Communications of the Korean Statistical Society
2010, Vol. 17, No. 1, 27–37

Coherent Forecasting in Binomial AR(p) Model

Hee-Young Kim[1,a], Yousung Park[b]

[a]Institute of Economics, Korea University; [b]Department of Statistics, Korea University

Abstract

This article concerns the forecasting in binomial AR(p) models which is proposed by Weiß (2009b) for time series of binomial counts. Our method extends to binomial AR(p) models a recent result by Jung and Tremayne (2006) for integer-valued autoregressive model of second order, INAR(2), with simple Poisson innovations. Forecasts are produced by conditional median which gives 'coherent' forecasts, and we estimate the forecast distributions of future values of binomial AR(p) models by means of a Monte Carlo method allowing for parameter uncertainty. Model parameters are estimated by the method of moments and estimated standard errors are calculated by means of block of block bootstrap. The method is fitted to log data set used in Weiß (2009b).

Keywords: Binomial thinning, binomial AR(p) model, block-of-blocks bootstrap.

1. Introduction

A binomial thinning operation developed by Steutel and van Harn (1979) is the most popular operation which is used to model integer-valued time series data. The purpose of the binomial thinning is to ensure the integer discreteness of the process. It is defined as $\alpha \circ X = \sum_{i=1}^{X} Y_i$, where Y_i are assumed to be *i.i.d.* Bernoulli random variables with $P(Y_i = 1) = \alpha$, $P(Y_i = 0) = 1 - \alpha$, and independent of X. The first integer-valued ARMA(INARMA) model using binomial thinning operation is INAR(1) model which is introduced by McKenzie (1985) and independently by Al-Osh and Alzaid (1987). Since INAR(1) model is subcritical Galton-Watson process with immigration, and also is related to the $M/M/\infty$ queueing system, there are many applications in diverse scientific fields ranging from medicine to economics, environmentology, insurance, *etc*. And there are various INARMA models, including INAR(p) models of higher auto-regressive order (Alzaid and Al-Osh, 1988, 1990; Du and Li, 1991; Al-Osh and Aly, 1992), generalized-INAR models (Latour, 1998; Brännäs and Hellström, 2001) and INARS model (Kim and Park, 2008; Park *et al.*, 2006).

But INAR(1) model cannot be directly applied to process with a finite range of counts, for example, binomial distribution. The reason is that any distribution of the discrete self-decomposable(DSD) family which contains the negative binomial distribution and the generalized Poisson distribution can be a possible marginal distribution of INAR(1) model.

McKenzie (1985) proposed binomial AR(1) model which is essentially designed to model binomial distribution, still using binomial thinning operation. Recently Weiß (2009a) used this model in statistical process control(SPC), and proposed various control charts to monitor correlated binomial distribution. Also Weiß (2009b) proposed a new class of p^{th} order autoregressive models, which coincide with the binomial AR(1) model for $p = 1$.

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2007-411-J03302).

[1] Corresponding author: Research Professor, Institute of Economics, Korea University, Anam-dong, Seongbuk-gu, Seoul, 136-701, Korea. E-mail: starkim@korea.ac.kr

In this paper, we are concerned with developing forecasting procedures in a binomial AR(p) model. In fact, there are some difficulty in deriving confidence intervals of forecasts in models using binomial thinning operations, because of the distributional complexity accrued from the binomial thinning operations, therefore particular methods were raised by researchers (Kim and Park, 2006a; Kim and Park, 2006b; Jung and Tremayne, 2006; Freeland and McCabe, 2004).

Freeland and McCabe (2004) emphasized the needs of forecasting the h-step ahead conditional distribution and coherent forecasting procedures, in cases where the variable is discrete and the cardinality of the support is small. The coherent forecasting procedures means preserving the integer structure of the data in generating the forecasts. Freeland and McCabe (2004) concerned with INAR(1) model, and later Jung and Tremayne (2006) extended Freeland and McCabe (2004)'s method to INAR(2) model.

In this article, we will adopt Jung and Tremayne (2006)'s approach to binomial AR(p) model. The set-up of the present paper is as follows. In Section 2, we briefly review the binomial AR(1) and the binomial AR(p) model. In Section 3, we estimate the forecast distributions of future values of binomial AR(p) models by means of a Monte Carlo method allowing for parameter uncertainty. Model parameters are estimated by the method of moments and estimated standard errors are calculated by means of block of block bootstrap. Section 4 gives an empirical example of forecasting procedure using the method in Section 3. Finally Section 5 provides some concluding remarks.

2. The Binomial Autoregressive Model

In Section 2, we briefly review elementary properties of binomial AR(1) and binomial AR(p) model.

2.1. The Binomial AR(1) model

Definition 1. *(Binomial AR(1) model, McKenzie, 1985) Fix $n \in \mathbb{N}$. Let $\pi \in (0,1)$, and $\rho \in [max(-\pi/(1-\pi), -(1-\pi)/\pi), 1]$. Define $\beta = \pi(1-\rho)$ and $\alpha = \beta + \rho$. The process $\{X_t\}$,*

$$X_t = \alpha \circ X_{t-1} + \beta \circ (n - X_{t-1}), \quad t \geq 1,\ X_0 \sim B(n, \pi) \tag{2.1}$$

is said to be a Binomial AR(1) model, where all thinnings are performed independently of each other, and the thinnings at time t are independent of $\{X_s, s < t\}$.

