obtained by minimizing the function

$$S(\boldsymbol{\theta}) = \sum_{t=p+1}^N (X_t - \mathbf{W}_t^\top \boldsymbol{\theta})^2 = \sum_{t=p+1}^N \left(X_t - na_0 - \sum_{i=1}^p a_i X_{t-i} \right)^2 \quad (5)$$

with respect to the vector $\boldsymbol{\theta}$, and they are given as

$$\hat{\boldsymbol{\theta}}^{\text{CLS}} = \left(\sum_{t=p+1}^N \mathbf{Z}_t \mathbf{W}_t^\top \right)^{-1} \sum_{t=p+1}^N X_t \mathbf{Z}_t. \quad (6)$$

Now we will derive the asymptotic properties of the CLS estimators $\hat{\boldsymbol{\theta}}^{\text{CLS}}$. First, we start with the consistency of these estimators.

Theorem 4 The CLS estimators $\hat{\boldsymbol{\theta}}^{\text{CLS}}$ given by (6) are strongly consistent estimators of the unknown parameter $\boldsymbol{\theta}$.

The proof of Theorem 4 is provided by Appendix A.4.

The asymptotic distribution of the CLS estimators $\hat{\boldsymbol{\theta}}^{\text{CLS}}$ given by (6) follows from the following theorem, the proof of which is provided by Appendix A.5.

Theorem 5 If the CLS estimators $\hat{\boldsymbol{\theta}}^{\text{CLS}}$ are given by (6), then

$$\sqrt{N}(\hat{\boldsymbol{\theta}}^{\text{CLS}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \mathbf{U}^{-1} \mathbf{R} \mathbf{U}^{-1}), \quad N \rightarrow \infty,$$

where $\mathbf{U} = \mathbf{E}(\mathbf{W}_t \mathbf{W}_t^\top)$ and $\mathbf{R} = n \mathbf{E}(\alpha_t (1 - \alpha_t) \mathbf{W}_t \mathbf{W}_t^\top)$.

Closed-forms expressions for the matrices $\mathbf{U}$ and $\mathbf{R}$ for higher order p are very cumbersome. Because of that, Appendix B provides closed-form expressions for the matrices $\mathbf{U}$ and $\mathbf{R}$ for the special case of the BINARCH(1) model, which may be used, in turn, to derive approximate standard errors of the CLS estimates.

3.3 Maximum likelihood type penalty function

This subsection is dedicated again to maximum likelihood estimators of the BINARCH(p) model. Instead of the standard approach to CML estimators, we shall use the maximum likelihood type penalty (MLTP) function (T1986) of the observed model given by

$$L(\boldsymbol{\theta}) = \sum_{t=p+1}^N \left\{ \log(n) + \log(\alpha_t) + \log(1 - \alpha_t) + \frac{(X_t - n\alpha_t)^2}{n\alpha_t(1 - \alpha_t)} \right\} = \sum_{t=p+1}^N \phi_t. \quad (7)$$

Tjøstheim [18] gave two motivations for using a penalty term in the conditional log-likelihood function. First, in the case of a conditional Gaussian process, the maximum likelihood penalty function $L(\boldsymbol{\theta})$ coincides with the conditional log-likelihood function of this process except a multiplicative constant. The second motivation is that ϕ_t has the marginal property, which can be used to derive asymptotic properties of the considered estimators. Differentiating the function $L(\boldsymbol{\theta})$ with respect to the parameters a_i , $i = 0, 1, \dots, p$, we obtain the estimators of the unknown parameters as the solutions of the nonlinear system of the equations

$$\sum_{t=p+1}^N \frac{\partial \phi_t}{\partial a_i} = \sum_{t=p+1}^N \frac{-X_t^2 + n\alpha_t + 2\alpha_t X_t^2 - 3n\alpha_t^2 + n^2\alpha_t^2 - 2n\alpha_t^2 X_t + 2n\alpha_t^3}{n\alpha_t^2(1 - \alpha_t)^2} \cdot \frac{\partial \alpha_t}{\partial a_i} = 0,$$

for $i = 0, 1, 2, \dots, p$. Now we will focus on deriving consistency and asymptotic properties of the estimators obtained by using above function. The notation introduced in previous section is retained.

Theorem 6 *The MLTP estimators obtained by minimizing (7) are strongly consistent estimators of unknown parameter θ .*

The proof of Theorem 6 is provided by Appendix A.6.

Finally, the asymptotic normality of the estimators obtained by minimizing (7) is established by the following theorem, the proof of which is provided by Appendix A.7.

Theorem 7 *If $\{\hat{\theta}^{\text{MLTP}}\}$ are the MLTP estimators obtained by minimizing (7), then*

$$\sqrt{N}(\hat{\theta}^{\text{MLTP}} - \theta) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}^{-1} + \mathbf{V}^{-1}\mathbf{S}\mathbf{V}^{-1}), \quad N \rightarrow \infty,$$

where

$$\mathbf{V} = \frac{1}{2n^2} \mathbf{E} \left(\frac{1 + 2(n-2)\alpha_t - 2(n-2)\alpha_t^2}{\alpha_t^2(1-\alpha_t)^2} \mathbf{W}_t \mathbf{W}_t^\top \right)$$

and

$$\mathbf{S} = \frac{1}{4n^4} \mathbf{E} \left(\left(\frac{-X_t^2 + n\alpha_t + 2\alpha_t X_t^2 - 3n\alpha_t^2 + n^2\alpha_t^2 - 2n\alpha_t^2 X_t + 2n\alpha_t^3}{\alpha_t^2(1-\alpha_t)^2} \right)^2 \mathbf{W}_t \mathbf{W}_t^\top \right) - \mathbf{V}.$$

4 Simulation study

In this section, we provide some results from a simulation study to check the finite-sample performance of the three estimation methods considered in the previous section. We simulated samples of size 500, and the number of replications is $m = 10,000$. We perform the estimation for the subsamples consisting of the first 50, 100, 200 and 500 elements, thus considering subsamples of four different sizes $N = 50$, $N = 100$, $N = 200$, and $N = 500$. We provide the estimation of the parameters of BINARCH(p) models for $p = 1$ and $p = 2$. For the case $p = 1$, we conduct simulations for the following cases: (1) $(n, a_0, a_1) = (5, 0.8, 0.1)$; (2) $(n, a_0, a_1) = (5, 0.6, 0.1)$; (3) $(n, a_0, a_1) = (5, 0.3, 0.6)$; (4) $(n, a_0, a_1) = (5, 0.1, 0.8)$; (5) $(n, a_0, a_1) = (5, 0.05, 0.9)$. We also consider four cases for $p = 2$: (6) $(n, a_0, a_1, a_2) = (5, 0.6, 0.2, 0.1)$; (7) $(n, a_0, a_1, a_2) = (5, 0.3, 0.5, 0.1)$; (8) $(n, a_0, a_1, a_2) = (5, 0.1, 0.1, 0.75)$; and (9) $(n, a_0, a_1, a_2) = (5, 0.1, 0.75, 0.1)$. For each choice of the true values of the parameters, we compute the estimates according to the three considered estimation methods for each generated subsample. The mean of the estimates and the mean absolute deviation errors provided within parentheses are given in Tables 1 and 2. From these tables, we can conclude that all three estimation methods give good estimates that quickly approach their true values as the size of the sample increases. Also, we can see that the mean absolute deviations are small and they decrease as the size of the sample increases.

In the case of the BINARCH(1) model, the best results are usually obtained by CML estimation. If a_1 is large and a_0 is small (cases 4 and 5), then MLTP estimation provides the best results for a_1 for samples of small size, while otherwise, the best results are provided by CML estimation.

In the case of the BINARCH(2) model, CML and CLS estimation give the best results, while MLTP estimation was superior only in one case. The CLS estimation is very sensitive to small true values of the parameters. In this case, if some CLS estimates are smaller than 0, we set these estimates to take the value 0. For this method of estimation, it is also possible that the sum of the estimates is equal or greater than 1. On the other hand, many statistical packages allow to minimize and maximize functions with boundary

Table 1: Mean of estimates and mean absolute deviation errors in parentheses for BINARCH(1) model.

