Available online at www.sciencedirect.com
SCIENCE DIRECT®

Statistics & Probability Letters 74 (2005) 77–88

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Least-squares estimation for bifurcating autoregressive processes

J. Zhou, I.V. Basawa*

Department of Statistics, University of Georgia, Athens, Georgia 30602, USA

Received 10 March 2004; received in revised form 17 April 2005
Available online 13 May 2005

Abstract

Bifurcating autoregressive processes are used to model each line of descent in a binary tree as a standard AR(p) process, allowing for correlations between nodes which share the same parent. Limit distributions of the least-squares estimators of the model parameters for a pth-order bifurcating autoregressive process (BAR(p)) are derived. An application to bifurcating integer-valued autoregression is given. A Poisson bifurcating model is introduced.

Keywords: Cell lineage data; Tree-indexed time series; Bifurcating autoregression; Least-squares estimators; Limit distributions; Integer-valued autoregression

1. Introduction

Bifurcating autoregressive models were introduced by Cowan and Staudte (1986) for cell lineage data where each individual in one generation gives rise to two offspring in the next generation. The Cowan–Staudte model views each line of descent as a first-order autoregressive (AR(1)) process with the added complication that the observations on the two sister cells who share the same parent are allowed to be correlated. Staudte et al. (1996) studied data sets in which the observed correlations between cousin cells were significant, thus necessitating higher-order

*Corresponding author. Tel.: +1 706 542 5232; fax: +1 706 542 3391.
E-mail address: ishwar@stat.uga.edu (I.V. Basawa).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2005.04.024

models. Huggins and Basawa (1999) proposed bifurcating ARMA(p,q) models to accommodate this extended dependence in the family tree. Huggins and Basawa (2000) discussed maximum likelihood estimation for a Gaussian bifurcating AR(p) process and established the consistency and asymptotic normality of the maximum likelihood estimators of the model parameters. Recently, Basawa and Zhou (2004) introduced non-Gaussian bifurcating autoregressive models and studied some preliminary estimation problems. Zhou and Basawa (2004) have discussed maximum likelihood estimation for an exponential bifurcating AR(1) process. In this paper, we consider the asymptotic properties of the least-squares (LS) estimators of parameters in a bifurcating AR(p) (BAR(p)) process.

The rest of the paper is organized as follows. The BAR(p) model and the LS estimators of the model parameters are presented in Section 2. The limit distributions of the LS estimators are derived in Section 3. Section 4 is concerned with an application to a bifurcating integer-valued AR(1) process. A Poisson bifurcating model is introduced in Section 5.

2. Least-squares estimation for BAR(p) processes

The pth-order bifurcating autoregressive process (BAR(p)) is defined by the equation

$$X_t = \phi_0 + \phi_1 X_{[t/2]} + \phi_2 X_{[t/4]} + \cdots + \phi_p X_{[t/2^p]} + \varepsilon_t, \tag{2.1}$$

where $\{(\varepsilon_{2t}, \varepsilon_{2t+1})\}$ is a sequence of independent identically distributed (i.i.d.) bivariate random variables with $E(\varepsilon_{2t}) = E(\varepsilon_{2t+1}) = 0$, $Var(\varepsilon_{2t}) = Var(\varepsilon_{2t+1}) = \sigma^2$, and $Corr(\varepsilon_{2t}, \varepsilon_{2t+1}) = \rho$. The notation $[u]$ denotes the largest integer less than or equal to u. As in Huggins and Basawa (1999), the bifurcating operator b is defined by

$$b^r u_t = \begin{cases} u_{[t/2^r]^*} & \text{if } t>0, \\ u_{t-r} & \text{if } t<0, \end{cases}$$

where $[t/2^r]^* = [t/2^r]$ if $(t/2^r) \geqslant 1$, and $[t/2^r]^* = [\log_2(t/2^r)] + 1$ if $(t/2^r) < 1$. This notation implies that the descendants of the initial cell are labeled according to their position in the binary tree and the ancestors of the initial cell are labeled $0, -1, -2, \ldots$. The BAR(p) process in (2.1) can then be represented as

$$\phi(b) X_t = \varepsilon_t + \phi_0, \tag{2.2}$$

where $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$. We assume that the roots of $\phi(z) = 0$ are greater than 1 in absolute value, so that we can write

$$X_t = \sum_{j=0}^{\infty} (\varepsilon_{[t/2^j]^*} + \phi_0)\psi_j, \tag{2.3}$$

where $\{\psi_j\}$ are the coefficients of z^j in the expansion of $\phi^{-1}(z)$. Moreover, $\sum_{j=0}^{\infty} |\psi_j| < \infty$. The coefficients ψ_j can be determined recursively as by Huggins and Basawa (1999). The autocovariances $Cov(X_t, X_s)$ are determined as discussed by Huggins and Basawa (1999).