The binomial AR(1) model can be interpreted as follows. There are n units, which are independently of each other, either in state 1 or state 0. Let X_t be the number of units being in state 1 at time t, and X_{t-1} be the number of units being in state 1 at time $t-1$. Then X_t is composed of two random components: the number of units which are are still in state 1 at time t, $\alpha \circ X_{t-1}$, and the number of units which moved from state 0 to state 1 at time t, $\beta \circ (n - X_{t-1})$.

$$X_t = \underbrace{\alpha \circ X_{t-1}}_{\text{still in state 1 at time } t} + \underbrace{\beta \circ (n - X_{t-1})}_{\text{from state 0 to state 1 at time } t}.$$

A binomial AR(1) model is a stationary Markov chain with $n+1$ states, and the properties of the binomial AR(1) model are stated as follows.

P.1 $E(X_t) = n\pi$, $\text{Var}(X_t) = n\pi(1-\pi)$, $\text{Corr}(X_t, X_{t-k}) = \rho^k$.

P.2 The transition probability, $P(X_t = k|X_{t-1} = l)$, $k, l = 0, \ldots, n$ is

$$\sum_{m=max(0,k+l-n)}^{min(k,l)} \binom{l}{m}\binom{n-l}{k-m}\alpha^m(1-\alpha)^{l-m}(1-\beta)^{n-l+m-k}.$$

P.3 $E(X_t|X_{t-1}) = \rho X_{t-1} + n\beta$.

P.4 $\mathrm{Var}(X_t|X_{t-1}) = \rho(1-\rho)(1-2\pi)X_{t-1} + n\beta(1-\beta)$.

Figure 1(a)–(f) shows simulated sample paths for binomial AR(1) processes for different parameter combinations. The realization in Figure 1(a) is generated using $n = 6$, $\pi = 0.5$, $\rho = 0.5$, the corresponding sample autocorrelation function(SACF) and sample partial autocorrelation function(SPACF) of that series are shown in Figure 1(b) and (c). The SACF decays exponentially and there are spike in the direction of positive at 1 lag in SPACF. In Figure 1(d), a simulated sample path is based on the parameter values $n = 6$, $\pi = 0.5$, $\rho = -0.5$. The corresponding SACF and SPACF are depicted in Figure 1(e) and (f). We can see oscillating behaviour in SACF, and there are spike in the direction of negative at 1 lag in SPACF.

2.2. The Binomial AR(p) model

Weiß (2009b) extended the binomial AR(1) model to a higher-order autoregressive model as a tool for modelling and generating sequences of dependent binomial process. To simplify the notation, Weiß (2009b) used a random function, $f_k(X_t)$, which is defined as $f_k(X_t) = \alpha \circ_{t+k} X_t + \beta \circ_{t+k} (n - X_t)$, where " $\circ_{t+k}$ " denotes that the thinning is performed at time $t + k$. Notice that the time index below the thinning operation indicates the time when the corresponding thinning is performed.

Definition 2. *(Binomial AR(p) model, Weiß, 2009b) Let* $\pi \in (0,1)$ *and* $\rho \in [max(-\pi/(1-\pi), -(1-\pi)/\pi), 1]$. *Define* $\beta = \pi(1-\rho)$ *and* $\alpha = \beta + \rho$. *Let* $\{\mathbf{D}_t\}$ *be an i.i.d. multinomial distribution with parameters* $\mathbf{D}_t = (D_{t,1}, \ldots, D_{t,p}) \sim MULT(1; \phi_1, \ldots, \phi_p)$. *Let a process* $\{X_t\}$ *with range* $\{0, \ldots, n\}$ *follow the recursion*

$$\begin{aligned} X_t &= D_{t,1}(\alpha \circ_t X_{t-1} + \beta \circ_t (n - X_{t-1})) + \cdots + D_{t,p}(\alpha \circ_t X_{t-p} + \beta \circ_t (n - X_{t-p})) \\ &= \sum_{i=1}^{p} D_{t,1} f_i(X_{t-i}). \end{aligned} \tag{2.2}$$

It is said to be a binomial AR(p) process, if conditions C.1 ∼ C.3 are satisfied. It is called a binomial AR(p)-Independent thinning process, if conditions C.1 ∼ C.4 are satisfied.

(C.1) The thinnings at time t are performed independently of each other and of $\{\mathbf{D}_t\}$.

(C.2) $\mathbf{D}_t = (D_{t,1}, \ldots, D_{t,p})$ *is independent of all* $\{X_s, s < t\}$ *and* $\{f_j(X_s), s < t, j = 1, \ldots, p\}$.

(C.3) The conditional probability

$$\begin{aligned} &P\Big(f_1(X_t) = i_1, \ldots, f_p(X_t) = i_p | X_t = x_t,\ X_{t-k} = x_{t-k}, k \geq 1; f_j(X_{t-k}) = z_{t-k}, k \leq 1, j = 1, \ldots, p\Big) \\ &= P\Big(f_1(X_t) = i_1, \ldots, f_p(X_t) = i_p | X_t = x_t\Big). \end{aligned}$$

(C.4) All thinnings $f_1(X_t), \ldots, f_p(X_t)$ *are conditionally independent, conditioned on* X_t.