Model	N	a_0^{CLS}	a_1^{CLS}	a_0^{CML}	a_1^{CML}	a_0^{MLTP}	a_1^{MLTP}
(1)	50	0.8263(0.1076)	0.0710(0.1178)	0.8009(0.0851)	0.1016(0.0903)	0.7935(0.0868)	0.1067(0.0972)
	100	0.8137(0.0758)	0.0848(0.0833)	0.8039(0.0669)	0.0973(0.0719)	0.7995(0.0695)	0.1002(0.0777)
	200	0.8073(0.0541)	0.0921(0.0591)	0.8044(0.0514)	0.0959(0.0556)	0.8025(0.0545)	0.0973(0.0602)
	500	0.8026(0.0343)	0.0972(0.0376)	0.8023(0.0340)	0.0976(0.0372)	0.8022(0.0369)	0.0979(0.0405)
(2)	50	0.6176(0.0807)	0.0741(0.1147)	0.5965(0.0598)	0.1014(0.0878)	0.5983(0.0648)	0.1047(0.0911)
	100	0.6089(0.0560)	0.0867(0.0797)	0.5996(0.0468)	0.0973(0.0692)	0.6004(0.0490)	0.0989(0.0708)
	200	0.6043(0.0401)	0.0934(0.0569)	0.6012(0.0371)	0.0965(0.0538)	0.6012(0.0378)	0.0972(0.0546)
	500	0.6018(0.0251)	0.0974(0.0356)	0.6015(0.0249)	0.0976(0.0353)	0.6014(0.0252)	0.0979(0.0359)
(3)	50	0.3463(0.0898)	0.5412(0.1076)	0.3432(0.0867)	0.5461(0.1031)	0.3370(0.0891)	0.5570(0.1060)
	100	0.3230(0.0594)	0.5711(0.0703)	0.3213(0.0576)	0.5739(0.0674)	0.3178(0.0595)	0.5799(0.0703)
	200	0.3111(0.0410)	0.5861(0.0485)	0.3103(0.0400)	0.5875(0.0467)	0.3085(0.0415)	0.5904(0.0491)
	500	0.3046(0.0255)	0.5941(0.0304)	0.3043(0.0247)	0.5947(0.0290)	0.3036(0.0256)	0.5959(0.0304)
(4)	50	0.1303(0.0494)	0.7396(0.0849)	0.1302(0.0485)	0.7401(0.0825)	0.1229(0.0486)	0.7553(0.0797)
	100	0.1135(0.0293)	0.7731(0.0490)	0.1131(0.0281)	0.7737(0.0468)	0.1083(0.0291)	0.7833(0.0473)
	200	0.1064(0.0193)	0.7877(0.0313)	0.1062(0.0183)	0.7880(0.0296)	0.1035(0.0195)	0.7932(0.0310)
	500	0.1025(0.0116)	0.7952(0.0184)	0.1025(0.0109)	0.7954(0.0173)	0.1013(0.0119)	0.7976(0.0187)
(5)	50	0.0922(0.0540)	0.8151(0.0976)	0.0907(0.0524)	0.8191(0.0923)	0.0886(0.0531)	0.8265(0.0884)
	100	0.0670(0.0272)	0.8663(0.0454)	0.0653(0.0249)	0.8697(0.0416)	0.0621(0.0249)	0.8765(0.0398)
	200	0.0571(0.0156)	0.8860(0.0243)	0.0556(0.0134)	0.8888(0.0212)	0.0535(0.0136)	0.8928(0.0211)
	500	0.0526(0.0089)	0.8950(0.0131)	0.0519(0.0074)	0.8964(0.0112)	0.0511(0.0080)	0.8979(0.0117)

conditions, which imply that the other two estimation methods give estimates that never take values outside $(0, 1)$, and that their sum is always less than 1.

The simplest method for use is the CLS estimation method, and the corresponding estimates are used as the starting values for other two methods. An interesting conclusion from our simulations is that, as expected, the efficiency of the CML estimates with respect to the CLS estimates is high.

5 Real-data example

In this section, we present a possible application of the novel BINARCH(p) model in the field of biostatistics. As a real-data set, we consider the infection counts as previously discussed by Weiß and Pollett [11]. This data set, taken from the “SurvStat” data base of the Robert-Koch-Institut [20], contains the number of districts with new cases of hantavirus infections per week in the year 2011 ($N = 52$ counts) reported in $n = 38$ Germany’s districts, i.e. the counts express the regional spread of the hantavirus infections. More than 200,000 infections by hantavirus are reported from all over the world per year. For most European countries, rising numbers of cases are reported, and in Germany, hantaviral infections meanwhile became the most common endemic rodent-borne human illness [21, 22]. The hantavirus mainly causes two diseases: The hantavirus cardiopulmonary syndrome (HCPS) with case-fatality rates $> 35\%$ is mainly reported from North, Central, and South America, while in Europe (and Asia), the less severe haemorrhagic fever with renal syndrome (HFRS; case-fatality rates $> 10\%$) is usually observed [21, 23]. Among the viral species causing HFRS in Central Europe are the Puumala virus (carried by the red bank vole) and the Dobrava virus (carried by the striped field mouse), see Schilling et al. [24]; Heyman et al. [22] for further details.

Let us return to the particular infection counts for 2011 in Germany. The minimum number of districts with new cases is 0 and the maximum number is 11. The sample mean is 4.173 and the sample variance is 7.793. As a result, the binomial index of dispersion is 2.098, which indicates overdispersion with respect to a binomial distribution (extra-binomial variation), also see Example 1. Therefore, it is plausible that Weiß and Pollett [11] found the binomial INARCH(1) model to be superior to the binomial AR(1) model. The

Table 2: Mean of estimates and mean absolute deviation errors in parentheses for BINARCH(2) model.