In particular, it is seen that

$$E(X_t) = \mu = \phi_0 \sum_{j=0}^{\infty} \psi_j = \phi_0 \left(1 - \sum_{i=1}^{p} \phi_i\right)^{-1}, \tag{2.4}$$

$$Var(X_t) = \gamma(0) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2,$$

$$Cov(X_t, X_{[t/2^k]^*}) = \gamma(k) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}, \quad k \geqslant 0. \tag{2.5}$$

Huggins and Basawa (2000) have discussed the consistency and asymptotic normality of the maximum likelihood estimators of the parameters in a BAR(p) process assuming Gaussian errors. Here, we consider the asymptotic properties of the LS estimators of $\phi = (\phi_0, \phi_1, \ldots, \phi_p)'$ and σ^2 without imposing any specific distributional assumption on $\{\varepsilon_t\}$. Let $Y_t = (1, X_{[t/2]}, \ldots, X_{[t/2^p]})^1$, $t \geqslant 2^p$. Then the LS estimator $\hat{\phi}$ of ϕ based on the observations $\{X_t, t = 2^p, 2^p+1, \ldots, n\}$ is seen to be

$$\hat{\phi} = \left(\sum_{t=2^p}^{n} Y_t Y_t'\right)^{-1} \sum_{t=2^p}^{n} Y_t X_t. \tag{2.6}$$

Define

$$\hat{\sigma}^2 = \frac{1}{(n - 2^P - p)} \sum_{t=2^p}^{n} (X_t - Y_t'\hat{\phi})^2. \tag{2.7}$$

We will derive the limit distributions of $\hat{\phi}$ and $\hat{\sigma}^2$ in the next section. A consistent estimator of ρ is given by

$$\hat{\rho} = \hat{\sigma}^{-2} \Sigma (X_{2t} - Y_{2t}'\hat{\phi})(X_{2t+1} - Y_{2t+1}'\hat{\phi}).$$

3. Limit distributions

Consider the following conditions:

(C.1) All the roots of $\phi(z) = 0$ are greater than 1 in absolute value.
(C.2) $E(\varepsilon_t^4) < \infty$, for all t.

Lemma 3.1. *Under* (C.1), *we have, as* $n \to \infty$,

(i) $(1/n)\sum_{t=1}^{n} X_t \xrightarrow{p} \mu$,
(ii) $(1/n)\sum_{t=1}^{n} (X_t - \mu)^2 \xrightarrow{p} \gamma(0)$,
(iii) $(1/n)\sum_{t=1}^{n} (X_t - \mu)(X_{[t/2^k]^*} - \mu) \xrightarrow{p} \gamma(k)$, *for* $k \geqslant 0$, *where* μ *and* $\gamma(k)$ *are defined in* (2.4) *and* (2.5), *respectively*.

Proof. Note that $\{\varepsilon_{[t/2^j]^*}\}, j=0,1,2,\ldots,$ are i.i.d. random variables with mean 0 and variance σ^2. The results then follow, via (2.3), as shown by Huggins and Basawa (2000). Also, see Brockwell and Davis (1987). □

Define $Z_t=(1,X_t,X_{[t/2]},\ldots,X_{[t/(2^{p-1})]})'$, and let $m=(n-1)/2=$ the number of triplets (X_t,X_{2t},X_{2t+1}) observed. We then have the following:

Lemma 3.2. *Under* (C.1)

$$\frac{1}{m}\sum_{t=2^{p-1}}^{m} Z_tZ_t' \xrightarrow{\text{p}} A \quad \textit{as } m\to\infty, \tag{3.1}$$

where A is a $(p+1)\times(p+1)$ matrix defined by

$$A=\begin{pmatrix} 1 & \mu & \mu & \ldots & \mu \\ \mu & a(0) & a(1) & \ldots & a(p-1) \\ \mu & a(1) & a(0) & \ldots & a(p-2) \\ \vdots & & & & \\ \mu & a(p-1) & a(p-2) & \ldots & a(0) \end{pmatrix}, \tag{3.2}$$

with μ defined in (2.4), $a(k)=\mu^2+\gamma(k)$, *and $\gamma(k)$ given by* (2.5).

Proof. The result follows from Lemma 2.1 after noting that

$$\Sigma Z_tZ_t'=\begin{pmatrix} m & \Sigma X_t & \Sigma X_{[t/2]} & \ldots & \Sigma X_{[t/(2^{p-1})]} \\ \Sigma X_t & \Sigma X_t^2 & \Sigma X_tX_{[t/2]} & \ldots & \Sigma X_tX_{[t/(2^{p-1})]} \\ \Sigma X_{[t/2]} & \Sigma X_{[t/2]}X_t & \Sigma X_{[t/2]}^2 & \ldots & \Sigma X_{[t/2]}X_{[t/(2^{p-1})]} \\ \vdots & & & & \\ \Sigma X_{[t/(2^{p-1})]} & \Sigma X_{[t/(2^{p-1})]}X_t & \Sigma X_{[t/(2^{p-1})]}X_{[t/2]} & \ldots & \Sigma X_{[t/(2^{p-1})]}^2 \end{pmatrix}. \qquad \square$$

The following version of the martingale central limit theorem will be used in the derivation of the limit distribution of the LS estimator.