With this construction, X_t can be defined by

$$X_t = \begin{cases} \alpha \circ_t X_{t-1} + \beta \circ_t (n - X_{t-1}), & \text{with probability } \phi_1, \\ \quad\vdots & \quad\vdots \\ \alpha \circ_t X_{t-p} + \beta \circ_t (n - X_{t-p}), & \text{with probability } \phi_p. \end{cases}$$

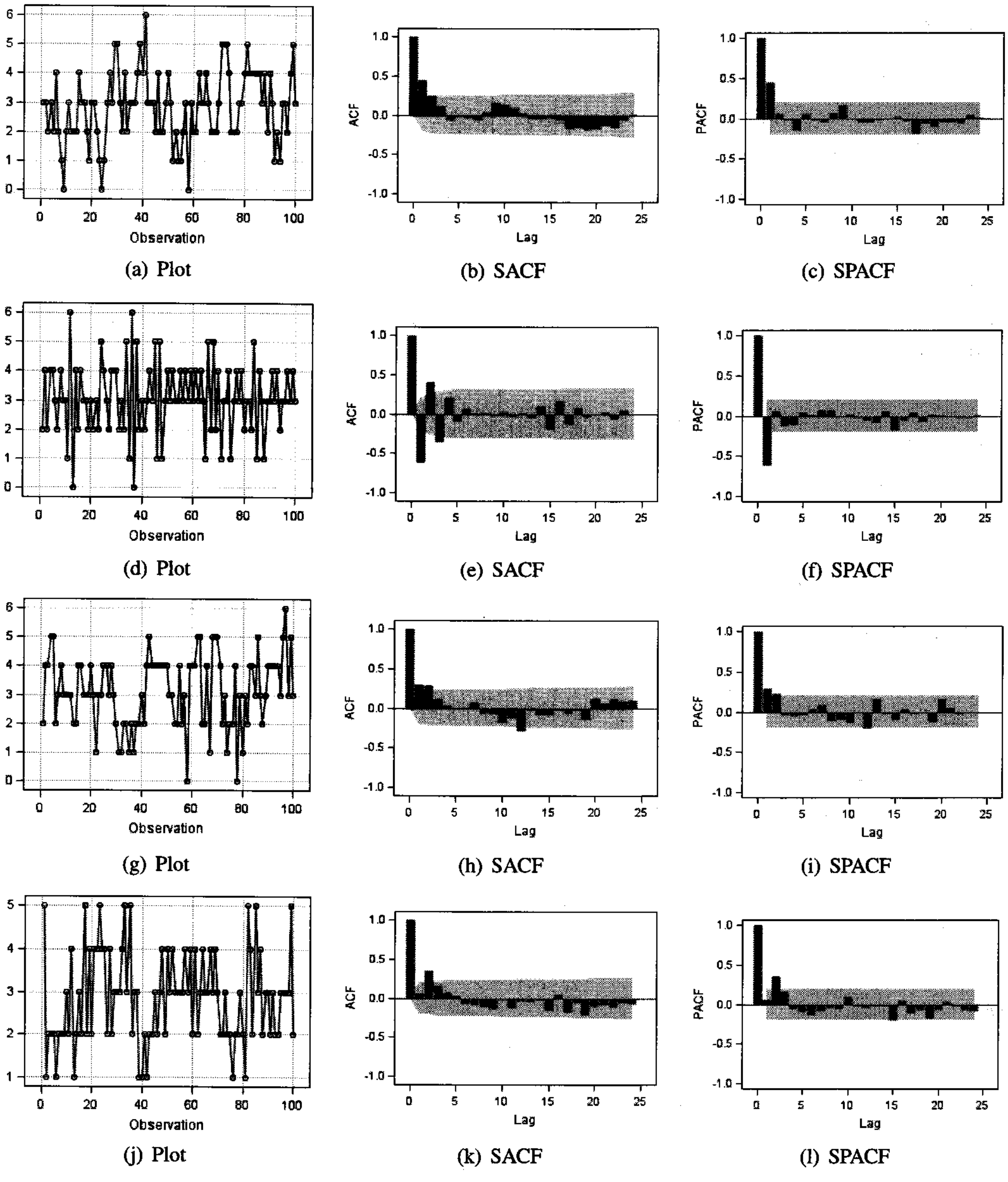

Figure 1: *Simulated sample paths and sample ACF and PACF of a Binomial AR(1) process and a Binomial AR(2) process: (a) simulated sample paths for a binomial AR(1) process, $n = 6, \pi = 0.5, \rho = 0.5$; (b) sample ACF of (a); (c) sample PACF of (a); (d) simulated sample paths for a binomial AR(1) process, $n = 6, \pi = 0.5, \rho = -0.5$; (e) sample ACF of (d); (f) sample PACF of (d); (g) simulated sample paths for a binomial AR(2) process, $n = 6, \pi = 0.5, \rho = 0.5, \phi_1 = 0.5$; (h) sample ACF of (g); (i) sample PACF of (g); (j) simulated sample paths for a binomial AR(2) process, $n = 6, \pi = 0.5, \rho = 0.5, \phi_1 = 0.2$; (k) sample ACF of (j); (l) sample PACF of (j).*

And due to the additional assumption *C.4*, Weiß (2009b) showed that the autocorrelation function of binomial AR(p)-Independent thinnings process satisfies the pth order difference equation

$$\rho(k) = \rho \sum_{i=1}^{p} \phi_i \rho(|k - i|), \tag{2.3}$$

and the k^{th} partial autocorrelation $\phi_{kk} = 0, k > p$. Hence, based on this fact, we can identify the model order p. This feature of a binomial AR(p)-Independent thinnings process makes it a satisfactory model for time series of binomial counts. Henceforth we will denote a binomial AR(p)-Independent thinnings process as BiAR(p)-Ind process.