Model	N	σ_0^{CLS}	σ_1^{CLS}	σ_2^{CLS}	σ_0^{CML}	σ_1^{CML}	σ_2^{CML}	σ_0^{MLTP}	σ_1^{MLTP}	σ_2^{MLTP}
(6)	50	0.6700(0.1456)	0.1706(0.1290)	0.0490(0.1246)	0.5905(0.0899)	0.2092(0.1151)	0.1022(0.0956)	0.5843(0.1007)	0.2114(0.1216)	0.1077(0.1016)
	100	0.6368(0.0998)	0.1845(0.0886)	0.0733(0.0862)	0.6007(0.0744)	0.2055(0.0883)	0.0933(0.0742)	0.5972(0.0820)	0.2059(0.0931)	0.0977(0.0793)
	200	0.6199(0.0686)	0.1918(0.0620)	0.0854(0.0601)	0.6055(0.0576)	0.2024(0.0640)	0.0911(0.0554)	0.6036(0.0628)	0.2027(0.0676)	0.0932(0.0594)
	500	0.6082(0.0428)	0.1970(0.0387)	0.0937(0.0380)	0.6053(0.0399)	0.1996(0.0392)	0.0943(0.0373)	0.6046(0.0432)	0.2003(0.0420)	0.0946(0.0406)
(7)	50	0.3676(0.1102)	0.4643(0.1267)	0.0489(0.1230)	0.3554(0.1001)	0.4449(0.1142)	0.0856(0.0864)	0.3476(0.1039)	0.4551(0.1199)	0.0901(0.0923)
	100	0.3323(0.0709)	0.4846(0.0877)	0.0739(0.0862)	0.3295(0.0695)	0.4755(0.0813)	0.0887(0.0703)	0.3249(0.0718)	0.4805(0.0856)	0.0912(0.0749)
	200	0.3167(0.0483)	0.4917(0.0612)	0.0869(0.0598)	0.3161(0.0477)	0.4890(0.0581)	0.0914(0.0540)	0.3135(0.0495)	0.4920(0.0611)	0.0925(0.0574)
	500	0.3061(0.0296)	0.4972(0.0376)	0.0949(0.0380)	0.3060(0.0289)	0.4971(0.0365)	0.0954(0.0364)	0.3052(0.0301)	0.4986(0.0384)	0.0953(0.0389)
(8)	50	0.2015(0.1160)	0.0549(0.1096)	0.6396(0.1270)	0.2101(0.1218)	0.0826(0.0805)	0.6087(0.1524)	0.2108(0.1260)	0.0840(0.0859)	0.6163(0.1500)
	100	0.1495(0.0652)	0.0824(0.0672)	0.6957(0.0721)	0.1600(0.0735)	0.0877(0.0576)	0.6832(0.0819)	0.1562(0.0745)	0.0875(0.0626)	0.6914(0.0813)
	200	0.1240(0.0387)	0.0923(0.0440)	0.7237(0.0440)	0.1279(0.0412)	0.0929(0.0398)	0.7211(0.0456)	0.1250(0.0425)	0.0926(0.0437)	0.7269(0.0465)
	500	0.1087(0.0210)	0.0976(0.0262)	0.7402(0.0244)	0.1087(0.0199)	0.0977(0.0238)	0.7409(0.0234)	0.1073(0.0212)	0.0975(0.0263)	0.7436(0.0248)
(9)	50	0.1691(0.0883)	0.7036(0.1402)	0.0493(0.1310)	0.1649(0.0833)	0.6724(0.1228)	0.0895(0.0906)	0.1575(0.0830)	0.6854(0.1260)	0.0892(0.0963)
	100	0.1316(0.0494)	0.7312(0.0910)	0.0754(0.0907)	0.1304(0.0474)	0.7206(0.0776)	0.0892(0.0711)	0.1248(0.0482)	0.7313(0.0826)	0.0881(0.0776)
	200	0.1145(0.0294)	0.7417(0.0617)	0.0884(0.0630)	0.1147(0.0286)	0.7388(0.0545)	0.0921(0.0551)	0.1115(0.0302)	0.7454(0.0599)	0.0908(0.0611)
	500	0.1058(0.0169)	0.7466(0.0382)	0.0955(0.0398)	0.1058(0.0161)	0.7466(0.0355)	0.0959(0.0364)	0.1043(0.0173)	0.7499(0.0395)	0.0949(0.0412)

Figure 2: Infection counts: (a) autocorrelation function; (b) partial autocorrelation function.

empirical autocorrelation and partial autocorrelation function (ACF and PACF, respectively) are given in Figure 2. Inspecting the plot of the PACF in more detail, however, an autoregressive model of order >1 appears to be reasonable. Hence, we shall now apply the novel BINARCH(p) model to the data, and we shall also compare it to the binomial AR(p) model (BINAR(p)) as introduced in Weiß [1], see Section 1 for further details.

First, we need to determine the order p . The plot of the PACF indicates that p may take values ≤ 3 . According to this, we consider models up to order 4 in our study. Next, we estimate the unknown parameters of both types of models by a maximum likelihood method. Finally, we compare the performance of these models. Often, one of the information criteria AIC or BIC is applied in this context. But since we have to use a *conditional* ML approach, where the number of terms involved in the log-likelihood function (4), namely $N - p$, decreases with increasing p , these information criteria are misleading in the present situation. Therefore, we shall evaluate the performance by comparing the root mean square errors of the models. Let $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_p$ represent the CML estimates of the parameters of the BINARCH(p) model. Then the root mean square error of the BINARCH(p) model is given by

$$RMS = \sqrt{\frac{1}{N-p} \sum_{t=p+1}^N \left(X_t - n\hat{a}_0 - \sum_{i=1}^p \hat{a}_i X_{t-i} \right)^2}.$$

On the other hand, let $\hat{\pi}, \hat{\rho}, \hat{\phi}_1, \dots, \hat{\phi}_p$ represent the CML estimates of the parameters of the BINAR(p) model. The estimates $\hat{\phi}_i, i = 1, 2, \dots, p$ satisfy the condition $\sum_{i=1}^p \hat{\phi}_i = 1$. Because of that, we only estimate the parameters $\phi_1, \phi_2, \dots, \phi_{p-1}$, while the estimate of the last parameter ϕ_p is obtained as $\hat{\phi}_p = 1 - \sum_{i=1}^{p-1} \hat{\phi}_i$. This implies that the estimated standard error for $\hat{\phi}_p$ will always be equal to 0. Now, the root mean square error of the BINAR(p) model is given by

$$RMS = \sqrt{\frac{1}{N-p} \sum_{t=p+1}^N \left(X_t - n\hat{\pi}(1-\hat{\rho}) - \hat{\rho} \sum_{i=1}^p \hat{\phi}_i X_{t-i} \right)^2}.$$

In Table 3, we give the CML estimates of the unknown parameters with the estimated standard errors in parentheses and the respective root mean square errors. Also, we give the negative log-likelihood functions for each fitted model. It becomes clear that the BINARCH models are always superior to their BINAR

Table 3: Maximum likelihood estimates of the parameters of the BINARCH(p) and BINAR(p) models with the corresponding root mean square errors and negative log-likelihood functions.

Model	Estimates					-log L	RMS
BINAR(1)	$\hat{\pi}$	$\hat{\rho}$	$\hat{\phi}_1$				
	0.1151	0.5353	1.0000			109.3891	2.1226
	(.01337)	(0.0707)	(0.0000)				
BINARCH(1)	$\hat{a}_0$	$\hat{a}_1$					
	0.0303	0.7476				103.6976	0.0981
	(.0111)	(0.1085)					
BINAR(2)	$\hat{\pi}$	$\hat{\rho}$	$\hat{\phi}_1$	$\hat{\phi}_2$			
	0.1176	0.6789	0.4283	0.5717		101.3702	1.9140
	(.0193)	(0.0688)	(0.1690)	(0.0000)			
BINARCH(2)	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$				
	0.0142	0.4685	0.4333			95.9150	1.8653
	(.0109)	(0.1396)	(0.1405)				
BINAR(3)	$\hat{\pi}$	$\hat{\rho}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$		
	0.1226	0.7556	0.2841	0.4478	0.2681	96.1284	1.8208
	(.0258)	(0.0639)	(0.1644)	(0.1634)	(0.0000)		
BINARCH(3)	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$			
	0.0110	0.3500	0.3216	0.2791		92.3401	1.7697
	(0.0108)	(0.1591)	(0.1506)	(0.1534)			
BINAR(4)	$\hat{\pi}$	$\hat{\rho}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	
	0.1235	0.7622	0.2909	0.4159	0.2594	0.0338	94.4509
	(0.0269)	(0.0657)	(0.1672)	(0.1896)	(0.1552)	(0.0000)	1.8359
BINARCH(4)	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$		
	0.0105	0.3599	0.3045	0.2941	0.000002	90.6773	1.7861
	(.0110)	(0.1644)	(0.1660)	(0.1620)	(0.1692)		

counterpart, which is plausible in view of the extra-binomial variation observed in the data. With increasing autoregressive order p (and hence increasing number of parameters), some of the parameter estimates are not significant anymore. This is mainly a problem of the sample size, which equals only $N = 52$ for the analyzed data set. The smallest root mean square error is obtained for the BINARCH(3) model, which indicates that our novel model with $p = 3$ is most appropriate for the infection counts. Recalling that the counts are sampled on a weekly base, the order 3 is indeed plausible, since both for the Puumala virus and the Dobrava virus, the incubation period is known to strongly vary, namely between 2 and 4 weeks [22]. Finally, in Figure 3, we present plots of the observed and expected values for each considered model, which highlight that the higher-order BINARCH models are best suited to adapt to the higher level observed in the second half of 2011.