Lemma 3.3. *Let $\{Y_t\}$, $t=1,2,\ldots,$ be a sequence of zero-mean vector martingale differences satisfying the following conditions*:

(a) $E(Y_tY_t')=\Omega_t$, *a positive definite matrix, and* $1/n\sum_{t=1}^n\Omega_t\to\Omega$, *a positive definite matrix.*
(b) $E(Y_{it}Y_{jt}Y_{lt}Y_{mt})<\infty$ *for all t, and all i,j,l,m, where Y_{rt} denotes the rth element of the vector Y_t.*
(c) $(1/n)\sum_{t=1}^n Y_tY_t'\xrightarrow{\text{d}}\Omega$.

Then, $(1/\sqrt{n})\sum_{t=1}^n Y_t\xrightarrow{\text{d}}\mathrm{N}(0,\Omega)$.

Proof. See, for instance, Proposition 7.9 in Hamilton (1994). □

Lemma 3.4. *Under* (C.1) *and* (C.2), *as* $m \to \infty$,

$$\frac{1}{\sqrt{m}} \sum_{t=2^{p-1}}^{m} Z_t V_t \xrightarrow{\mathrm{d}} \mathrm{N}(0, \sigma^2(1+\rho)A),$$

where A *is defined in Lemma* 3.2, *and* $V_t = (1/\sqrt{2})(\varepsilon_{2t} + \varepsilon_{2t+1})$.

Proof. Let $\mathscr{F}_t = \sigma\{\varepsilon_j : j \leqslant 2t+1\}$. It can be verified that $\sum_{t=2^{p-1}}^{m} Z_t V_t$ is a zero-mean martingale with respect to $\mathscr{F}_t$. In order to verify the central limit theorem for martingales, we now check the conditions of Lemma 3.3.

(a) From (2.5), we have $E(Z_t Z_t' V_t^2) = E(Z_t Z_t') E(V_t^2) = A\sigma^2(1+\rho)$, where A is defined in Lemma 3.2. It can be verified that A is a positive definite matrix. Hence, condition (a) is satisfied.
(b) $E(V_t^4 Z_{it} Z_{jt} Z_{kt} Z_{lt}) < \infty$, for all i, j, k, l, where Z_{rt} is the rth element of the vector Z_t. Condition (b) holds from Proposition 7.10 of Hamilton (1994) under assumption (C.2).
(c) $(1/m)\sum_{t=2^{p-1}}^{m} V_t^2 Z_t Z_t' \xrightarrow{\mathrm{p}} \sigma^2(1+\rho)A$. In order to verify (c), consider

$$\begin{aligned} \frac{1}{m} \sum_{t=2^{p-1}}^{m} V_t^2 Z_t Z_t' &= \frac{1}{m} \sum_{t=2^{p-1}}^{m} [V_t^2 - \sigma^2(1+\rho)] Z_t Z_t' + \sigma^2(1+\rho) \frac{1}{m} \sum_{t=2^{p-1}}^{m} Z_t Z_t' \\ &= U_{1m} + U_{2m}, \quad \text{say.} \end{aligned}$$

We have $U_{1m} = (1/m)\sum_{t=2^{p-1}}^{m} W_t$, where $W_t = (V_t^2 - \sigma^2(1+\rho)) Z_t Z_t'$. For any $(p+1)$-vector λ, we have $\lambda' U_{1m} \lambda = (1/m)\sum_{t=2^{p-1}}^{m} \lambda' W_t \lambda$. It is easily verified that $E(\lambda' W_t \lambda | \mathscr{F}_{t-1}) = 0$, and $\{\lambda' W_t \lambda\}$ is a stationary martingale difference sequence with $E(\lambda' W_t \lambda)^2 < \infty$ (see (b)). Consequently, by the law of large numbers for martingales (see Hall and Heyde (1980)) we conclude that $\lambda' U_{1m} \lambda \xrightarrow{\mathrm{p}} 0$, and hence $U_{1m} \xrightarrow{\mathrm{p}} 0$.

From Lemma 3.2, $(1/m)\Sigma Z_t Z_t' \xrightarrow{\mathrm{p}} A$, and hence $U_{2m} \xrightarrow{\mathrm{p}} \sigma^2(1+\rho)A$. Consequently, condition (c) is verified. The desired limit in Lemma 3.4 then follows from Lemma 3.3. □

The limit distribution of $\hat{\phi}$ is given below.