Figure 1(g)–(l) depicts the simulated observations for BiAR(2)-Ind model for different parameter combinations. Figure 1(g) is that of $n = 6, \pi = 0.5, \rho = 0.5, \phi_1 = 0.5$, and Figure 1(j) is that of $n = 6, \pi = 0.5, \rho = 0.5, \phi_1 = 0.2$. In both cases, all parameters are the same, except for ϕ_1. We can observe the SPACF for each series(Figure 1(i) and (l)) is essentially zero for lags larger than 2. The SACF(Figure 1(h)) and the SPACF(Figure 1(i)) of $n = 6, \pi = 0.5, \rho = 0.5, \phi_1 = 0.5$ has spike at lag1 and lag2, but those(Figure 1(k) and (l)) of $n = 6, \pi = 0.5, \rho = 0.5, \phi_1 = 0.2$ have only at lag2.

For estimation and forecasting purposes the conditional probabilities $P(X_{T+h} = x_{T+h}|X_T = x_T, X_{T-1} = x_{T-1}, \ldots)$ play a key role. Weiß (2009b) presented 1-step ahead conditional distribution of BiAR(p)-Ind model, but when $h \geq 2$, h-step ahead conditional distribution is not easily obtained because of thinning operations.

3. Forecasting in the Binomial AR(p)-Independent Model

Prediction is one of the main goals in time series analysis. Conditional on $\{X_1, \ldots, X_T\}$, the minimum mean square error(MSE) predictor of $X_{T+h}, h > 0$ is given by the conditional mean of $E(X_{T+h}|X_T, \ldots, X_1)$. Despite its optimality property, forecasting based on the conditional expectation suffer from the fact that it could not produce integer value in the count data context. In contrast, the conditional median could produce coherent forecasts, preserving the integer structure of the data in generating the forecasts. The coherent forecasting in integer-valued time series model was first developed by Freeland and McCabe (2004) for a INAR(1) model, and Jung and Tremayne (2006) provided the methods for a INAR(2) model.

Also they all stressed the need of the forecasts for each point mass of the distribution, in the case of analysis of low count time series. A density forecast is an estimate of the future probability distribution of a random variable, conditional on the data available at the time the forecast is made. A density forecasting is receiving increasing interest even in continuous time series analysis, in particular macroeconomics and finance (Tay and Wallis, 2000). In this section, we adopt the procedure in Jung and Tremayne (2006) to obtain coherent forecasting and the prediction for point mass of the distribution, in BiAR(p)-Ind model.

These schemes use bootstrap techniques to estimate asymptotic standard errors of the YW estimates and to forecast distributions of future values allowing for parameter uncertainty. Summarizing, the steps for obtaining bootstrap prediction intervals are:

Step 1: Let $\{X_1, \ldots, X_T\}$ denote the observations, and choose a model order p by examination the patterns in the SACF and the the SPACF, because a BiAR(p)-Ind process satisfies (2.3).

Step 2: Estimate the model parameters by the method of moments based on the Yule-Walker equation (2.3), that is,

$$\hat{\pi} = \frac{\sum_{i=1}^{T} X_i}{T},$$
$$\hat{\rho}(k) = \hat{\rho} \sum_{i=1}^{p} \hat{\phi}_i \hat{\rho}(|k-i|), \quad \text{on condition that } \sum_{i=1}^{p} \phi_i = 1. \tag{3.1}$$

Step 3: To estimate the asymptotic standard error of the Yule-Walker estimator and to allow for the uncertainty in parameter estimation in estimating the forecast distribution, implement the blocks of blocks bootstrap approach suggested by Künsch (1989). To implement resampling blocks of blocks, define a new $(p+1)$-variate process $Y_i = (X_i, \ldots, X_{i+p})'$, and define $\mathbf{Y}$ using $\{X_1, \ldots, X_T\}$ as follows.

$$\mathbf{Y} = (Y_1, Y_2, \ldots, Y_{T-p}) = \begin{pmatrix} X_1 & X_2 & \cdots & X_{T-p} \\ X_2 & X_3 & \cdots & X_{T-p+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p+1} & X_{2+p} & \cdots & X_T \end{pmatrix} \stackrel{Let}{=} \begin{pmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,T-p} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,T-p+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p+1,1} & X_{p+1,2} & \cdots & X_{p+1,T-p} \end{pmatrix}. \quad (3.2)$$

Note that the sample lag covariance of order $k, 0 \le k \le p$ is

$$\hat{\gamma}(k) = \frac{1}{T-k}\sum_{i=1}^{T-k}\left(X_i - \bar{X}\right)\left(X_{i+k} - \bar{X}\right) \simeq \frac{1}{T-k}\sum_{i=1}^{T-k}\left(X_{1,i} - \bar{X}_{1.}\right)\left(X_{1+k,i} - \bar{X}_{1+k.}\right). \quad (3.3)$$