To further check the adequacy of the BINARCH(3) model, we use the parametric bootstrap based on the fitted model introduced by Tsay [25], which was also considered in Jung and Tremayne [26]; Weiß [27]. For parameter values $a_0 = 0.0110$, $a_1 = 0.3500$, $a_2 = 0.3216$ and $a_3 = 0.2791$, we simulate 10,000 samples of size 52 from the BINARCH(3) model. For each simulated sample, we compute the sample ACF, and for each fixed lag, we derive the 2.5% and 97.5% quantiles. By using these quantiles, we draw the bootstrap confidence intervals in Figure 4. From this graph, we can conclude that the BINARCH(3) model adequately describes the autocorrelation structure of the infection counts. Certainly, the bootstrap confidence intervals are rather wide, but this was to be expected since the sample size is small in our data example.

If we consider the other two estimation methods, the CLS method and the MLTP method, we obtain the following results for the BINARCH(3) model. The CLS estimates are $\hat{a}_0 = 0.0094$, $\hat{a}_1 = 0.2713$, $\hat{a}_2 = 0.3379$, and $\hat{a}_3 = 0.3277$, with the RMS = 1.7634. We can see that this method gives little smaller RMS than the CML method. On the other hand, the MLTP estimates are $\hat{a}_0 = 0.0032$, $\hat{a}_1 = 0.4298$, $\hat{a}_2 = 0.2995$, and $\hat{a}_3 = 0.2596$, with RMS = 1.7917. This method gives the largest RMS.

Figure 3: Plots of observed and expected values for each considered model.

Figure 4: ACF for infection counts with 95% bootstrap confidence intervals.

6 The binomial INGARCH model

In this section, we generalize the binomial INARCH(p) model and consider an integer-valued time series model for $\{X_t\}_{t \in \mathbb{Z}}$ following (1) with $\mathcal{F}_{t-1}$ and n defined as in Section 2, while α_t is generated as

$$\alpha_t = a_0 + \frac{1}{n} \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^q b_j \alpha_{t-j}, \quad \forall t \in \mathbb{Z}, \quad (8)$$

where $a_0 > 0$, $a_1 \geq 0$, $\dots$, $a_p \geq 0$, $b_1 \geq 0$, $\dots$, $b_q \geq 0$, $p \in \{1, 2, \dots\}$ and $q \in \{0, 1, 2, \dots\}$. Similarly as in the case of the BINARCH(p) model, we suppose that the parameters a_i and b_j , $i = 0, 1, \dots, p$, and $j = 1, 2, \dots, q$, satisfy the inequality $a_0 + \sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1$, which implies that $\{\alpha_t\}$ given by (8) is well-defined. We will say that the time series $\{X_t\}_{t \in \mathbb{Z}}$ given by (1) and (8) is the *binomial integer-valued GARCH model* and we will denote it as BINGARCH(p, q).

Obviously, for $q = 0$, we obtain BINARCH(p) model from before. Another important special case is $p = q = 1$, the BINGARCH(1, 1) model, which is also an instance of the observation-driven models as introduced by Davis and Liu [28], see Example 3 therein.

In the next theorem, we show that the model given by (1) and (8) is a first-order stationary process.

Theorem 8 The BINGARCH(p, q) process given by (1) and (8) is a first-order stationary process, where

$$\mu \equiv \mathbf{E}(X_t) = \frac{na_0}{1 - \sum_{i=1}^p a_i - \sum_{j=1}^q b_j}.$$

The proof of Theorem 8 is provided in Appendix A.8. For the particular case of a BINGARCH(1, 1) process, Davis and Liu [28] even established strict stationarity. Furthermore, also consistency and asymptotic normality of the maximum likelihood estimators are proven in that work.

Now, we derive and discuss the autocovariance structure of the BINGARCH(p, q) model given by (1) and (8).

Theorem 9 Let $\{X_t\}_{t \in \mathbb{Z}}$ be a stationary BINGARCH(p, q) process given by (1) and (8). The autocovariance functions $\gamma_X(k) = \text{Cov}(X_t, X_{t-k})$ and $\gamma_\alpha(k) = \text{Cov}(\alpha_t, \alpha_{t-k})$ satisfy the following equations:

$$\gamma_X(k) = \sum_{i=1}^p a_i \gamma_X(|k-i|) + \sum_{j=1}^{\min(k-1, q)} b_j \gamma_X(k-j) + n^2 \sum_{j=k}^q b_j \gamma_\alpha(j-k), \quad k \geq 1,$$

$$\gamma_\alpha(k) = \sum_{i=1}^{\min(k, p)} a_i \gamma_\alpha(k-i) + \frac{1}{n^2} \sum_{i=k+1}^p a_i \gamma_X(i-k) + \sum_{j=1}^q b_j \gamma_\alpha(|k-j|), \quad k \geq 0.$$

The proof of Theorem 9 is provided in Appendix A.9.

Example 2 For the stationary BINGARCH(1,1) model [28], we derive the autocovariance structure explicitly. From Theorem 9, we have that the autocovariance structure of the process $\{X_t\}_{t \in \mathbb{Z}}$ is given by

$$\gamma_X(1) = a_1 \gamma_X(0) + n^2 b_1 \gamma_\alpha(0), \quad (9)$$

$$\gamma_X(k) = (a_1 + b_1)^{k-1} \gamma_X(1), \quad k \geq 2.$$

On the other hand, the autocovariance structure of the process $\{\alpha_t\}_{t \in \mathbb{Z}}$ is given by

$$\gamma_\alpha(0) = \frac{a_1}{n^2} \gamma_X(1) + b_1 \gamma_\alpha(1), \quad (10)$$

$$\gamma_\alpha(k) = (a_1 + b_1)^k \gamma_\alpha(0), \quad k \geq 1. \quad (11)$$

Substituting eqs (9) and (11) for $k=1$ into eq. (10), we obtain that

$$\frac{a_1^2}{n^2} \gamma_X(0) + [(2a_1 + b_1)b_1 - 1] \gamma_\alpha(0) = 0.$$

Also from eq. (3), we have that

$$\gamma_X(0) - n(n-1) \gamma_\alpha(0) = \frac{na_0(1-a_0-a_1-b_1)}{(1-a_1-b_1)^2}.$$

Solving the last two equations with respect to $\gamma_X(0)$ and $\gamma_\alpha(0)$, we obtain that the variances of the random variables X_t and α_t are given as, respectively,

$$\text{Var}(X_t) = \frac{n^2 a_0 (1-a_0-a_1-b_1) (1-2a_1b_1-b_1^2)}{(1-a_1-b_1)^2 [a_1^2 + n(1-(a_1+b_1)^2)]},$$

$$\text{Var}(\alpha_t) = \frac{a_0 a_1^2 (1-a_0-a_1-b_1)}{(1-a_1-b_1)^2 [a_1^2 + n(1-(a_1+b_1)^2)]}.$$

Note that these variances are well-defined. Now the autocovariance and the autocorrelation functions of the process $\{X_t\}_{t \in \mathbb{Z}}$ are given as, respectively,

$$\gamma_X(k) = \frac{n^2 a_0 a_1 (1-a_0-a_1-b_1) (1-a_1b_1-b_1^2)}{(1-a_1-b_1)^2 [a_1^2 + n(1-(a_1+b_1)^2)]} (a_1+b_1)^{k-1}, \quad k \geq 1,$$

$$\rho_X(k) = \frac{a_1 (1-a_1b_1-b_1^2)}{1-2a_1b_1-b_1^2} (a_1+b_1)^{k-1}, \quad k \geq 1.$$

From the last two equations, we can conclude that the autocovariance and autocorrelation functions are positive and exponentially decaying in time.

