Adv. Appl. Prob. **38,** 559–578 (2006)
Printed in Northern Ireland

A SIMPLE INTEGER-VALUED BILINEAR TIME SERIES MODEL

P. DOUKHAN,* *CREST*

A. LATOUR,** *Université Pierre Mendès-France*

D. ORAICHI,*** *CHU Sainte-Justine*

Abstract

In this paper, we extend the integer-valued model class to give a nonnegative integer-valued bilinear process, denoted by INBL(p, q, m, n), similar to the real-valued bilinear model. We demonstrate the existence of this strictly stationary process and give an existence condition for it. The estimation problem is discussed in the context of a particular simple case. The method of moments is applied and the asymptotic joint distribution of the estimators is given: it turns out to be a normal distribution. We present numerical examples and applications of the model to real time series data on meningitis and *Escherichia coli* infections.

Keywords: Integer-valued process; bilinear model

2000 Mathematics Subject Classification: Primary 62M10; 60G10
Secondary 62F10; 62-07

1. Introduction

As was pointed out in [22, p. 3], discrete-valued time series are commonly encountered in practice. In the last two decades, many developments have been made in this field. Consequently, tools specifically designed for discrete-valued series are now available for data analysis.

There has been a real effort to define a family of models that are structurally simple, sufficiently versatile, and also accessible. Pioneering work must be mentioned. Several articles have dealt with statistical data expressed in terms of counts taken sequentially in time and correlated. Many authors have tackled the problem of integer-valued time series analysis. Jacobs and Lewis in [16], [17], and [18] presented and applied the so-called discrete autoregressive moving average models. Some autoregressive moving average models for dependent sequences of Poisson counts were suggested in [8], [21], [23], and [24]. In [2], Alzaid and Al-Osh introduced integer-valued pth-order autoregressive (INAR(p)) models and, in [1], integer-valued qth-order moving average (INMA(q)) models. In [11], Du and Li gave the first rigorous construction of an integer-valued autoregressive process. Gauthier and Latour in [14] and Latour in [19] and [20] developed a more general version of the INAR(p) model, denoted by GINAR(p). Park and Kim in [25] studied the properties of the INMA(q) model,

Received 8 May 2001; revision received 9 March 2006.
* Postal address: Laboratoire de Statistique, CREST, Timbre J340, 3 avenue Pierre Larousse, 92240 Malakoff Cedex, France. Email address: paul.doukhan@ensae.fr
** Postal address: Laboratoire de Statistique et Analyse de Données, Université Pierre Mendès-France, Bâtiment Sciences Humaines et Mathématiques, 1251 avenue Centrale, BP 47, 38040 Grenoble Cedex 09, France.
Email address: alain.latour@upmf-grenoble.fr
*** Postal address: Centre de Recherche du CHU Sainte-Justine, 3175 chemin de la Côte-Sainte-Catherine, Montréal, QC H3T 1C5, Canada. Email address: driss.oraichi@recherche-ste-justine.qc.ca

while Dion *et al.*in [9] established links between some models used in integer-valued time series analysis and branching processes.

More recently, a simple integer-valued generalized autoregressive conditional heteroskedastic-type model of orders p and q (the INGARCH(p, q) model) has been proposed. See [13] for some results on this model and its application in epidemiology. This model was also studied in [28] and applied in finance to model the number of transactions taking place during a short interval of time. Streettin [30] has derived some stationarity results for the INGARCH(1, 1) model. Like a model proposed by Davis *et al.* in [6], this is an observation-driven model.

In this paper, as in many other ones, the Steutel–van Harn operator is used. Let us recall the definition of this operator from [14].

Definition 1.1. (*Steutel–van Harn operator.*) Let $\{Y_j\}_{j\in\mathbb{N}}$ be a sequence of independent and identically distributed nonnegative integer-valued variables with mean α and variance λ, independent of X, which is a nonnegative integer-valued variable. The Steutel–van Harn operator $\alpha\circ$ is defined by

$$\alpha \circ X = \begin{cases} \displaystyle\sum_{i=1}^{X} Y_i & \text{if } X > 0, \\ 0 & \text{otherwise.} \end{cases}$$

The sequence $\{Y_i\}_{i\in\mathbb{N}}$ is called a counting sequence. Note that, as indicated in Definition 1.1, the mean of the summands $\{Y_i\}$ associated with the operator $\alpha\circ$ is denoted by α. Suppose that $\beta\circ$ is another Steutel–van Harn operator based on a counting sequence $\{\tilde{Y}_i\}_{i\in\mathbb{N}}$. The operators $\alpha\circ$ and $\beta\circ$ are said to be independent if and only if the counting sequences $\{Y_i\}_{i\in\mathbb{N}}$ and $\{\tilde{Y}_i\}_{i\in\mathbb{N}}$ are mutually independent.