Fix a block size $l, 1 < l < T-p-2$, and define the blocks in terms of $Y_i's$ as $B_j = (Y_j, \ldots, Y_{j+l-1}), 1 \le j \le T-l+1$. Select k blocks randomly from the collections $\{B_i : 1 \le i \le T-p-l+1\}$ to generate $Y_1^*, \ldots, Y_l^*; Y_{l+1}^*, \ldots, Y_{2l}^*; \ldots, Y_T^*$, where $T = kl$. Then the bootstrap version of $\hat{\gamma}(k)$ based on bootstrap sample is

$$\hat{\gamma}^*(k) = \frac{1}{T-k}\sum_{i=1}^{T-k}\left(X_{1,i}^* - \bar{X}_{1.}^*\right)\left(X_{1+k,i}^* - \bar{X}_{1+k.}^*\right), \quad 0 \le k \le p. \quad (3.4)$$

Step 4: Calculate the Yule-Walker estimator $\hat{\boldsymbol{\vartheta}}^* = (\hat{\pi}^*, \hat{\rho}^*, \hat{\phi}_1^*, \ldots, \hat{\phi}_p^*)$ using (3.4).

Step 5: Repeat Step 3–4 B_1 times, and store $\{\hat{\boldsymbol{\vartheta}}_b^*, b = 1, \ldots, B_1\}$

Step 6: Draw one parameter estimates vector randomly with replacement from the set $\{\hat{\boldsymbol{\vartheta}}_b^*, b = 1, \ldots, B_1\}$ in Step 5, and generate $\hat{X}_{T+h}$ using selected parameter estimates $\hat{\boldsymbol{\vartheta}}^* = (\hat{\pi}^*, \hat{\rho}^*, \hat{\phi}_1^*, \ldots, \hat{\phi}_p^*)$ as follows.

$$\hat{X}_{T+h} = \hat{D}_{t,1}^*\left(\hat{\alpha}^* \circ_t \hat{X}_{t+h-1} + \hat{\beta}^* \circ_t \left(n - \hat{X}_{t+h-1}\right)\right) + \cdots + \hat{D}_{t,p}^*\left(\hat{\alpha}^* \circ_t \hat{X}_{t+h-p} + \hat{\beta}^* \circ_t \left(n - \hat{X}_{t+h-p}\right)\right),$$

where $\hat{X}_{T+j} = X_{T+j}, j \le 0$ and $(\hat{D}_{t,1}^*, \ldots, \hat{D}_{t,p}^*) \sim \text{MULT}(1; \hat{\phi}_1^*, \ldots, \hat{\phi}_p^*)$.

Step 7: For each selected parameter estimates $\hat{\boldsymbol{\vartheta}}^* = (\hat{\pi}^*, \hat{\rho}^*, \hat{\phi}_1^*, \ldots, \hat{\phi}_p^*)$, repeat Step 6 B_2 times to obtain a single estimate of the forecast distribution and its median, say, $\widehat{\text{Med}}(h)$. This step proposes point forecasts based on the integer-valued median of the forecast distribution.

Step 8: Repeat Step 6–7 B_3 times to assess the variability of the point masses of the forecast distribution and to incorporate parameter uncertainty. Therefore, we obtain $n+1$ sets as as follows:

$$\left\{P^{*(i)}\left(X_{T+h}^* = 0\right)\right\}_{i=1}^{B^0}; \ \left\{P^{*(i)}\left(X_{T+h}^* = 1\right)\right\}_{i=1}^{B^1}; \cdots; \left\{P^{*(i)}\left(X_{T+h}^* = n\right)\right\}_{i=1}^{B^n}. \quad (3.5)$$

Note that the number of elements of $n+1$ sets are different, and theses number are not B_3. The reason is that there is not the bootstrap sample proportion of particular value depending on the selected parameter estimates $\hat{\boldsymbol{\vartheta}}^* = (\hat{\pi}^*, \hat{\rho}^*, \hat{\phi}_1^*, \ldots, \hat{\phi}_p^*)$ in Step 6. The endpoints of the $100 \times (1-\alpha)\%$ confidence intervals for the estimated probabilities of the forecast distribution are given by quantiles of each of these $n+1$ sets.