Remark 2. Using the results from the Example 2, we can show that the BINGARCH(1,1) model can handle underdispersion, equidispersion or overdispersion with respect to a Poisson distribution. We have that

$$I_{\text{Pois}} = \frac{n(1-a_0-a_1-b_1)(1-2a_1b_1-b_1^2)}{(1-a_1-b_1)[a_1^2 + n(1-(a_1+b_1)^2)]}.$$

When the sum $a_0 + a_1 + b_1$ is close to 1, we have that I_{Pois} is close to 0. Since $1 - a_0 - a_1 - b_1 < 1 - a_1 - b_1$, we obtain that

$$I_{\text{Pois}} < \frac{n(1 - 2a_1b_1 - b_1^2)}{a_1^2 + n(1 - (a_1 + b_1)^2)} < n,$$

since $1 - 2a_1b_1 - b_1^2 < a_1^2 + n(1 - (a_1 + b_1)^2)$. So the range of I_{Pois} is bounded.

For example, when $(a_0, a_1, b_1, n) = (0.1, 0.493686, 0.1, 10)$, we obtain the equidispersion, when $(a_0, a_1, b_1, n) = (0.52, 0.42, 0.05, 10)$, we obtain that $I_{\text{Pois}} = 0.0226$ (underdispersion), and when $(a_0, a_1, b_1, n) = (0.001, 0.99, 0.001, 10)$ we obtain that $I_{\text{Pois}} = 7.6523$ (overdispersion).

Expressing the dispersion behaviour with respect to a binomial distribution, in contrast, we get

$$I_{\text{Bin}} = 1 + \frac{(n-1)\text{Var}(\alpha_t)}{\mu/n(1-\mu/n)} = 1 + \frac{(n-1)a_1^2}{n(1 - (a_1 + b_1)^2) + a_1^2},$$

i.e. we always have extra-binomial variation.

7 Conclusions

The BINARCH approach offers a way of modelling time series of counts with a finite range which exhibit extra-binomial variation. We introduced an extension of the basic BINARCH(1) model to arbitrary orders p , with a generalization to a full BINGARCH(p, q) model. After having discussed stochastic properties of these models, we analyzed three approaches for parameter estimation, and we successfully applied our novel model to a time series of counts of infections by the hantavirus.

Acknowledgments: The authors are very grateful to the Referees and to Professor Dimitris Karlis for their valuable suggestions and comments, which greatly improved this manuscript.

Conflict of Interest: The authors have declared no conflict of interest.

Appendix A Proofs

A.1 Proof of Theorem 1

The sequence of random variables $\{X_t\}_{t \in \mathbb{Z}}$ is a p th order Markov process and, hence, the vector-valued process

$$\mathbf{X}_t := (X_t, \dots, X_{t-p+1})^\top \quad \text{for } t \in \mathbb{Z}$$

constitutes a finite 1st order Markov process, i.e. a finite Markov chain. Its 1-step-ahead transition probabilities are

$$\begin{aligned} p_{\mathbf{k}|\mathbf{l}}(\mathbf{a}) &:= P(\mathbf{X}_t = \mathbf{k} | \mathbf{X}_{t-1} = \mathbf{l}) = \delta_{k_2 l_1} \cdots \delta_{k_p l_{p-1}} \cdot P(X_t = k_1 | \mathbf{X}_{t-1} = \mathbf{l}) \\ &= \delta_{k_2 l_1} \cdots \delta_{k_p l_{p-1}} \cdot \binom{n}{k_1} \alpha_0^{k_1} (1 - \alpha_0)^{n - k_1} \quad \text{with } \alpha_0 = a_0 + \frac{1}{n} \sum_{i=1}^p a_i l_i. \end{aligned} \quad (12)$$

Let $\mathbf{k}_i = (k_i, k_{i+1}, \dots, k_{i+p-1})^\top$. Then the p -step-ahead transition probabilities are

$$P(\mathbf{X}_t = \mathbf{k}_0 | \mathbf{X}_{t-p} = \mathbf{k}_p) = \prod_{j=0}^{p-1} P(X_{t-j} = k_j | X_{t-j-1} = k_{j+1}, \dots, X_{t-j-p} = k_{j+p})$$

$$= \prod_{j=0}^{p-1} \binom{n}{k_j} \alpha_j^{k_j} (1 - \alpha_j)^{n-k_j} \quad \text{with } \alpha_j = a_0 + \frac{1}{n} \sum_{i=1}^p a_i l_{j+i},$$

which are truly larger than 0 because of $a_0 > 0$. Hence, the finite Markov chain $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ is primitive, implying that it is also irreducible and aperiodic, and therefore ergodic with a unique stationary distribution [29]. Since the range of X_t is finite, any moments exist, and the strict stationarity of $\{X_t\}_{t \in \mathbb{Z}}$ also implies its second-order stationarity. $\square$

A.2 Proof of Theorem 2

Let $k \geq 1$. From the definition of the BINARCH(p) process $\{X_t\}_{t \in \mathbb{Z}}$, we have that

$$\begin{aligned} \gamma_X(k) &= \mathbf{E}[(X_t - \mu)(X_{t-k} - \mu)] = \mathbf{E}[(X_{t-k} - \mu)(\mathbf{E}(X_t | F_{t-1}) - \mu)] = \text{Cov}(n\alpha_t, X_{t-k}) \\ &= n \text{Cov}(\alpha_t, X_{t-k}) = n \text{Cov}\left(a_0 + \frac{1}{n} \sum_{i=1}^p a_i X_{t-i}, X_{t-k}\right) = \sum_{i=1}^p a_i \gamma_X(|k-i|). \end{aligned} \quad (13)$$

Let us consider now the variance of the random variable X_t . According to (3) and the fact that $\mathbf{E}(\alpha_t) = \mu/n$, we need the variance of the random variable α_t for derivation of $\text{Var}(X_t)$. Thus, we have from (2) and (13) that

$$\text{Var}(\alpha_t) = \text{Var}\left(a_0 + \frac{1}{n} \sum_{i=1}^p a_i X_{t-i}\right) = \frac{1}{n^2} \sum_{i=1}^p a_i \sum_{j=1}^p a_j \gamma_X(|i-j|) = \frac{1}{n^2} \sum_{i=1}^p a_i \gamma_X(i).$$

Replacing the last expression in (3), we obtain the expression for $\gamma_X(0)$ $\square$.

A.3 Proof of Theorem 3

To prove Theorem 3, we have to check if Condition 5.1 in Billingsley [19] is satisfied. First, the transition probabilities $p_{\mathbf{k}|\mathbf{l}}(\boldsymbol{\theta})$ from (12) are truly positive for each $(\mathbf{k}, \mathbf{l}) \in D$, where

$$D = \{(\mathbf{k}, \mathbf{l}) \in \{0, \dots, n\}^{p \times 2} \mid k_2 = l_1, \dots, k_p = l_{p-1}\}.$$

Obviously, the set D is independent of $\boldsymbol{\theta}$. Secondly, the non-zero probabilities $p_{\mathbf{k}|\mathbf{l}}(\boldsymbol{\theta})$ are polynomials in $\boldsymbol{\theta}$ and, thus, continuously differentiable in $\boldsymbol{\theta}$ up to any order. Also, see above, the transition matrix is primitive. Defining $d := |D| = (n+1)^{p+1}$, it suffices to check if the $d \times (p+1)$ -matrix

$$(\partial p_{\mathbf{k}|\mathbf{l}}(\boldsymbol{\theta}) / \partial a_u) \quad \text{with } (\mathbf{k}, \mathbf{l}) \in D \text{ and } u = 0, \dots, p$$

has rank $p+1$ for all $\boldsymbol{\theta}$. For this purpose, consider the submatrix corresponding to

$$\mathbf{k} = (n, l_1, \dots, l_{p-1})^\top, \text{ and } \mathbf{l} = \mathbf{0} \text{ or } \mathbf{l} = (n, 0, \dots, 0)^\top \text{ or } \dots \text{ or } \mathbf{l} = (0, \dots, 0, n)^\top,$$

i.e. the submatrix corresponding to the transition probabilities

$$a_0^n, (a_0 + a_1)^n, \dots, (a_0 + a_p)^n.$$

This submatrix equals

$$\begin{pmatrix} n a_0^{n-1} & 0 & \dots & 0 \\ n(a_0 + a_1)^{n-1} & n(a_0 + a_1)^{n-1} & & 0 \\ \vdots & & \ddots & 0 \\ n(a_0 + a_1)^{n-1} & 0 & & n(a_0 + a_p)^{n-1} \end{pmatrix},$$

the determinant of which is $\neq 0$ and, thus, has full rank $p+1$.