We would like to extend the integer-valued model class to give a nonnegative integer-valued bilinear process, denoted by INBL(p, q, m, n), similar to the real-valued bilinear process presented by Tong in [31, pp. 114–115]. A time series $\{X_t\}_{t\in\mathbb{Z}}$ is generated by a bilinear model if it satisfies the equation

$$X_t = \alpha + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} c_j \varepsilon_{t-j} + \sum_{k=1}^{m}\sum_{\ell=1}^{n} b_{\ell k}(\varepsilon_{t-\ell} X_{t-k}) + \varepsilon_t, \tag{1.1}$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of independent, identically distributed random variables, usually but not always with zero mean, and where α, a_i $(i = 1, \ldots, p)$, c_j $(j = 1, \ldots, q)$, and $b_{\ell k}$ $(\ell = 1, \ldots, n,\ k = 1, \ldots, m)$, are real constants. In (1.1) we can 'formally' substitute Steutel–van Harn operators for some of the parameters, giving an equation of the form

$$X_t = \sum_{i=1}^{p} a_i \circ X_{t-i} + \sum_{j=1}^{q} c_j \circ \varepsilon_{t-j} + \sum_{k=1}^{m}\sum_{\ell=1}^{n} b_{\ell k} \circ (\varepsilon_{t-\ell} X_{t-k}) + \varepsilon_t, \tag{1.2}$$

where the operators $a_i\circ$ $(i = 1, \ldots, p)$, $c_j\circ$ $(j = 1, \ldots, q)$, and $b_{\ell k}\circ$ $(\ell = 1, \ldots, n,\ k = 1, \ldots, m)$ are mutually independent and $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of independent, identically distributed nonnegative integer-valued random variables of finite mean μ and finite variance σ^2, independent of the operators.

In (1.2) it may seem more appropriate to write $\alpha_j^{(t)}\circ$ instead of $\alpha_j\circ$, to explicitly indicate that there exists a sequence of variables $\{Y_k^{(t,j)}\}_{k\in\mathbb{N}}$ for all t. However, because (1.2) more closely resembles the standard equation of the bilinear model in its present form, we prefer to use the notation $\alpha_j\circ$.

Model (1.2) has never been studied before (to the authors' knowledge) and is quite complicated. We believe that it is better to restrict our discussion to the first-order bilinear model

$$X_t = a \circ X_{t-1} + b \circ (\varepsilon_{t-1} X_{t-1}) + \varepsilon_t, \tag{1.3}$$

where the sequences involved in the operators $a\circ$ and $b\circ$ are respectively of means a and b and variances α and β. Let Y and $\tilde{Y}$ respectively denote generic variables used in $a\circ$ and $b\circ$. It should be pointed out that using the Steutel–van Harn operator in (1.2) instead of using the usual multiplication means that we are not allowed simply to invoke known results established for the classical real-valued bilinear process, as was done in [15], [26], [27], and [31, pp. 114ff.].

The structure of the paper is as follows. In Section 2 we demonstrate the existence of a strictly stationary process satisfying (1.3). In Section 3 a sufficient condition for second-order stationarity is obtained. In Section 4 we discuss the problem of the estimation of the parameters in a model where $\varepsilon_t,\ t \in \mathbb{Z}$, has a Poisson distribution with mean μ and the sequences involved in the operators $a\circ$ and $b\circ$ have Poisson distributions with respective means a and b. In Section 5 the asymptotic distribution of estimators is derived. In Section 6 we present numerical examples and applications of this model to real time series.

2. Existence of a strictly stationary bilinear process

In Theorem 2.1 we give a sufficient condition under which there is a strictly stationary process $\{X_t\}_{t\in\mathbb{Z}}$ that satisfies (1.3) and is such that ε_t is independent of $X_s,\ s < t$.

Theorem 2.1. *If $a + b\mu < 1$ then there exists a unique strictly stationary process $\{X_t\}_{t\in\mathbb{Z}}$ that satisfies (1.3) and is such that ε_t is independent of $X_s,\ s < t$.*

The proof of this theorem is based on several results that we shall prove first. For each t, let us introduce a sequence of random variables, $\{X_t^{(n)}\}_{n\in\mathbb{N}}$, that will be used in the proof of Theorem 2.1 to generate a strictly stationary solution to (1.3). Let

$$X_t^{(n)} = \begin{cases} 0, & n < 0, \\ \varepsilon_t, & n = 0, \\ a^{(t)} \circ X_{t-1}^{(n-1)} + b^{(t)} \circ (\varepsilon_{t-1} X_{t-1}^{(n-1)}) + \varepsilon_t, & n > 0. \end{cases} \tag{2.1}$$

The notation $a^{(t)}\circ$ and $b^{(t)}\circ$ indicates that the counting sequences $\{Y_k^{(t)}\}_{k\in\mathbb{N}}$ and $\{\tilde{Y}_k^{(t)}\}_{k\in\mathbb{N}}$ used in the operators $a\circ$ and $b\circ$ are fixed at time t. We will show that the sequence $\{X_t^{(n)}\}_{n\in\mathbb{Z}}$ has an almost-sure limit, denoted hereafter by X_t, for all t. We will prove that the limit process $\{X_t\}_{t\in\mathbb{Z}}$ satisfies the conditions of Theorem 2.1. To simplify the proof of the main result we demonstrate the following lemmas, which concern the sequence defined by (2.1).

Lemma 2.1. *The sequence $\{X_t^{(n)}\}_{n\in\mathbb{Z}}$ is nondecreasing for all $t \in \mathbb{Z}$.*

Proof. We prove this result by induction. For $n = 0$, we have

$$X_t^{(1)} = a^{(t)} \circ X_{t-1}^{(0)} + b^{(t)} \circ (\varepsilon_{t-1} X_{t-1}^{(0)}) + \varepsilon_t \geq \varepsilon_t = X_t^{(0)}.$$

Now suppose that $X_t^{(k)} \geq X_t^{(k-1)}$ for all t and for all $k \leq n-1$. Since ε_{t-k} is a nonnegative integer-valued random variable, using the induction hypothesis yields

$$X_{t-1}^{(n-1)} \geq X_{t-1}^{(n-2)} \quad \text{and} \quad \varepsilon_{t-1} X_{t-1}^{(n-1)} \geq \varepsilon_{t-1} X_{t-1}^{(n-2)},$$

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

and by definition of $a^{(t)}\circ$ and $b^{(t)}\circ$ we obtain

$$a^{(t)}\circ X_{t-1}^{(n-1)} \geq a^{(t)}\circ X_{t-1}^{(n-2)} \quad \text{and} \quad b^{(t)}\circ(\varepsilon_{t-1}X_{t-1}^{(n-1)}) \geq b^{(t)}\circ(\varepsilon_{t-1}X_{t-1}^{(n-2)}).$$

Consequently we can write $X_t^{(n)} \geq X_t^{(n-1)}$, and the proof follows by induction on n.