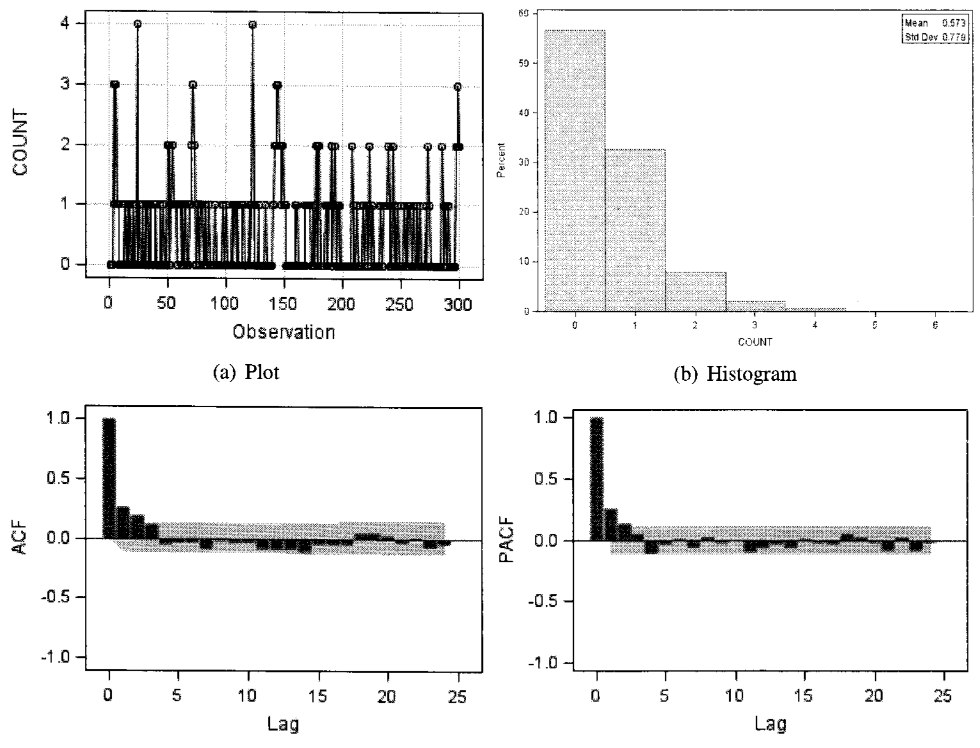

Figure 2: *Time series plot, histogram, sample ACF and PACF of the access counts: (a) Time series plot; (b) histogram; (c) sample ACF; (d) sample PACF.*

4. Empirical Application

In this section, we will apply the methods in Section 3 to the data used in Weiß (2009b) who originally introduced binomial AR(p) model. The data are the number of access, say X_t, to the home directory of six sever of the Department of Statistics of the University of Würzburg for each minute. The analyzed data are collected on 29^{th} day of November, 2005 and consist of 661 observations. Evidently, X_t take value in $\{0, 1, \ldots, 6\}$.

We divide the whole data set into two different parts as modelling and forecasting. The set $\{X_t : 357 \leq t \leq 656\}$ is used in modelling, and the set $\{X_t : 657 \leq t \leq 661\}$ is used in forecasting. Therefore, we use 300 observations as sample sizes, *i.e.*, $T = 300$, and the prediction horizon h are $h = 1, \ldots, 10$.

The Figure 2 displays histogram, plot, SACF and SPACF of the 300 observations. Although the possible range of X_t is from 0 to 6, the observed counts are composed of 56.67% 0, 32.67% 1, 8.00% 2, 2.00% 3 and 0.67% 4. To begin with we consider both BiAR(2)-Ind model and BiAR(3)-Ind model, and we will obtain YW estimates for both models by solving (3.1), and will estimate asymptotic standard errors of YW estimates by bootstrap method described in Section 3. And we will choose one between the two, and will compute the forecast distribution for the selected model. For the block length l, we choose a fixed number 10 after a careful examination of Table 8.2 in Davison and Hinkley (1997). We set $B_1 = 1000$ in Step 5, $B_2 = 1499$ in Step 7 and $B_3 = 100$ in Step 8, and all Monte Carlo simulations are performed using the IML procedure in SAS version 9.2.

Table 1 shows that the YW estimates and the estimated values of the bootstrap standard errors for

Table 1: YW estimates and the estimated values of the bootstrap standard errors for the YW estimates for Binomial AR(2) and Binomial AR(3) model

Parameter	$p=2$		$p=3$	
	YW Estimate	Bootstrap Std. Err.†	YW Estimate	Bootstrap Std. Err.†
π	0.0955556	0.0095526	0.0955556	0.0097945
ρ	0.3594186	0.1141124	0.3868873	0.1119591
ϕ_1	0.6091725	0.1524441	0.5503524	0.1563003
ϕ_2	0.3908275	0.1524441	0.3388118	0.1269084
ϕ_3			0.1108358	0.1390467

†: $\sqrt{\sum_{b=1}^{B} 1/(B-1)(\hat{\theta}^{*(b)} - \bar{\hat{\theta}}^{*})^2}$, where $\bar{\hat{\theta}}^{*} = 1/B \sum_{b=1}^{B} \hat{\theta}^{*(b)}$, with $\hat{\theta}^{*(b)}$ the bootstrap YW estimator evaluated on the b^{th} bootstrap replication and $B = 1000$

Table 2: Results of median forecasts for the 100 simulation runs

h	Binomial AR(2) $\widetilde{Med}$	Observed Value
$h=1$	1:89%, 2:11%	1
$h=2$	0:23%, 1:77%	1
$h=3$	0:60%, 1:40%	0
$h=4$	0:89%, 1:11%	1
$h=5$	0:89%, 1:11%	0
$h=6$	0:89%, 1:11%	
$h=7$	0:89%, 1:11%	
$h=8$	0:89%, 1:11%	
$h=9$	0:89%, 1:11%	
$h=10$	0:89%, 1:11%	