So Condition 5.1 in Billingsley [19] is satisfied, and there exists a consistent CML estimator of θ that is also asymptotically normally distributed [19, Theorems 2.1 and 2.2].

A.4 Proof of Theorem 4

We will prove this theorem by using Theorem 3.1 in Tjøstheim [18]. Let us show that all the conditions C1–C3 of this theorem are satisfied. According to the results of the previous section, the BINARCH(p) model $\{X_t\}_{t \in \mathbb{Z}}$ is strictly stationary ergodic process with finite moments. Next, the function $g_t(\theta) = \mathbf{E}(X_t | F_{t-1}) = na_0 + \sum_{i=1}^p a_i X_{t-i}$ is almost surely three times continuously differentiable with respect to the vector θ in an open set, which contains true vector θ^0 . The condition C1 is trivially satisfied since

$$\mathbf{E} \left(\frac{\partial g_t(\theta)}{\partial a_0} \right)^2 = n^2; \mathbf{E} \left(\frac{\partial g_t(\theta)}{\partial a_i} \right)^2 = \mathbf{E}(X_{t-i}^2) < \infty, \quad i = 1, 2, \dots, p; \mathbf{E} \left(\frac{\partial^2 g_t(\theta)}{\partial a_i \partial a_j} \right)^2 = 0,$$

for $i, j \in \{0, 1, \dots, p\}$. In a similar way, we can show that the condition C3 is satisfied. Now we will prove that the condition C2 is satisfied, i.e. that all the components of the vector $\partial g_t(\theta) / \partial \theta = (n, X_{t-1}, \dots, X_{t-p})^\top$ are linearly independent. This means that we need to prove that if $d_0, d_1, \dots, d_p$ are arbitrary real numbers such that

$$\mathbf{E} \left(\sum_{i=0}^p d_i \frac{\partial g_t(\theta)}{\partial a_i} \right)^2 = \mathbf{E} \left(nd_0 + \sum_{i=1}^p d_i X_{t-i} \right)^2 = 0, \quad (14)$$

then $d_0 = d_1 = \dots = d_p = 0$. We will prove that the condition C2 is satisfied following the proof of Lemma 6 [30]. Condition (14) implies that $nd_0 + \sum_{i=1}^p d_i X_{t-i} \stackrel{\text{a.s.}}{=} 0$. Now, for simplicity, let us suppose that $d_1 \neq 0$. Then

$$X_{t-1} = -\frac{nd_0}{d_1} - \sum_{i=2}^p \frac{d_i}{d_1} X_{t-i} \text{ and}$$

$$P(X_{t-1} = x_{t-1} | X_{t-i} = x_{t-i}, \quad i \geq 2) = 1(nd_0 + \sum_{i=1}^p d_i x_{t-i} = 0) \neq \binom{n}{x_{t-1}} \alpha_{t-1}^{x_{t-1}} (1 - \alpha_{t-1})^{n-x_{t-1}},$$

since X_{t-1} for given $X_{t-i}, i \geq 2$, has the binomial distribution with parameters n and α_{t-1} . Thus, we obtain a contradiction which implies that $d_1 = 0$. In a similar way, we can show that all $d_i = 0, i = 0, 1, \dots, p$.

Thus all the conditions of Theorem 3.1 in Tjøstheim [18] are satisfied, which implies that the CLS estimators $\hat{\theta}^{\text{CLS}}$ given by (6) are strongly consistent estimators of the unknown parameter θ $\square$.

A.5 Proof of Theorem 5

We only need to prove that all the elements of the matrix $\mathbf{R}$ are finite. Then the condition D1 of Theorem 3.2 [18] will be satisfied, which implies the asymptotic normality of the conditional least squares estimators $\hat{\theta}^{\text{CLS}}$. Let us denote the joint moments of the random variables $X_{t-i}, X_{t-j}, X_{t-k}$ and $X_{t-l}, i < j < k < l$, by $\mu_{i,j,k,l}(r_1, r_2, r_3, r_4) = \mathbf{E} \left(X_{t-i}^{r_1} X_{t-j}^{r_2} X_{t-k}^{r_3} X_{t-l}^{r_4} \right), r_1 \geq 0, r_2 \geq 0, r_3 \geq 0, r_4 \geq 0$. Then the elements of the matrix $\mathbf{R}$ are the joint moments $\mu_{i,j,k,l}(r_1, r_2, r_3, r_4)$, where $1 \leq r_1 + r_2 + r_3 + r_4 \leq 4$. By applying the Lemma 2 [30] when $r_i \geq 1$ with $s_i = \sum_{j=1}^i r_j, j \in \{2, 3, 4\}$, we obtain that

$$\mu_{i,j}(r_1, r_2) \leq (\mu_i(s_2))^{\frac{r_1}{s_2}} (\mu_j(s_2))^{\frac{r_2}{s_2}} < \infty,$$

$$\mu_{i,j,k}(r_1, r_2, r_3) \leq (\mu_i(s_3))^{\frac{r_1}{s_3}} (\mu_j(s_3))^{\frac{r_2}{s_3}} (\mu_k(s_3))^{\frac{r_3}{s_3}} < \infty,$$

$$\mu_{i,j,k,l}(r_1, r_2, r_3, r_4) \leq (\mu_i(S_4))^{r_1/s_4} (\mu_j(S_4))^{r_2/s_4} (\mu_k(S_4))^{r_3/s_4} (\mu_l(S_4))^{r_4/s_4} < \infty,$$

since the r th order moments of the random variable X_t are finite. Thus, we have proven that all elements of the matrix $\mathbf{R}$ are finite. $\square$

A.6 Proof of Theorem 6

To prove this theorem, we will follow the proof of Theorem 5.1 in Tjøstheim [18], i.e. we will show that the conditions E1–E3 of the mentioned theorem are satisfied. As it was mentioned earlier in the proof of Theorem 4, the BINARCH(p) model $\{X_t\}_{t \in \mathbb{Z}}$ is strictly stationary and ergodic process with finite moments of any order. Also, functions $g_t(\boldsymbol{\theta}) = \mathbf{E}(X_t | \mathcal{F}_{t-1}) = na_0 + \sum_{i=1}^p a_i X_{t-i}$ and $\text{Var}(X_t | \mathcal{F}_{t-1}) = n\alpha_t(1 - \alpha_t)$ are almost surely three times continuously differentiable with respect to the vector $\boldsymbol{\theta}$ in an open set which contains true vector $\boldsymbol{\theta}^0$. Having in mind that $\frac{\partial \alpha_t}{\partial a_0} = 1$, $\frac{\partial \alpha_t}{\partial a_i} = \frac{1}{n} X_{t-i}$, $i = 1, 2, \dots, p$, and the results given in Theorem 5 (recall Lemma 2 in Zhu and Wang [30]), it is easily seen that