Theorem 3.1. *Under* (C.1) *and* (C.2), *we have*

$$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \sigma^2(1+\rho)A^{-1}) \quad \text{as } n \to \infty.$$

Proof. We have

$$\begin{aligned} \sqrt{n}(\hat{\phi} - \phi) &= \left(\frac{1}{n} \sum_{t=2^p}^{n} Y_t Y_t' \right)^{-1} \left[\frac{1}{\sqrt{n}} \sum_{t=2^p}^{n} Y_t \varepsilon_t \right] \\ &= \left(\frac{1}{m} \sum_{t=2^{p-1}}^{m} Z_t Z_t' \right)^{-1} \left[\frac{1}{\sqrt{m}} \sum_{t=2^{p-1}}^{m} Z_t V_t \right] + o_p(1). \end{aligned}$$

The result then follows from Lemmas 3.2, 3.4 and Slutsky's theorem. □

Remark. It may be noted that if the errors $\{(\varepsilon_{2t}, \varepsilon_{2t+1})\}$ are bivariate normal, the limit distribution of the LS estimator $\hat{\phi}$ is the same as that of the maximum likelihood estimator derived by Huggins and Basawa (2000).

The next theorem gives the limit distribution of $\hat{\sigma}^2$.

Theorem 3.2. *Under* (C.1) *and* (C.2), *we have, as* $n \to \infty$,

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{\mathrm{d}} \mathrm{N}(0, u_4 + u_{22} - 2\sigma^4),$$

where $u_4 = E(\varepsilon_t^4)$ *and* $u_{22} = E(\varepsilon_{2t}^2 \varepsilon_{2t+1}^2)$.

Proof. We have

$$\begin{aligned}
\sum_{t=2^p}^{n} (X_t - Y_t'\hat{\phi})^2 &= \sum_{t=2^p}^{n} (X_t - Y_t'\phi - Y_t'(\hat{\phi} - \phi))^2 \\
&= \sum_{t=2^p}^{n} \varepsilon_t^2 - 2(\hat{\phi} - \phi)' \sum_{t=2^p}^{n} Y_t \varepsilon_t + (\hat{\phi} - \phi)' \left(\sum_{t=2^p}^{n} Y_t Y_t' \right) (\hat{\phi} - \phi) \\
&= \sum_{t=2^p}^{n} \varepsilon_t^2 - (\hat{\phi} - \phi)' \left(\sum_{t=2^p}^{n} Y_t Y_t' \right) (\hat{\phi} - \phi).
\end{aligned}$$

Hence,

$$\begin{aligned}
\sqrt{n}(\hat{\sigma}^2 - \sigma^2) &\simeq \frac{1}{\sqrt{n}} \sum_{t=2^p}^{n} (\varepsilon_t^2 - \sigma^2) - \sqrt{n}(\hat{\phi} - \phi)' \left(\frac{1}{n} \sum_{t=2^p}^{n} Y_t Y_t' \right) (\hat{\phi} - \phi) \\
&= W_{1n} + W_{2n}, \quad \text{say.}
\end{aligned}$$

Note that $W_{2n} \xrightarrow{\mathrm{p}} 0$, since $(1/n)\sum_{t=2^p}^{n} Y_t Y_t' \xrightarrow{\mathrm{p}} A$, and $\sqrt{n}(\hat{\phi} - \phi) = \mathrm{O_p}(1)$.

We have

$$\begin{aligned}
W_{1n} = \frac{1}{\sqrt{n}} \sum_{t=2^p}^{n} (\varepsilon_t^2 - \sigma^2) &\simeq \frac{1}{\sqrt{m}} \sum_{t=2^{p-1}}^{m} \left(\frac{\varepsilon_{2t}^2 + \varepsilon_{2t+1}^2 - 2\sigma^2}{\sqrt{2}} \right) \\
&\xrightarrow{\mathrm{d}} \mathrm{N}(0, u_4 + u_{22} - 2\sigma^4).
\end{aligned}$$

This completes the proof. □

The limit distribution of $\hat{\rho}$ can be obtained in a similar manner which is omitted. We now illustrate Theorem 3.1 by two examples.

Example 1. BAR(1) model

Consider the model

$$X_t = \phi_0 + \phi_1 X_{[t/2]} + \varepsilon_t, \quad \phi_0 \neq 0 \text{ and } |\phi_1| < 1.$$

The LS estimators are given by

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{m} U_t (X_t - \bar{X})}{\sum_{t=1}^{m} (X_t - \bar{X})^2} \quad \text{where } U_t = \frac{\varepsilon_{2t} + \varepsilon_{2t+1}}{2} \text{ and } \bar{X} = \frac{1}{m} \sum_{t=1}^{m} X_t,$$

$$\hat{\phi}_0 = \bar{U} - \hat{\phi}_1 \bar{X} \quad \text{where } \bar{U} = \frac{1}{m} \sum_{t=1}^{m} U_t.$$