Table 3: 95% CI for estimated probabilities of the forecast distribution based on Binomial AR(2)

h	0	1	2	3	4	5	6
$h=1$	(7.805, 46.498)	(31.021, 44.630)	(12.675, 41.027)	(1.735, 17.612)	(0.067, 2.536)	(0.067, 0.234)	.
$h=2$	(15.811, 53.436)	(34.223, 43.296)	(9.006, 35.157)	(0.867, 8.539)	(0.067, 1.067)	(0.067, 0.133)	.
$h=3$	(24.016, 59.306)	(31.955, 41.628)	(7.005, 27.018)	(0.600, 8.072)	(0.067, 1.067)	(0.067, 0.200)	.
$h=4$	(29.953, 60.374)	(31.421, 42.829)	(6.738, 22.548)	(0.534, 5.670)	(0.067, 0.867)	(0.067, 0.133)	0.067
$h=5$	(34.089, 60.774)	(30.954, 41.695)	(6.805, 19.480)	(0.467, 4.870)	(0.067, 0.734)	(0.067, 0.133)	0.067
$h=6$	(37.825, 61.641)	(30.954, 40.827)	(6.538, 17.745)	(0.534, 4.003)	(0.067, 0.600)	(0.067, 0.133)	.
$h=7$	(39.760, 61.975)	(30.487, 40.827)	(5.670, 16.211)	(0.600, 3.869)	(0.067, 0.600)	(0.067, 0.133)	.
$h=8$	(41.628, 61.107)	(31.154, 40.027)	(6.204, 15.944)	(0.534, 3.402)	(0.067, 0.400)	0.067	.
$h=9$	(42.962, 61.508)	(30.487, 39.293)	(6.404, 15.143)	(0.534, 3.736)	(0.067, 0.467)	0.067	.
$h=10$	(43.162, 61.241)	(30.554, 39.093)	(6.204, 15.143)	(0.534, 2.802)	(0.067, 0.400)	0.067	.

the YW estimates. We note that all estimated parameters in BiAR(2)-Ind model are significant, but $\hat{\phi}_3$ in BiAR(3)-Ind model is not, therefore, we will present h-step ahead estimated forecast distribution for the BiAR(2)-Ind model for the access counts data.

Table 2 present results of the median forecasts from the $B_3 = 100$ simulation in the second column and the observed value in the third column. Each median value is obtained from the $B_2 = 1499$ forecast values, so it is the 750^{th} ranked forecast value and $B_2 = 1499$ forecast values are generated from randomly selected parameter estimates vector. For $h = 1$, 89% of 100 median forecasts are 1 and 11% are 2, so the mode of the estimated sampling distribution of the median forecasts from $B_3 = 100$ replications is 1 and this value coincide with the observed count value. For $h = 2, 3, 5$, the mode of $B_3 = 100$ replications coincide with the observed count value.

Figure 3 displays the box-whisker plots of one, two, three, four, five and ten ahead estimated forecast distributions for the BiAR(2)-Ind model for the access counts. It is based on taking the estimated probabilities for each value $\{0, 1, 2, \ldots, 6\}$ from each of the $B_3 = 100$ estimates of the forecast distribution. And Table 3 gives more information, that is, 95% confidence intervals(CI) for