Lemma 2.2. *The process $\{X_t^{(n)}\}_{t\in\mathbb{Z}}$ is strictly stationary for all $n\in\mathbb{N}$.*

Proof. According to [5, p. 12], to show that the process $\{X_t^{(n)}\}_{t\in\mathbb{Z}}$ is strictly stationary it suffices to show that the two vectors $(X_1^{(n)},\dots,X_\ell^{(n)})'$ and $(X_{1+h}^{(n)},\dots,X_{\ell+h}^{(n)})'$ are identically distributed. It is clear that

$$\begin{pmatrix} X_1^{(0)} \\ \vdots \\ X_\ell^{(0)} \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_\ell \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} X_{1+h}^{(0)} \\ \vdots \\ X_{\ell+h}^{(0)} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1+h} \\ \vdots \\ \varepsilon_{\ell+h} \end{pmatrix}$$

are identically distributed, since $(\varepsilon_1^{(n+1)},\dots,\varepsilon_\ell^{(n+1)})'$ and $(\varepsilon_{1+h}^{(n+1)},\dots,\varepsilon_{\ell+h}^{(n+1)})'$ are identically distributed. Hence, the process $\{X_t^{(0)}\}_{t\in\mathbb{Z}}$ is strictly stationary. Now suppose that the process $\{X_t^{(r)}\}_{t\in\mathbb{Z}}$ is strictly stationary for all r such that $1\leq r\leq n$. We then have

$$\begin{aligned}\begin{pmatrix} X_1^{(n+1)} \\ \vdots \\ X_\ell^{(n+1)} \end{pmatrix} &= \begin{pmatrix} a^{(1)}\circ & \cdots & 0\circ \\ \vdots & \ddots & \vdots \\ 0\circ & \cdots & a^{(\ell)}\circ \end{pmatrix} \begin{pmatrix} X_0^{(n)} \\ \vdots \\ X_{\ell-1}^{(n)} \end{pmatrix} \\ &\quad + \begin{pmatrix} b^{(1)}\circ & \cdots & 0\circ \\ \vdots & \ddots & \vdots \\ 0\circ & \cdots & b^{(\ell)}\circ \end{pmatrix} \begin{pmatrix} X_0^{(n)}\varepsilon_0 \\ \vdots \\ X_{\ell-1}^{(n)}\varepsilon_{\ell-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_\ell \end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\begin{pmatrix} X_{1+h}^{(n+1)} \\ \vdots \\ X_{\ell+h}^{(n+1)} \end{pmatrix} &= \begin{pmatrix} a^{(1+h)}\circ & \cdots & 0\circ \\ \vdots & \ddots & \vdots \\ 0\circ & \cdots & a^{(\ell+h)}\circ \end{pmatrix} \begin{pmatrix} X_h^{(n)} \\ \vdots \\ X_{\ell+h-1}^{(n)} \end{pmatrix} \\ &\quad + \begin{pmatrix} b^{(1+h)}\circ & \cdots & 0\circ \\ \vdots & \ddots & \vdots \\ 0\circ & \cdots & b^{(\ell+h)}\circ \end{pmatrix} \begin{pmatrix} X_h^{(n)}\varepsilon_h \\ \vdots \\ X_{\ell+h-1}^{(n)}\varepsilon_{\ell+h-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1+h} \\ \vdots \\ \varepsilon_{\ell+h} \end{pmatrix}.\end{aligned}$$

By the induction hypothesis and the property of the random vectors involved in the right-hand sides of the two preceeding equalities, the vectors $(X_1^{(n+1)},\dots,X_\ell^{(n+1)})'$ and $(X_{1+h}^{(n+1)},\dots,X_{\ell+h}^{(n+1)})'$ are identically distributed.

Lemma 2.3. *The vectors $(X_t^{(n)},X_t^{(n-1)})'$ and $(X_{t+h}^{(n)},X_{t+h}^{(n-1)})'$ are identically distributed for all $n,h\in\mathbb{N}$.*

Proof. The proof is similar to the previous one, and is thus omitted.

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

Let $k_n = \mathrm{E}[X_t^{(n)} - X_t^{(n-1)}]$. By Lemma 2.3 and the structure of the process $\{X_t{}^{(n)}\}_{t\in\mathbb{Z}}$, we conclude that k_n is independent of t.

Lemma 2.4. *The sequence* $\{k_n\}_{n\in\mathbb{Z}}$ *is a geometric sequence with ratio* $a + b\mu$.