From Theorem 3.1, we have

$$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{\text{d}} \mathrm{N}(0, \sigma^2(1+\rho)A^{-1}),$$

where

$$A = \begin{pmatrix} 1 & \phi_0/(1-\phi_1) \\ \phi_0/(1-\phi_1) & \sigma^2/(1-\phi_1^2) + (\phi_0/(1-\phi_1))^2 \end{pmatrix}.$$

If $\phi_0 = 0$, we have $\hat{\phi}_1 = \sum_{t=1}^m U_t X_t / \sum_{t=1}^m X_t^2$, and $A = EX_t^2 = \sigma^2/(1-\phi_1^2)$. Consequently, we have, for $\phi_0 = 0$,

$$\sqrt{n}(\hat{\phi}_1 - \phi_1) \xrightarrow{\text{d}} \mathrm{N}(0, (1+\rho)(1-\phi_1^2)).$$

Example 2. BAR(2) model

For the model

$$X_t = \phi_0 + \phi_1 X_{[t/2]} + \phi_2 X_{[t/4]} + \varepsilon_t,$$

we have under (C.1) and (C.2),

$$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{\text{d}} \mathrm{N}(0, \sigma^2(1+\rho)A^{-1}),$$

where

$$A = \begin{pmatrix} 1 & \mu & \mu \\ \mu & a(0) & a(1) \\ \mu & a(1) & a(0) \end{pmatrix}.$$

In particular, when $\phi_0 = 0$ and $\phi = (\phi_1, \phi_2)'$, we have

$$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{\text{d}} \mathrm{N}(0, (1+\rho)B),$$

where

$$B = \begin{pmatrix} 1-\phi_2^2 & -\phi_1(1-\phi_2) \\ -\phi_1(1-\phi_2) & 1-\phi_2^2 \end{pmatrix}.$$

3.1. Mean-centered process

We now consider the mean-centered version of the model in (2.1). Model (2.1) can be rewritten as

$$X_t - \mu = \phi_1(X_{[t/2]} - \mu) + \phi_2(X_{[t/2^2]} - \mu) + \cdots + \phi_p(X_{[t/2^p]} - \mu) + \varepsilon_t, \tag{3.3}$$

where $\mu = \phi_0(1 - \sum_{i=1}^p \phi_i)^{-1}$.

Define

$$\hat{\mu} = \hat{\phi}_0 \left(1 - \sum_{i=1}^p \hat{\phi}_i\right)^{-1}. \tag{3.4}$$

Let $\beta = (\mu, \phi_1, \phi_2, \ldots, \phi_p)'$. We then have

$$(\hat{\beta} - \beta) = D(\hat{\phi} - \phi) + o_p(1), \tag{3.5}$$

where $\phi = (\phi_0, \phi_1, \ldots, \phi_p)'$,

$$D = \begin{pmatrix} c & c\mu & c\mu & \ldots & c\mu \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{3.6}$$

and $c = (1 - \sum_{i=1}^{p} \phi_i)^{-1}$. The limit distribution of $\hat{\beta}$ is given next.

Theorem 3.3. *Under* (C.1) *and* (C.2), *we have*

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathrm{p}} \mathrm{N}(0, \sigma^2(1+\rho)DA^{-1}D') \quad \text{as } n \to \infty,$$

where A is defined in (3.2) *and D in* (3.6).

Proof. The result follows from Theorem 3.1 and (3.5). $\square$

Remark. It is easily verified that

$$DA^{-1}D' = \begin{pmatrix} c^2 & \mathbf{0} \\ \mathbf{0} & \Gamma^{-1} \end{pmatrix}, \tag{3.7}$$

where

$$\Gamma = \begin{pmatrix} \gamma(0) & \gamma(1) & \ldots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \ldots & \gamma(p-2) \\ \vdots & & & \\ \gamma(p-1) & \gamma(p-2) & \ldots & \gamma(0) \end{pmatrix}. \tag{3.8}$$

In order to check (3.7), first note that

$$A = \begin{pmatrix} 1 & \mu\boldsymbol{u}' \\ \mu\boldsymbol{u} & \Gamma + \mu^2\boldsymbol{u}\boldsymbol{u}' \end{pmatrix} = P'\Sigma P,$$

where $\boldsymbol{u} = (1, 1, \ldots, 1)'$ is a $(p \times 1)$ unit vector,

$$P = \begin{pmatrix} 1 & \mu\boldsymbol{u}' \\ \mathbf{0} & I \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \Gamma \end{pmatrix}.$$

Also,

$$D = \begin{pmatrix} c & c\mu\boldsymbol{u}' \\ \mathbf{0} & I \end{pmatrix} = QP,$$

where

$$Q = \begin{pmatrix} c & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix}.$$

We thus have

$$DA^{-1}D' = (QP)(P'\Sigma P)^{-1}(QP)' = Q\Sigma^{-1}Q' = \begin{pmatrix} c^2 & \mathbf{0} \\ \mathbf{0} & \Gamma^{-1} \end{pmatrix}. \tag{3.9}$$

Hence, the result in (3.7) is verified.