$$\begin{aligned} \mathbf{E} \left| \frac{\partial \phi_t}{\partial a_i} \right| &\leq \mathbf{E} \left(\frac{X_t^2 + n\alpha_t + 2\alpha_t X_t^2 + 3n\alpha_t^2 + n^2\alpha_t^2 + 2n\alpha_t^2 X_t + 2n\alpha_t^3}{n\alpha_t^2(1 - \alpha_t)^2} \right) \\ &= \mathbf{E} \left(\mathbf{E} \left(\frac{X_t^2 + n\alpha_t + 2\alpha_t X_t^2 + 3n\alpha_t^2 + n^2\alpha_t^2 + 2n\alpha_t^2 X_t + 2n\alpha_t^3}{n\alpha_t^2(1 - \alpha_t)^2} \middle| \mathcal{F}_{t-1} \right) \right) \\ &= \mathbf{E} \left(\frac{(1 + 2\alpha_t)[n\alpha_t(1 - \alpha_t) + n^2\alpha_t^2] + n\alpha_t + 3n\alpha_t^2 + n^2\alpha_t^2 + 2n^2\alpha_t^3 + 2n\alpha_t^3}{n\alpha_t^2(1 - \alpha_t)^2} \right) \\ &= 2\mathbf{E} \left(\frac{(1 + 2\alpha_t)(1 + n\alpha_t)}{\alpha_t(1 - \alpha_t)^2} \right) < \infty, \end{aligned}$$

since $0 < a_0 \leq \alpha_t \leq \sum_{i=0}^p a_i < 1$. Similarly, based on the fact that $\frac{\partial^2 \alpha_t}{\partial a_i \partial a_j} = 0$, $i, j \in \{0, 1, \dots, p\}$, we find

$$\mathbf{E} \left| \frac{\partial^2 \phi_t}{\partial a_i \partial a_j} \right| < \infty, \quad i, j \in \{0, 1, \dots, p\}.$$

We conclude that condition E1 is satisfied. In a similar way, it can be shown that condition E3 is also satisfied. The only thing left to prove is that condition E2 is fulfilled, i.e. if we assume that for arbitrary real numbers $d_0, d_1, \dots, d_p$, it holds that

$$\mathbf{E} \left(\frac{\left(d_0 n + \sum_{i=1}^p d_i X_{t-i} \right)^2}{n\alpha_t(1 - \alpha_t)} \right) + \mathbf{E} \left(\frac{(1 - 2\alpha_t) \left(d_0 n + \sum_{i=1}^p d_i X_{t-i} \right)^2}{n\alpha_t(1 - \alpha_t)} \right) = 0, \quad (15)$$

then $d_0 = d_1 = \dots = d_p = 0$. Equality (15) implies

$$\left(d_0 n + \sum_{i=1}^p d_i X_{t-i} \right)^2 \frac{\alpha_t(1 - \alpha_t)(n + 4) + 1}{(n\alpha_t(1 - \alpha_t))^2} \stackrel{\text{a.s.}}{=} 0.$$

It is easy to conclude that $\alpha_t(1 - \alpha_t)(n + 4) + 1 \neq 0$, and further implies that $nd_0 + \sum_{i=1}^p d_i X_{t-i} \stackrel{\text{a.s.}}{=} 0$. The rest of the proof is equivalent to the proof of Theorem 4. $\square$

A.7 Proof of Theorem 7

Following the proof of Theorem 5 and using the fact that $0 < a_0 \leq \alpha_t \leq \sum_{i=0}^p a_i < 1$, we obtain that all the elements of the matrix $\mathbf{S}$ are finite. This implies that condition F1 of Theorem 5.2 in Tjøstheim [18] is satisfied, which proves the asymptotic normality of the estimators $\{\hat{\boldsymbol{\theta}}^{\text{MLTP}}\}$ obtained by minimizing (7). $\square$

A.8 Proof of Theorem 8

Let $\mu_t = E(X_t)$. Using the fact that $\mathbf{E}(\alpha_t) = \mu_t/n$, we obtain from the eq. (8) that

$$\mu_t = na_0 + \sum_{i=1}^p a_i \mu_{t-i} + \sum_{j=1}^q b_j \mu_{t-j}. \quad (16)$$

The eq. (16) represents a non-homogeneous difference equation. According to Goldberg [31], this equation has a finite stable solution which is independent of t if all the roots of the equation

$$1 - \sum_{i=1}^p a_i z^{-i} - \sum_{j=1}^q b_j z^{-j} = 0$$

lie inside the unit circle. Since the parameters a_i , $i = 0, 1, \dots, p$, and b_j , $j = 1, 2, \dots, q$, are non-negative, it follows that the roots lie inside the unit circle if $\sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1$. Finally, this condition is satisfied since $a_0 + \sum_{i=1}^p a_i + \sum_{j=1}^q b_j < 1$. $\square$

A.9 Proof of Theorem 9

The proof of Theorem 9 is an immediate consequence of the following Lemma 1, which provides the dependence between the random variables X_t and α_{t-k} for $k \in \mathbb{Z}$.

Lemma 1 Let $\{X_t\}_{t \in \mathbb{Z}}$ be a BINGARCH(p, q) model given by (1) and (8). Then the covariance function between the random variables X_t and α_{t-k} , $k \in \mathbb{Z}$, is given by

$$\text{Cov}(X_t, \alpha_{t-k}) = \begin{cases} n \text{Cov}(\alpha_t, \alpha_{t-k}), & k \geq 0, \\ \frac{1}{n} \text{Cov}(X_t, X_{t-k}), & k < 0. \end{cases}$$

Proof. Let us first derive the covariance function of the random variables X_t and X_{t-k} , $k > 0$. Since $\{X_t\}_{t \in \mathbb{Z}}$ is a first-order stationary process, we have that $\mu = \mathbf{E}(X_t)$, for all $t \in \mathbb{Z}$. Since $\mathbf{E}(X_t | \mathcal{F}_{t-1}) = n\alpha_t$ and X_{t-k} is a $\mathcal{F}_{t-1}$ -measurable function, we obtain that

$$\text{Cov}(X_t, X_{t-k}) = \mathbf{E}[(X_t - \mu)(X_{t-k} - \mu)] = \mathbf{E}[(X_{t-k} - \mu)(\mathbf{E}(X_t | \mathcal{F}_{t-1}) - \mu)] = n \text{Cov}(\alpha_t, X_{t-k}),$$

which implies that $\text{Cov}(X_t, \alpha_{t-k}) = \frac{1}{n} \text{Cov}(X_t, X_{t-k})$, for $k < 0$.

Let us now derive the covariance $\text{Cov}(X_t, \alpha_{t-k})$ for $k \geq 0$. Let A_t represents the σ -field generated by the random variables $\{\alpha_t\}_{t \geq 0}$. Since $\mathbf{E}(X_t | A_t) = \mathbf{E}(\mathbf{E}(X_t | A_t, \mathcal{F}_{t-1}) | A_t) = n\alpha_t$, the random variable α_{t-k} is an A_t -measurable function and $\mathbf{E}(\alpha_t) = \mu/n$, we obtain that

$$\begin{aligned} \text{Cov}(X_t, n\alpha_{t-k}) &= \mathbf{E}[(X_t - \mu)(n\alpha_{t-k} - \mu)] = \mathbf{E}[(n\alpha_{t-k} - \mu)(\mathbf{E}(X_t | A_t, \mathcal{F}_{t-1}) | A_t) - \mu)] \\ &= \text{Cov}(n\alpha_t, n\alpha_{t-k}), \end{aligned}$$

which implies that $\text{Cov}(X_t, \alpha_{t-k}) = n \text{Cov}(\alpha_t, \alpha_{t-k})$ for $k \geq 0$. $\square$

Appendix B CLS estimators of the BINARCH(1) model

Proposition 1 below provides closed-form expressions for the matrices $\mathbf{U}$ and $\mathbf{R}$ for BINARCH(1) model. Before, we provide the following lemma with the expressions for the first four moments, which will be used to derive the elements of the matrices $\mathbf{U}$ and $\mathbf{R}$.