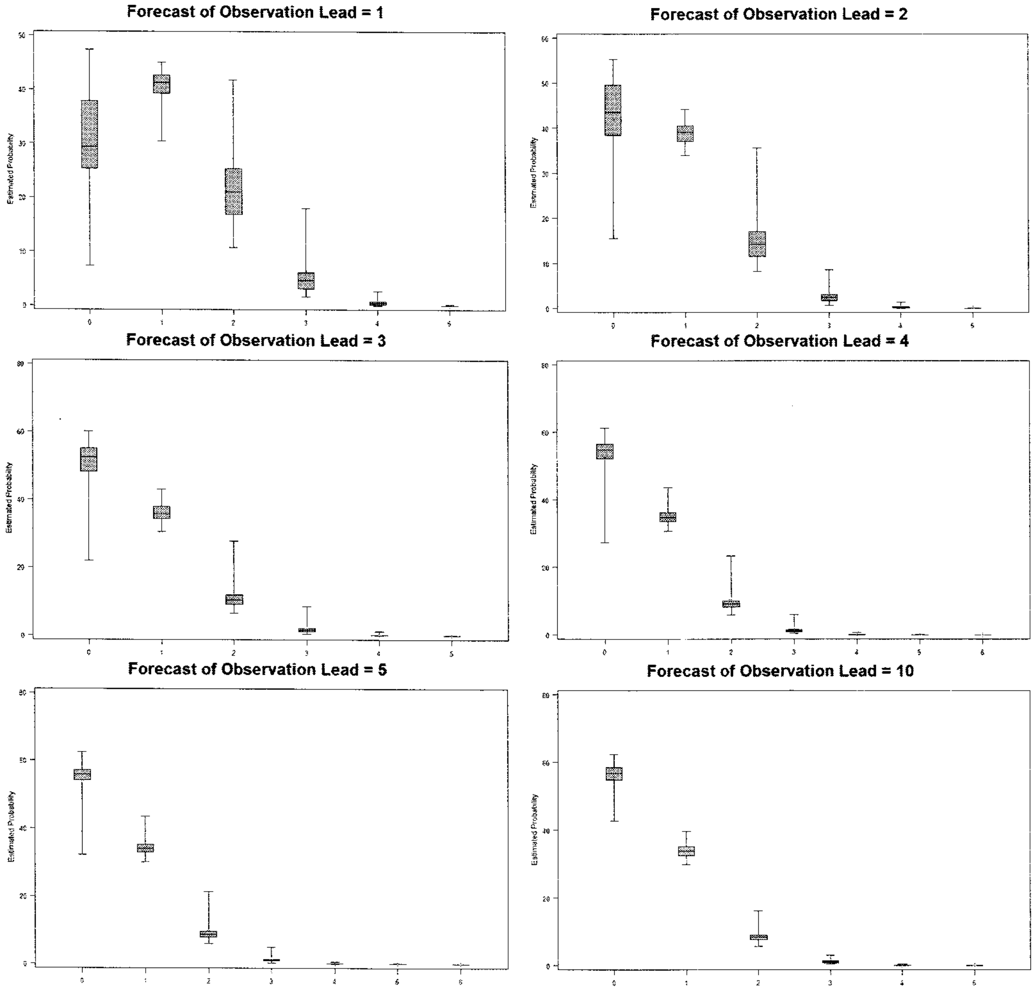

Figure 3: *Box-Whisker plots of one, two, three, four, five and ten ahead estimated forecast distributions for the Binomial AR(2) model for the access counts.*

the estimated probabilities at the different mass points. Notice that as lead time h gets greater, the CI intervals for the estimated probabilities at the different mass points are similar. The reason is that since BiAR(2)-Ind model is stationary, the forecast distributions approach to the marginal distributions. Also we can note that the observed frequencies, 56.67% 0, 32.67% 1, 8.00% 2, 2.00% 3, are within the corresponding 95% CI limits at $h = 10$.

5. Conclusion

Since the integer-valued ARMA(INARMA) models using thinning operations have a correlation structure similar to standard ARMA models, and may be interpreted as a queue, a birth and death process or a branching process with immigration, INARMA models have great appeal for modeling time series of counts. Recently Weiß (2009a) and Weiß (2009b) proposed autoregressive models for time series of binomial counts which has a special case having usual AR(p)-like autocorrelation structure.

In this paper, we focused on the methods forecasting procedures in a binomial AR(p) model, in particular coherent forecasting procedures. Our approach relies on the method introduced by Jung and Tremayne (2006). We estimated the forecast distributions of future values of binomial AR(p) models by means of a Monte Carlo method allowing for parameter uncertainty. We employed it to real data set which are the number of access to the home directory of six sever of the Department of Statistics of the University of Würzburg for each minute. The result demonstrated its usefulness in binomial AR(p) model.

Acknowledgement

The author is grateful to Christian H. Weiß for providing the dataset used in Section 4.

References

Al-Osh, M. A., Aly, E. -E. and A. A. (1992). First-order autoregressive time series with negative binomial and geometric marginals, *Communications in Statistics-Theory and Methods,* **21**, 2483–2492.

Al-Osh, M. A. and Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process, *Journal of Time Series Analysis,* **8**, 261–275.

Alzaid, A. A. and Al-Osh, M. A. (1988). First-order integer-valued autoregressive process: distributional and regression properties, *Statistica Neerlandica,* **42**, 53–61.

Alzaid, A. A. and Al-Osh, M. A. (1990). An integer-valued p^{th}-order autoregressive structure (INAR (p)) process, *Journal of Applied Probability,* **27**, 314–324.

Böckenholt, U. (1999). Mixed INAR(1) Poisson regression models: Analyzing heterogeneity and serial dependencies in longitudinal count data, *Journal of Econometricsc,* **89**, 317–338.

Brännäs, K. and Hellström, J. (2001). Generalized integer-valued autoregression, *Econometric Reviews,* **20**, 425–443.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application,* Cambridge University Press, Cambridge.

Du, J. -G. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model, *Journal of Time Series Analysis,* **12**, 129–142.

Freeland, R. and McCabe, B. P. M. (2004). Forecasting discrete valued low count time series, *International Journal of Forecasting,* **20**, 427–434.

Jung, R. and Tremayne, A. (2006). Coherent forecasting in integer time series models, *International Journal of Forecasting,* **22**, 223–238.

Kim, H. -Y. and Park, Y. (2006a). Prediction mean squared error of the poisson inar(1) process with estimated parameters, *Journal of the Korean Statistical Society,* **35**, 37–47.

Kim, H. -Y. and Park, Y. (2006b). Bootstrap confidence intervals for the INAR(1) process, *The Korean Communications in Statistics,* **13**, 343–358.

Kim, H. -Y. and Park, Y. (2008). A non-stationary integer-valued autoregressive model, *Statistical Papers,* **49**, 485-502.

Künsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations, *The Annals of Statistics,* **17**, 1217–1241.

Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process, *Journal of Time Series Analysis,* **19**, 439–455.

McKenzie, E. (1985). Some simple models for discrete variate series, *Water Resources Bulletin,* **21**, 645–650.

Park, Y., Choi J. W. and Kim, H.-Y. (2006). Forecasting cause-age specific mortality using two random processes, *Journal of the American Statistical Association,* **101**, 472–483.

Steutel, F. W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability, *The Annals of Probability,* **7**, 893–899.

Tay, A. S. and Wallis, K. F. (2000). Density forecasting: A survey, *Journal of Forecasting,* **19**, 235–254.

Weiß, C. H. (2009a). Monitoring correlated processes with binomial marginals, *Journal of Applied Statistics,* **36**, 391–414.

Weiß, C. H. (2009b). A new class of autoregressive models for time series of binomial counts, *Communications in Statistics - Theory and Methods,* **38**, 447–460.

Received November 2009; Accepted December 2009