It then follows that

$$\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{\text{d}} \mathrm{N}(0, c^2\sigma^2(1+\rho)),$$

and

$$\sqrt{n}(\hat{\phi}^* - \phi^*) \xrightarrow{\text{d}} \mathrm{N}(0, \sigma^2(1+\rho)\Gamma^{-1}),$$

where $\phi^* = (\phi_1, \phi_2, \ldots, \phi_p)'$. Moreover, $\hat{\mu}$ is asymptotically independent of $\hat{\phi}^*$. It can further be noted that

$$A^{-1} = P^{-1}\Sigma^{-1}(P^{-1})' = \begin{pmatrix} 1 + \mu^2\boldsymbol{u}'\Gamma^{-1}\boldsymbol{u} & -\mu\boldsymbol{u}'\Gamma^{-1} \\ -\mu\Gamma^{-1}\boldsymbol{u} & \Gamma^{-1} \end{pmatrix}. \tag{3.10}$$

Example 1 (Continued). The centered version of the BAR(1) model is

$$X_t - \mu = \phi_1(X_{[t/2]} - \mu) + \varepsilon_t \quad \text{where } \mu = \phi_0(1-\phi_1)^{-1}.$$

It follows from Theorem 3.3 that

$$\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{\text{d}} \mathrm{N}(0, \sigma^2(1+\rho)(1-\phi_1)^{-2})$$

and

$$\sqrt{n}(\hat{\phi}_1 - \phi_1) \xrightarrow{\text{d}} \mathrm{N}(0, (1+\rho)(1-\phi_1^2)).$$

Moreover, $\hat{\mu}$ is asymptotically independent of $\hat{\phi}_1$.

4. Integer-valued bifurcating autoregressive model

In this section, we introduce an extension of the first-order integer-valued autoregression (INAR(1)) (see Al-Osh and Alzaid, 1987) to a binary tree-indexed process and discuss LS estimation for the model parameters. Consider the process $\{X_t\}$ satisfying the relation

$$X_t = \phi_1 \circ X_{[t/2]} + \varepsilon_t \quad 0 < \phi_1 < 1, \tag{4.1}$$

where $\phi_1 \circ X_{[t/2]}$ denotes the binomial thinning operation defined by

$$\phi_1 \circ X_{[t/2]} = \sum_{i=1}^{X_{[t/2]}} Y_i, \tag{4.2}$$

where $\{Y_i\}$, $i = 1, 2, \ldots,$ are i.i.d. Bernoulli random variables with $P(Y_i = 1) = \phi_1$ and $P(Y_i = 0) = 1 - \phi_1$, $0 < \phi_1 < 1$. The error process $\{\varepsilon_t\}$ is characterized by the fact that $\{(\varepsilon_{2t}, \varepsilon_{2t+1})\}$, $t = 1, 2, \ldots,$ are i.i.d. integer-valued bivariate random variables with $E(\varepsilon_{2t}) = E(\varepsilon_{2t+1}) = \phi_0$, $Var(\varepsilon_{2t}) = Var(\varepsilon_{2t+1}) = \sigma^2$ and $Corr(\varepsilon_{2t}, \varepsilon_{2t+1}) = \rho$. It is readily verified from (4.1) that

$$E(X_t | X_{[t/2]}) = \phi_0 + \phi_1 X_{[t/2]}, \quad \phi_0 > 0, \tag{4.3}$$

and

$$Var(X_t | X_{[t/2]}) = \phi_1(1 - \phi_1) X_{[t/2]} + \sigma^2. \tag{4.4}$$

The conditional least-squares (CLS) estimators of ϕ_0 and ϕ_1 are obtained by minimizing $\sum_{t=2}^{n} (X_t - \phi_0 - \phi_1 X_{[t/2]})^2$ with respect to ϕ_0 and ϕ_1, and these are the same as the LS estimators $\hat{\phi}_0$ and $\hat{\phi}_1$ for the BAR(1) model given in Example 1 in Section 3. It can be verified from (4.3) and (4.4) that the unconditional stationary moments are given by

$$\mu = E(X_t) = \phi_0(1 - \phi_1)^{-1} \tag{4.5}$$

and

$$\gamma(0) = Var(X_t) = (\mu\phi_1(1 - \phi_1) + \sigma^2)(1 - \phi_1^2)^{-1}. \tag{4.6}$$

Using basically similar arguments as those for the centered BAR(1) example at the end of Section 3, one can verify that

$$\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \sigma^2(1 + \rho)(1 - \phi_1)^{-2})$$

and

$$\sqrt{n}(\hat{\phi}_1 - \phi_1) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \sigma^2(1 + \rho)\gamma^{-1}(0))$$

where $\gamma(0)$ is given by (4.6). Moreover, $\hat{\mu}$ is asymptotically independent of $\hat{\phi}_1$. Even though some of the time series asymptotics used in the previous section are not directly applicable for the model in (4.1), one can use the fact that $\{X_t\}$ is an ergodic Markov chain (see Grunwald et al., 2000) and standard Markov chain asymptotics can then be used to establish the above results. The details are omitted.