Lemma 2 Let $\{X_t\}_{t \in \mathbb{Z}}$ be the BINARCH(1) process given by (1) and (2). Then $\mu_1 \equiv \mathbf{E}(X_t) = nA_1$, $A_1 \equiv \mathbf{E}(\alpha_t) = a_0/(1 - a_1)$, and the second, third and fourth moments μ_i and A_i , $i = 2, 3, 4$, of the random variables X_t and α_t , respectively, are given as solutions of the following equations

$$\begin{aligned} \begin{pmatrix} \mu_2 \\ A_2 \end{pmatrix} &= \begin{pmatrix} 1 & -n(n-1) \\ \frac{a_1^2}{n^2} & -1 \end{pmatrix}^{-1} \begin{pmatrix} nA_1 \\ -a_0^2 - 2a_0a_1A_1 \end{pmatrix} \\ \begin{pmatrix} \mu_3 \\ A_3 \end{pmatrix} &= \begin{pmatrix} 1 & -n(n-1)(n-2) \\ \frac{a_1^3}{n^3} & -1 \end{pmatrix}^{-1} \begin{pmatrix} nA_1 + 3n(n-1)A_2 \\ -a_0^3 - 3a_0^2a_1A_1 - \frac{3a_0a_1^2}{n^2}\mu_2 \end{pmatrix} \\ \begin{pmatrix} \mu_4 \\ A_4 \end{pmatrix} &= \begin{pmatrix} 1 & -24C_4^n \\ \frac{a_1^4}{n^4} & -1 \end{pmatrix}^{-1} \begin{pmatrix} nA_1 + 7n(n-1)A_2 + 6n(n-1)(n-2)A_3 \\ -a_0^4 - 4a_0^3a_1A_1 - \frac{6a_0^2a_1^2}{n^2}\mu_2 - \frac{4a_0a_1^3}{n^3}\mu_3 \end{pmatrix}, \end{aligned}$$

where $C_4^n = n(n-1)(n-2)(n-3)/24$.

The proof of Lemma 2 follows after standard calculations, which are based on the fact that the random variable X_t for given X_{t-1} has the binomial distribution with parameters n and α_t , and expressions for its moments about 0.

Proposition 1 Let $\{X_t\}_{t \in \mathbb{Z}}$ be the BINARCH(1) process given by (1) and (2). Then the matrix $\mathbf{U}$ is given by

$$\mathbf{U} = \begin{pmatrix} n^2 & n\mu_1 \\ n\mu_1 & \mu_2 \end{pmatrix},$$

where μ_1 and μ_2 are given as in the previous Lemma. The matrix $\mathbf{R}$ follows as

$$\mathbf{R} = n \begin{pmatrix} n^2(A_1 - A_2) & na_0(1 - a_0)\mu_1 + a_1(1 - 2a_0)\mu_2 - \frac{a_1^2}{n}\mu_3 \\ na_0(1 - a_0)\mu_1 + a_1(1 - 2a_0)\mu_2 - \frac{a_1^2}{n}\mu_3 & a_0(1 - a_0)\mu_2 + \frac{a_1}{n}(1 - 2a_0)\mu_3 - \frac{a_1^2}{n^2}\mu_4 \end{pmatrix}.$$

To derive the matrix $\mathbf{U}$, we need the first- and second-order moments of the random variable X_{t-1} . On the other hand, for derivation of the matrix $\mathbf{R}$, we need up to the fourth-order moment of the random variable X_{t-1} , since

$$\mathbf{E}(\alpha_t(1 - \alpha_t)X_{t-1}^i) = a_0(1 - a_0)\mathbf{E}(X_{t-1}^i) + \frac{a_1}{n}(1 - 2a_0)\mathbf{E}(X_{t-1}^{i+1}) - \frac{a_1^2}{n^2}\mathbf{E}(X_{t-1}^{i+2}), \quad i = 0, 1, 2.$$

All these moments can be obtained by using the results of the previous Lemma 2.

References

1. Weiß CH. A new class of autoregressive models for time series of binomial counts. *Commun Stat Theor Method* 2009;38:447–60.
2. Weiß CH, Kim HY. Binomial AR (1) processes: moments, cumulants, and estimation. *Statistics* 2013;47:494–510.
3. Weiß CH, Kim HY. Parameter estimation for binomial AR (1) models with applications in finance and industry. *Stat Papers* 2013;54:563–90.
4. Weiß CH, Pollett PK. Chain binomial models and binomial autoregressive processes. *Biometrics* 2012;68:815–24.
5. Steutel FW, van Harn K. Discrete analogues of self-decomposability and stability. *Ann Prob* 1979;7:893–9.
6. McKenzie E. Some simple models for discrete variate time series. *Water Resour Bull* 1985;21:645–50.

7. Cui Y, Lund R. Inference in binomial AR (1) models. *Stat Prob Lett* 2010;80:1985–90.
8. Scotto MG, Weiß CH, Silva ME, Pereira I. Bivariate binomial autoregressive models. *J Multivar Anal* 2014;125:233–51.
9. Al-Osh MA, Alzaid AA. Binomial autoregressive moving average models. *Stochastic Models* 1991;7:261–82.
10. Ferland R, Latour A, Oraichi D. Integer-valued GARCH process. *J Time Ser Anal* 2006;27:923–42.
11. Weiß CH, Pollett PK. Binomial autoregressive processes with density-dependent thinning. *J Time Ser Anal* 2014;35:115–32.
12. Zhu F. A negative binomial integer-valued GARCH model. *J Time Ser Anal* 2011;32:54–67.
13. Zhu F. Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *J Math Anal Appl* 2012;389:58–71.
14. Xu H-Y, Xie M, Goh TN, Fu X. A model for integer-valued time series with conditional overdispersion. *Comput Stat Data Anal* 2012;56:4229–42.
15. Gonçalves E, Mendes-Lopes N, Silva F. Infinitely divisible distributions in integer-valued GARCH models. *J Time Ser Anal* 2015;36:503–27.
16. Held L, Hofmann M, Höhle M, Schmid V. A two-component model for counts of infectious diseases. *Biostatistics* 2006;7:422–37.
17. Höhle M, Paul M. Count data regression charts for the monitoring of surveillance time series. *Comput Stat Data Anal* 2008;52:4357–68.
18. Tjøstheim D. Estimation in nonlinear time series models. *Stochastic Processes Appl* 1986;21:251–73.
19. Billingsley P. Statistical inference for Markov processes. Chicago: University of Chicago Press, 1961.
20. Robert-Koch-Institut. <http://www3.rki.de/SurvStat>, data status: June 12, 2013.
21. Faulde M, Sobe D, Kimmig P, Scharninghausen J. Renal failure and hantavirus infection in Europe. *Nephrol Dialysis Transpl* 2000;15:751–3.
22. Heyman P, Vaheri A, Lundkvist Å, Avsic-Zupanc T. Hantavirus infections in Europe: from virus carriers to a major public-health problem. *Expert Rev Anti-Infect Therapy* 2009;7:205–17.
23. MacNeil A, Comer JA, Ksiazeka TG, Rollin PE. Sin Nombre virus-specific immunoglobulin M and G kinetics in hantavirus pulmonary syndrome and the role played by serologic responses in predicting disease outcome. *J Infect Dis* 2010;202:242–6.
24. Schilling S, Emmerich P, Klempa B, Auste B, Schnaith E, Schmitz H, et al. Hantavirus disease outbreak in Germany: limitations of routine serological diagnostics and clustering of virus sequences of human and rodent origin. *J Clin Microbiol* 2007;45:3008–14.
25. Tsay RS. Model checking via parametric bootstraps in time series analysis. *Appl Stat* 1992;41:1–15.
26. Jung RC, Tremayne AR. Convolution-closed models for count time series with applications. *J Time Ser Anal* 2011;32:268–80.
27. Weiß CH. A Poisson INAR (1) model with serially dependent innovations. *Metrika* 2015;78:829–51.
28. Davis RA, Liu H. Theory and inference for a class of observation-driven models with application to time series of counts. *eprint arXiv:1204.3915*. 2012.
29. Seneta E. Non-negative matrices and Markov chains, 2nd ed. New York: Springer Verlag, 1983.
30. Zhu F, Wang D. Estimation and testing for a Poisson autoregressive model. *Metrika* 2011;73:211–30.
31. Goldberg S. Difference equations. New York: Wiley, 1958.