Proof. Because the sequence $\{X_t^{(n)}\}_{n\in\mathbb{Z}}$ is nondecreasing, we have the following equality in distribution:

$$X_t^{(n)} - X_t^{(n-1)} \overset{\mathrm{D}}{=} a^{(t)} \circ (X_{t-1}^{(n-1)} - X_{t-1}^{(n-2)}) + b^{(t)} \circ (\varepsilon_{t-1}(X_{t-1}^{(n-1)} - X_{t-1}^{(n-2)})).$$

By taking expectations on both sides of this equality and using the properties of the Steutel–van Harn operator, we have

$$k_n = \mathrm{E}[X_t^{(n)} - X_t^{(n-1)}] = a\,\mathrm{E}[X_t^{(n-1)} - X_t^{(n-2)}] + b\,\mathrm{E}[\varepsilon_{t-1}(X_{t-1}^{(n-1)} - X_{t-1}^{(n-2)})].$$

From (2.1) we observe that, for all $j = 1, \dots, n-1$, $X_{t-j}^{(n-j)} - X_{t-j}^{(n-j-1)}$ depends only on $\varepsilon_{t-j-1}, \dots, \varepsilon_{t-n}$ and the sequence involved in the operator. Hence,

$$\mathrm{E}[\varepsilon_{t-1}(X_{t-1}^{(n-1)} - X_{t-1}^{(n-2)})] = \mu\,\mathrm{E}[X_{t-1}^{(n-1)} - X_{t-1}^{(n-2)}]$$

and, so,

$$k_n = (a + b\mu)k_{n-1} = (a + b\mu)^{n-1}k_1,$$

where $k_1 = \mathrm{E}[X_{t-n+1}^{(1)} - X_{t-n+1}^{(0)}] = a\mu + b(\sigma^2 + \mu^2)$.

Now we prove that the sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ has a unique almost surely nonnegative integer-valued limit, X_t, for all t. The process $\{X_t\}_{t\in\mathbb{Z}}$ satisfies the conditions of Theorem 2.1.

Almost sure convergence of $\{X_t^{(n)}\}_{n\in\mathbb{N}}$. Let $(\Omega, \mathcal{F}, \mathrm{P})$ be the common probability space on which the relevant random variables are defined. Since the sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ is a nondecreasing sequence of nonnegative integers, we have

$$\lim_{n\to\infty} X_t^{(n)}(\omega) = X_t(\omega) \quad \text{for all } \omega \in \Omega.$$

It remains to show that X_t is almost surely finite. To do so, it suffices to show that the set $A_\infty = \{\omega \colon X_t(\omega) = \infty\}$ is such that $\mathrm{P}\{A_\infty\} = 0$. We observe that

$$A_\infty = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \limsup_{n\to\infty} A_n,$$

where $A_n = \{\omega \colon X_t^{(n)}(\omega) - X_t^{(n-1)}(\omega) > 0\}$. On the one hand, we have

$$k_n = \mathrm{E}[X_t^{(n)} - X_t^{(n-1)}] \geq \sum_{k=1}^{\infty} \mathrm{P}\{\omega \colon X_t^{(n)}(\omega) - X_t^{(n-1)}(\omega) = k\} = \mathrm{P}\{A_n\}.$$

On the other hand, in Lemma 2.4 we showed that $k_n = (a + b\mu)^{n-1}k_1$. Consequently, if $a + b\mu < 1$ then the series $\sum_{n\geq 1} k_n$ converges and, hence, the series $\sum_{n\geq 1} \mathrm{P}(A_n)$ also converges. Applying the Borel–Cantelli lemma yields $\mathrm{P}\{A_\infty\} = 0$, from which we conclude that X_t is almost surely finite and that the process $\{X_t\}_{t\in\mathbb{Z}}$ satisfies the conditions of Theorem 2.1.

Strict stationarity. According to Lemma 2.3, the process $\{X_t^{(n)}\}_{t\in\mathbb{Z}}$ is strictly stationary. Because X_t is the almost-sure limit of the sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ for all t, it is obvious that the process $\{X_t\}_{t\in\mathbb{Z}}$ is also strictly stationary.

Nonnegative integer-valuedness. Since X_t is the almost-sure limit of the nondecreasing integer-valued sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ for all t, we can find an $N_t > 0$ such that $X_t^{(n)} - X_t^{(m)} = 0$ for all m and n with $m \geq n \geq N_t$, and clearly $X_t^{(n)} = X_t$ for all $n \geq N_t$. Thus, X_t is a nonnegative integer-valued random variable.

Independence. Let $\Upsilon(t)$ denote all of the sequences involved in the operators, and let $\mathcal{F}(\cdot)$ denote the smallest σ-algebra that makes measurable the random variables it takes as arguments. With this notation, and from the structure of the process $\{X_t\}_{t\in\mathbb{Z}}$, for all $s < t$ we have

$$\begin{aligned}\mathcal{F}(X_s, X_{s-1}, \dots) &\subset \mathcal{F}(X_{t-1}, X_{t-2}, \dots)\\ &\subset \mathcal{F}(\varepsilon_{t-1}, \Upsilon(t-1), \varepsilon_{t-2}, \Upsilon(t-2), \dots),\end{aligned} \tag{2.2}$$

from which we deduce that ε_t is independent of $X_s,\ s < t$.

Uniqueness. Let $\{Z_t\}_{t\in\mathbb{Z}}$ be another process satisfying (1.3) that is strictly stationary and such that ε_t is independent of $Z_s,\ s < t$. We will demonstrate that $X_t = Z_t$. It suffices to show that the set $B_\infty = \{\omega\colon |X_t(\omega) - Z_t(\omega)| > 0\}$ is of probability 0. Note that $B_\infty = \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} B_m = \limsup_{n\to\infty} B_n$, where $B_n = \{\omega\colon |X_t^{(n)}(\omega) - Z_t(\omega)| > 0\}$. The following notation will be used:

$$\begin{aligned} W_t^{(0)} &= \varepsilon_t, \qquad & W_t^{(n)} &= |X_t^{(n)}(\omega) - Z_t(\omega)|,\\ L_t^{(0)} &= 0, \qquad & L_t^{(n)} &= \min(X_t^{(n)}, Z_t).\end{aligned}$$