5. Bifurcating Poisson model

As an example of the bifurcating INAR(1) model of Section 4, we present here a Poisson bifurcating model, and study some of its properties. Consider the model in (4.1) with $\{(\varepsilon_{2t}, \varepsilon_{2t+1})\}$ having a bivariate Poisson distribution defined by

$$P(\varepsilon_{2t} = y_1, \varepsilon_{2t+1} = y_2) = \mathrm{e}^{-(\theta_1+\theta_2+\theta_3)} \sum_{i=0}^{y_1 \wedge y_2} \frac{\theta_1^{y_1-i}\theta_2^{y_2-i}\theta_3^{i}}{(y_1 - i)!(y_2 - i)!i!}, \tag{5.1}$$

where $y_1 \wedge y_2 = \min(y_1, y_2)$, $\theta_i > 0$, $i = 1, 2, 3$, and $y_j = 0, 1, 2, \ldots, j = 1, 2$. The marginal distributions of ε_{2t} and ε_{2t+1} are then Poisson with means $\theta_1 + \theta_3$ and $\theta_2 + \theta_3$, respectively, and $Cov(\varepsilon_{2t}, \varepsilon_{2t+1}) = \theta_3$. The joint moment generating function of $(\varepsilon_{2t}, \varepsilon_{2t+1})$ is seen to be

$$M(t_1, t_2) = \exp[\theta_3(e^{t_1+t_2} - 1) + \theta_1(e^{t_1} - 1) + \theta_2(e^{t_2} - 1)]. \tag{5.2}$$

See, for instance, Johnson et al. (1997). We now choose the following parameterization:

$$\theta_1 = \theta_2 = (1 - \rho)\phi_0 \text{ and } \theta_3 = \rho\phi_0 \quad \text{with } 0 < \rho < 1, \ \phi_0 > 0.$$

We then get $E(\varepsilon_{2t}) = E(\varepsilon_{2t+1}) = Var(\varepsilon_{2t}) = Var(\varepsilon_{2t+1}) = \phi_0$, and $Corr(\varepsilon_{2t}, \varepsilon_{2t+1}) = \rho$.

The conditional distribution of X_t given $X_{[t/2]}$ is obtained from (4.1) and (5.1), and it is seen to be

$$p(x_t|x_{[t/2]}) = e^{-\phi_0} \sum_{i=0}^{x_t \wedge x_{[t/2]}} \frac{\phi_0^{(x_t - i)}}{(x_t - i)!} \binom{x_{[t/2]}}{i} \phi_1^i (1 - \phi_1)^{(x_{[t/2]} - i)}. \tag{5.3}$$

We have, from (4.3) and (4.4),

$$E(X_t|X_{[t/2]}) = \phi_0 + \phi_1 X_{[t/2]}$$

and

$$Var(X_t|X_{[t/2]}) = \phi_1(1 - \phi_1)X_{[t/2]} + \phi_0.$$

The CLS estimators of ϕ_0 and ϕ_1 are then obtained as discussed in Section 4.

The likelihood function is given by

$$L_n(\phi_0, \phi_1, \rho) = p(x_1) \prod_{t=1}^{m} p(x_{2t}, x_{2t+1}|x_t),$$

where m is the total number of triplets (x_t, x_{2t}, x_{2t+1}) observed, and $p(x_{2t}, x_{2t+1}|x_t)$ is the conditional distribution of (X_{2t}, X_{2t+1}) given X_t. However, $p(x_{2t}, x_{2t+1}|x_t)$ does not have a simple form. The conditional moment generating function of (X_{2t}, X_{2t+1}) given X_t is given below.

Lemma 5.1. *The conditional moment generating function of* (X_{2t}, X_{2t+1}) *given* X_t *is*

$$M_{(X_{2t}, X_{2t+1})|X_t}^{(t_1,t_2)} = [\phi_1 e^{t_1+t_2} + (1 - \phi_1)]^{X_t} M_{(\varepsilon_{2t}, \varepsilon_{2t+1})}(t_1, t_2),$$

where $M_{(\varepsilon_{2t}, \varepsilon_{2t+1})}(t_1, t_2)$ *is given by* (5.2).

Proof. We have

$$\begin{aligned} E[e^{t_1 X_{2t} + t_2 X_{2t+1}}|X_t] &= E[e^{t_1 \sum_{i=1}^{X_t} Y_i + t_1 \varepsilon_{2t} + t_2 \sum_{i=1}^{X_t} Y_i + t_2 \varepsilon_{2t+1}}|X_t] \\ &= E[e^{(t_1+t_2)\sum_{i=1}^{X_t} Y_i}|X_t] E(e^{t_1 \varepsilon_{2t} + t_2 \varepsilon_{2t+1}}) \\ &= [\phi_1 e^{t_1+t_2} + (1 - \phi_1)]^{X_t} M_{(\varepsilon_{2t}, \varepsilon_{2t+1})}(t_1, t_2), \end{aligned}$$

since conditional on X_t, $\sum_{i=1}^{X_t} Y_i$ is a binomial random variable with parameters (X_t, ϕ_1). □

Next, we obtain the unconditional joint distribution of (X_{2t}, X_{2t+1}) for the model given by (4.1) and (5.1). This turns out to be a bivariate Poisson distribution.

Lemma 5.2. *The joint distribution of* (X_{2t}, X_{2t+1}) *is a bivariate Poisson with* $E(X_{2t}) = E(X_{2t+1}) = \phi_0/(1-\phi_1)$, *and* $Cov(X_{2t}, X_{2t+1}) = (\rho + \phi_1/(1-\phi_1))\phi_0$.

Proof. The joint moment generating function of (X_{2t}, X_{2t+1}) is given by

$$\begin{aligned} M_{(X_{2t},X_{2t+1})}(t_1,t_2) &= E[M_{(X_{2t},X_{2t+1})}(t_1,t_2)|X_t] \\ &= M_{(\varepsilon_{2t},\varepsilon_{2t+1})}(t_1,t_2)E[(\phi_1 e^{t_1+t_2} + (1-\phi_1)^{X_t}]. \end{aligned} \tag{5.4}$$

Next, note that the marginal distribution of X_t is Poisson with mean $\phi_0/(1-\phi_1)$. This is seen from representing X_t in (4.1) in terms of $\{\varepsilon_{[t/2^j]}\}$, $j = 0, 1, \ldots,$

$$X_t = \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{[t/2^j]},$$

and noting that $\{\varepsilon_{[t/2^j]}\}$, $j = 0, 1, 2, \ldots,$ is a sequence of i.i.d. Poisson random variables with mean ϕ_0. Consequently,

$$\begin{aligned} E[(\phi_1 e^{t_1+t_2} + (1-\phi_1))^{X_t}] &= E[e^{sX_t}] \quad \text{where } s = \log(\phi_1 e^{t_1+t_2} + (1-\phi_1)) \\ &= \exp\left[\frac{\phi_0}{1-\phi_1}(e^s - 1)\right] = \exp\left[\frac{\phi_0}{1-\phi_1}(\phi_1 e^{t_1+t_2} - \phi_1)\right]. \end{aligned} \tag{5.5}$$

Substituting (5.5) in (5.4), and simplifying, we get the moment generating function of the bivariate Poisson distribution given in (5.2) with

$$\theta_1 = \theta_2 = (1-\rho)\phi_0 \quad \text{and} \quad \theta_3 = (\phi_1/(1-\phi_1) + \rho)\phi_0.$$

The result in the lemma then follows. □

References

Al-Osh, M.A., Alzaid, A.A., 1987. First-order integer-valued autoregressive (INAR(1)) process. J. Time Ser. Anal. 8, 261–275.

Basawa, I.V., Zhou, J., 2004. Non-Gaussian bifurcating models and quasi-likelihood estimation, to appear in J. Appl. Probab. Spec. 41A.

Brockwell, P.J., Davis, R.A., 1987. Time Series: Theory and Methods. Springer, New York.

Cowan, R., Staudte, R.G., 1986. The bifurcating autoregression model in cell lineage studies. Biometrics 42, 769–783.

Grunwald, G.K., Hyndman, R.J., Tedesco, L., Tweedie, R.L., 2000. Non-Gaussian conditional linear AR(1) models. Austral. New Zealand J. Statist. 42, 479–495.

Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and Its Applications. Academic Press, New York.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Huggins, R.M., Basawa, I.V., 1999. Extensions of the bifurcating autoregressive model for cell lineage studies. J. Appl. Probab. 36, 1225–1233.

Huggins, R.M., Basawa, I.V., 2000. Inference for the extended bifurcating autoregressive model for cell lineage studies. Austral. New Zealand J. Statist. 42, 423–432.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1997. Discrete Multivariate Distributions. Wiley, New York.

Staudte, R.G., Zhang, J., Huggins, R.M., Cowan, R., 1996. A re-examination of the cell lineage data of E.O. Powell. Biometrics 52, 1214–1222.

Zhou, J., Basawa, I.V., 2004. Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors, J. Time Ser. Anal., to appear.