On the one hand, we have

$$\begin{aligned} W_t^{(n)} &= |a^{(t)} \circ X_{t-1}^{(n-1)} - a^{(t)} \circ Z_{t-1} + b^{(t)} \circ (\varepsilon_{t-1} X_{t-1}^{(n-1)}) - b^{(t)} \circ (\varepsilon_{t-1} Z_{t-1})|\\ &= \sum_{i=1}^{W_{t-1}^{(n-1)}} Y_{i+L_t^{n-1}}^{(t)} + \sum_{i=1}^{\varepsilon_{t-1} W_{t-1}^{(n-1)}} \tilde{Y}_{i+\varepsilon_{t-1} L_t^{n-1}}^{(t)}\\ &\overset{\mathrm{D}}{=} a^{(t)} \circ W_{t-1}^{(n-1)} + b^{(t)} \circ (\varepsilon_{t-1} W_{t-1}^{(n-1)}),\end{aligned}$$

where, recall, $\{Y_n{}^{(t)}\}_{n\in\mathbb{N}}$ and $\{\tilde{Y}_n^{(t)}\}_{n\in\mathbb{N}}$ respectively denote the sequences involved in the operators $a^{(t)}\circ$ and $b^{(t)}\circ$. From the structure of $W_t^{(n)}$, we observe that ε_{t-1} is independent of $W_{t-1}^{(n-1)}$. By using the expected value properties of the Steutel–van Harn operator we thus find that

$$\mathrm{E}[W_t^{(n)}] = (a + b\mu)^n\, \mathrm{E}[W_{t-n}^{(0)}],$$

where $\mathrm{E}[W_{t-n}^{(0)}] = \mathrm{E}[\varepsilon_t] = \mu$. On the other hand, we have

$$\mathrm{E}[W_t^{(n)}] \geq \sum_{k=1}^{\infty} \mathrm{P}\{\omega\colon |X_t^{(n)}(\omega) - Z_t(\omega)| = k\} = \mathrm{P}\{B_n\}.$$

Consequently,

$$\sum_{n\geq 1} \mathrm{P}\{B_n\} \leq \mu \sum_{n\geq 1} (a + b\mu)^n < \infty,$$

since $a + b\mu < 1$. This means that X_t is almost surely unique.

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

3. Stationarity condition

In the previous section, we proved that the process $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationary. To conclude that this process is second-order stationary, it suffices to show that the first two moments of X_t exist.

Remark 3.1. Since the processes $\{X_t\}_{t\in\mathbb{Z}}$ and $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are strictly stationary, the processes $\{X_t\varepsilon_t\}_{t\in\mathbb{Z}}$, $\{X_t^2\varepsilon_t\}_{t\in\mathbb{Z}}$, and $\{X_t\varepsilon_t\}_{t\in\mathbb{Z}}^2$ are also strictly stationary.

Proposition 3.1. *Let the process* $\{X_t\}_{t\in\mathbb{Z}}$ *satisfy the conditions of Theorem 2.1. Then* $\mathrm{E}[X_t]$ *exists.*

Proof. From (1.3) and by using properties of the Steutel–van Harn operator, we have

$$\mathrm{E}[X_t] = a\,\mathrm{E}[X_{t-1}] + b\,\mathrm{E}[\varepsilon_{t-1}X_{t-1}] + \mu. \tag{3.1}$$

Observe that

$$\varepsilon_t X_t = (a \circ X_{t-1})\varepsilon_t + (b \circ (\varepsilon_{t-1}X_{t-1}))\varepsilon_t + \varepsilon_t^2.$$

Using the independence property of ε_t, we also have

$$\mathrm{E}[X_t\varepsilon_t] = a\mu\,\mathrm{E}[X_{t-1}] + b\mu\,\mathrm{E}[X_{t-1}\varepsilon_{t-1}] + \sigma^2 + \mu^2. \tag{3.2}$$

From Remark 3.1, we conclude that

$$\mathrm{E}[X_t\varepsilon_t] = \mathrm{E}[X_{t-1}\varepsilon_{t-1}].$$

Consequently,

$$\mathrm{E}[X_{t-1}\varepsilon_{t-1}] = \frac{a\mu\,\mathrm{E}[X_{t-1}] + \sigma^2 + \mu^2}{1 - b\mu}.$$

Finally, using the fact that the process $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationarity, we obtain

$$\mathrm{E}[X_t] = \frac{b\sigma^2 + \mu}{1 - (a + b\mu)}.$$

This expectation exists because $a + b\mu < 1$ by the existence condition.

From now on, we will omit the superscript in the operator and we will write $a\circ$ and $b\circ$ instead of $a^{(t)}\circ$ and $b^{(t)}\circ$.

Proposition 3.2. *Let the process* $\{X_t\}_{t\in\mathbb{Z}}$ *satisfy the conditions of Theorem 2.1, and suppose that* ε_t *has a finite fourth moment. If* $(a + b\mu)^2 + b\sigma^2$ *then* $\mathrm{E}[{X_t}^2]$ *exists.*

Proof. Observe that

$$\mathrm{E}[X_t^2] = \mathrm{E}[B_{t-1}^2 + 2B_{t-1}\varepsilon_t + \varepsilon_t^2],$$

where $B_{t-1} = a \circ X_{t-1} + b \circ (X_{t-1}\varepsilon_{t-1})$. Using the independence property of ε_t, we have

$$\mathrm{E}[X_t^2] = \mathrm{E}[B_{t-1}^2] + 2\mu\,\mathrm{E}[B_{t-1}] + \sigma^2 + \mu^2, \tag{3.3}$$

and using the properties of the Steutel–van Harn operator yields

$$\begin{aligned}\mathrm{E}[B_{t-1}^2] &= \mathrm{E}[(a \circ X_{t-1})^2] + 2\,\mathrm{E}[(a \circ X_{t-1})(b \circ (X_{t-1}\varepsilon_{t-1}))] + \mathrm{E}[(b \circ (X_{t-1}\varepsilon_{t-1}))^2] \\ &= \alpha\,\mathrm{E}[X_{t-1}] + a^2\,\mathrm{E}[X_{t-1}^2] + 2ab\,\mathrm{E}[X_{t-1}^2\varepsilon_{t-1}] + \beta\,\mathrm{E}[X_{t-1}\varepsilon_{t-1}] \\ &\quad + b^2\,\mathrm{E}[(X_{t-1}\varepsilon_{t-1})^2].\end{aligned}$$

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

From Remark 3.1, we obtain

$$\begin{aligned}\mathrm{E}[X_{t-1}^2\varepsilon_{t-1}] &= \mathrm{E}[X_t^2\varepsilon_t]\\ &= \mathrm{E}[(B_{t-1}^2 + 2B_{t-1}\varepsilon_t + \varepsilon_t^2)\varepsilon_t]\\ &= \mu\,\mathrm{E}[B_{t-1}^2] + 2(\sigma^2+\mu^2)\,\mathrm{E}[B_{t-1}] + \mathrm{E}[\varepsilon_t^3]\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}[(X_{t-1}\varepsilon_{t-1})^2] &= \mathrm{E}[(X_t\varepsilon_t)^2]\\ &= \mathrm{E}[(B_{t-1}^2 + 2B_{t-1}\varepsilon_t + \varepsilon_t^2)\varepsilon_t^2]\\ &= (\sigma^2+\mu^2)\,\mathrm{E}[B_{t-1}^2] + 2\,\mathrm{E}[\varepsilon_t^3]\,\mathrm{E}[B_{t-1}] + \mathrm{E}[\varepsilon_t^4].\end{aligned}$$

Let

$$\begin{aligned}C_t^{(1)} &= \alpha\,\mathrm{E}[X_{t-1}] + 4ab(\sigma^2+\mu^2)\,\mathrm{E}[B_{t-1}] + 2ab\,\mathrm{E}[\varepsilon_t^3],\\ C_t^{(2)} &= \beta\,\mathrm{E}[X_{t-1}\varepsilon_{t-1}] + 2b^2\,\mathrm{E}[\varepsilon_t^3] + b^2\,\mathrm{E}[\varepsilon_t^4].\end{aligned}$$

Consequently,

$$\mathrm{E}[B_{t-1}^2] = [2ab\mu + b^2(\sigma^2+\mu^2)]\,\mathrm{E}[B_{t-1}^2] + a^2\,\mathrm{E}[X_{t-1}^2] + C_t^{(1)} + C_t^{(2)},$$

whence

$$\mathrm{E}[B_{t-1}^2] = \frac{a^2\,\mathrm{E}[X_{t-1}^2] + C_t^{(1)} + C_t^{(2)}}{1-(2ab\mu + b^2(\sigma^2+\mu^2))}.$$

By the strict stationarity property of the process $\{X_t\}_{t\in\mathbb{Z}}$, and using (3.3), we conclude that

$$\begin{aligned}\mathrm{E}[X_t^2] &= \frac{C_t^{(1)} + C_t^{(2)} + (1-(2ab\mu + b^2(\sigma^2+\mu^2)))(2\mu\,\mathrm{E}[B_{t-1}] + \sigma^2+\mu^2)}{1-(a^2+2ab\mu + b^2(\sigma^2+\mu^2))}\\ &= \frac{C_t^{(1)} + C_t^{(2)} + (1-(2ab\mu + b^2(\sigma^2+\mu^2)))(2\mu\,\mathrm{E}[B_{t-1}] + \sigma^2+\mu^2)}{1-((a+b\mu)^2 + b^2\sigma^2)}.\end{aligned}$$

If $(a+b\mu)^2 + b^2\sigma^2 < 1$ then the numerator and the denominator both become positive, and $\mathrm{E}[X_t^2]$ exists.

We can now state the following theorem.

Theorem 3.1. *If $(a+b\mu)^2 + b^2\sigma^2 < 1$ then there exists a unique second-order, strictly stationary process that satisfyies (1.3) and is such that ε_t is independent of $X_s,\ s<t$.*

3.1. Sufficient condition for the existence of $\mathrm{E}[X_t^p]$

As we will see later, to have asymptotic normality of the estimator, the existence of $\mathrm{E}[X^p]$ for $p>4$ is required. For a random variable X, let $\|X\|_p = \mathrm{E}[|X|^p]^{1/p}$. The aim of this section is to give a sufficient condition for the existence of $\mathrm{E}[X_t^p]$. In view of 1.3, let $m_p^p = \mathrm{E}[X_t^p]$. Obviously,

$$m_p \le \|a\circ X_{t-1}\|_p + \|b\circ(\varepsilon_{t-1}X_{t-1})\|_p + \|\varepsilon_t\|_p. \tag{3.4}$$

For the first term on the right-hand side of (3.4), we have

$$\|a\circ X_{t-1}\|_p^p = \mathrm{E}[(a\circ X_{t-1})^p] = \mathrm{E}\Bigg[\sum_{j=1}^{X_{t-1}} Y_{jt}\Bigg].$$

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

By using the convexity of the function $f(z) = z^p$, we find that

$$\|a \circ X_{t-1}\|_p^p \leq \mathrm{E}\bigg[X_{t-1}^{p-1}\bigg(\sum_{j=1}^{X_{t-1}} Y_{jt}^p\bigg)\bigg] = \mathrm{E}[Y^p]\,\mathrm{E}[X_{t-1}^p].$$

Hence, by strict stationarity, we obtain

$$\|a \circ X_t\|_p \leq \|X_t\|_p \|Y\|_p.$$

For the second term on the right-hand side of (3.4), by the same argument we find that

$$\|b \circ (\varepsilon_{t-1}X_{t-1})\|_p^p \leq \mathrm{E}[(\varepsilon_{t-1}X_{t-1})^p]\,\mathrm{E}[\tilde{Y}^p].$$

Let us introduce the following notation: $n_p^p = \mathrm{E}[(\varepsilon_{t-1}X_{t-1})^p]$, $a_p^p = \mathrm{E}[Y^p]$, $b_p^p = \mathrm{E}[\tilde{Y}^p]$, and $\mu_p^p = \mathrm{E}[\varepsilon^p]$. We can then write

$$m_p \leq a_p m_p + b_p n_p + \mu_p. \tag{3.5}$$

By substituting $a \circ X_{t-2} + b \circ (\varepsilon_{t-2}X_{t-2}) + \varepsilon_{t-1}$ for X_{t-1}, we obtain

$$n_p^p = \mathrm{E}[[\varepsilon_{t-1}(a \circ X_{t-2} + b \circ (\varepsilon_{t-2}X_{t-2}) + \varepsilon_{t-1})]^p],$$

from which we deduce that

$$\begin{aligned} n_p &\leq \|\varepsilon_{t-1} a \circ X_{t-2}\|_p + \|\varepsilon_{t-1}(b \circ (\varepsilon_{t-2}X_{t-2}))\|_p + \|\varepsilon_{t-1}^2\|_p \\ &\leq \|\varepsilon\|_p \|Y\|_p m_p + \|\varepsilon\|_p \|\tilde{Y}\|_p n_p + \|\varepsilon^2\|_p. \end{aligned}$$

Isolating n_p leads to

$$n_p \leq \frac{\|\varepsilon\|_p \|Y\|_p m_p + \|\varepsilon^2\|_p}{1 - \|\varepsilon\|_p \|\tilde{Y}\|_p} = \frac{\mu_p a_p m_p + \mu_{2p}^2}{1 - \mu_p b_p}.$$

Relation (3.5) becomes

$$m_p \leq \bigg(a_p + \frac{b_p \mu_p a_p}{1 - \mu_p b_p}\bigg) m_p + \frac{b_p \mu_{2p}^2}{1 - b_p \mu_p} + \mu_p \leq \frac{a_p m_p + \mu_p + b_p(\mu_{2p}^2 - \mu_p^2)}{1 - \mu_p b_p},$$

and isolating m_p gives

$$m_p \leq \frac{\mu_p + b_p(\mu_{2p}^2 - \mu_p^2)}{1 - (a_p + \mu_p b_p)}.$$

Under the hypothesis that a_p, μ_p, and b_p exist, we conclude that m_p exists if

$$a_p + \mu_p b_p < 1 \Leftrightarrow \|Y\|_p + \|\varepsilon\|_p \|\tilde{Y}\|_p < 1. \tag{3.6}$$

In the next section we assume that the random variables Y, $\tilde{Y}$, and ε are all Poisson and, thus, that all their moments exist. In that case, $\mathrm{E}[X_t^p]$ exists if $a_p + \mu_p b_p < 1$.

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

4. Parameter estimation

In the estimation procedure discussed in this section, we assume that the distribution of the random variables of the sequence $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is $\mathcal{P}(\mu)$, the Poisson distribution with parameter μ, and that the distributions of variables of the sequences involved in the operators $a\circ$ and $b\circ$ are respectively $\mathcal{P}(a)$ and $\mathcal{P}(b)$. Even though similar results might hold for many other distributions, we prefer to investigate the Poisson case because it arises very naturally in many counting processes, in the same way that the Gaussian distribution arises in the continuous case. Also, the Poisson distribution is computationally more tractable than other distributions when dealing with integer-valued processes. Therefore, we expect that the Poisson distribution for ε_t in the INBL(p, q, m, n) process plays a role similar to that of the Gaussian distribution in the classical BL(p, q, m, n) model. The estimation problem associated with the INBL$(1, 0, 1, 1)$ process is more complicated than that associated with the BL$(1, 0, 1, 1)$ process. However, we have successfully developed some higher-order moments for the simple, nonnegative integer-valued bilinear model, so we can apply the method of moments.

Assume first that the existence and stationarity conditions hold, i.e. that $(a+b\mu)^2+b^2\mu < 1$, and let $\gamma(0)$ be the variance of the process. After some tedious calculations, we obtain

$$\mathrm{E}[X_t X_{t+1}] = (a + b\mu)\,\mathrm{E}[X_t^2] + 2b\mu\mu_X + b\mu + \mu\mu_X$$

and, for $k \geq 2$,

$$\mathrm{E}[X_t X_{t+k}] = (a + b\mu)\,\mathrm{E}[X_t X_{t+k-1}] + b\mu\mu_X + \mu\mu_X,$$

where $\mu_X = \mathrm{E}[X_t]$. Consequently, we have the following expressions:

$$\gamma(1) = (a + b\mu)\gamma(0) + (a + b\mu)\mu_X^2 - \mu_X^2 + 2b\mu\mu_X + b\mu + \mu\mu_X, \tag{4.1}$$

$$\gamma(k) = (a + b\mu)\gamma(k-1) + (a + b\mu)\mu_X^2 - \mu_X^2 + b\mu\mu_X + \mu\mu_X, \qquad k \geq 2. \tag{4.2}$$

Since ε_t is $\mathcal{P}(\mu)$-distributed, by the proof of Proposition 3.1 we have

$$\mu_X = \frac{b\mu + \mu}{1 - (a + b\mu)}.$$

Therefore, by simple substitution, we obtain

$$(a + b\mu)\mu_X^2 - \mu_X^2 + 2b\mu\mu_X + b\mu + \mu\mu_X = b\mu(\mu_X + 1)$$

and

$$(a + b\mu)\mu_X^2 - \mu_X^2 + b\mu\mu_X + \mu\mu_X = 0.$$

Thus, (4.1) and (4.2) imply that

$$a + b\mu = \frac{\gamma(k)}{\gamma(k-1)}$$

and

$$b\mu = \frac{\gamma(1) - (a + b\mu)\gamma(0)}{\mu_X + 1}.$$

By defining $A = a + b\mu$ and $B = b\mu$, since $\mu_X = (b\mu + \mu)/(1 - (a + b\mu))$ we deduce that

$$\mu = \mu_X(1 - A) - B, \qquad a = A - B, \quad \text{and} \quad b = \frac{B}{\mu}.$$

Given the observations $X_1, \dots, X_n$, let

$$\bar{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t \tag{4.3}$$

and

$$\hat{\gamma}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (X_{t+k} - \bar{X})(X_t - \bar{X}). \tag{4.4}$$

Note that we use n^{-1} instead of $(n-k)^{-1}$ as the normalizing constant for our estimator of the autocovariance $\gamma(k)$. From (4.3) and (4.4) we obtain the moment estimators $\hat{\mu}$, $\hat{a}$, and $\hat{b}$ of the corresponding parameters μ, a, and b as follows:

$$\hat{\mu} = \bar{X}(1 - \hat{A}) - \hat{B}, \tag{4.5}$$

$$\hat{a} = \hat{A} - \hat{B}, \tag{4.6}$$

$$\hat{b} = \frac{\hat{B}}{\hat{\mu}}, \tag{4.7}$$

where $\hat{A} = \hat{\gamma}(2)/\hat{\gamma}(1)$ and $\hat{B} = (\hat{\gamma}(1) - \hat{A}\hat{\gamma}(0))/(\bar{X} + 1)$.

Theorem 4.1. *The moment estimators $\hat{\mu}$, $\hat{a}$, and $\hat{b}$, defined in (4.5), (4.6), and (4.7), are strongly consistent.*

Proof. To demonstrate that the moment estimators $\hat{\mu}$, $\hat{a}$, and $\hat{b}$ are strongly consistent it suffices to prove that the process $\{X_t\}_{t\in\mathbb{Z}}$ is ergodic. As we saw previously, from (2.2) we have

$$\mathcal{F}(X_t, X_{t-1}, \dots) \subset \mathcal{F}(\varepsilon_t, \Upsilon(t), \varepsilon_{t-1}, \Upsilon(t-1), \dots).$$

Hence,

$$\bigcap_{t=0}^{\infty} \mathcal{F}(X_t, X_{t-1}, \cdots) \subset \bigcap_{t=0}^{\infty} \mathcal{F}(\varepsilon_t, \Upsilon(t), \varepsilon_{t-1}, \Upsilon(t-1), \dots). \tag{4.8}$$

Because the right-hand side of (4.8) is the tail of a σ-field of independent random variables (ε_t and $\Upsilon(t)$), the probability of any event in it is 0 or 1, from which we conclude that any event in the σ-field of the left-hand side is also of probability 0 or 1. Thus, from [32], the process $\{X_t\}_{t\in\mathbb{Z}}$ is ergodic. Since the process $\{X_t\}_{t\in\mathbb{Z}}$ is stationary and ergodic, we conclude that the estimators $\bar{X}_n$ and $\hat{\gamma}(k)$ are strongly consistent. Consequently, we deduce that the moment estimators $\hat{\mu}$, $\hat{a}$, and $\hat{b}$ are strongly consistent.

5. Asymptotic distribution of the estimators

5.1. Weak dependence of the process

To obtain the asymptotic distribution of the estimators given in Section 4, we will use some weak dependence results; see [7] or [10]. Let us denote by $\mathcal{P}_0$ the common probability space on which are defined the variables Y_{it} and $\tilde{Y}_{it}\varepsilon_t$, $t \le 0$, such that

$$\mathcal{F}_0 = \sigma(Y_{it}, \tilde{Y}_{it}, \varepsilon_t : t \le 0),$$

and let $\mathcal{P}$ denote the common probability space on which are defined the variables Y_{it}, Z_{jt}, and ε_t, $t \in \mathbb{Z}$, such that $\mathcal{F} = \sigma(Y_{it}, \tilde{Y}_{it}, \varepsilon_t : t \in \mathbb{Z})$. Let $s_t = \mathrm{E}_{\mathcal{F}_0}[X_t]$, $s = \mathrm{E}[X_t]$, $u_t = \mathrm{E}_{\mathcal{F}_0}[\varepsilon_t X_t]$,

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

and $u = \mathrm{E}[\varepsilon_t X_t]$. In the notation of [7], we will calculate $\theta_t = \mathrm{E}[|s_t - s|]$. Obviously (see (3.1) and (3.2)),

$$\begin{aligned} s_t &= a s_{t-1} + b u_{t-1} + \mu, \\ u_t &= a\mu s_{t-1} + b\mu u_{t-1} + \sigma^2 + \mu^2, \end{aligned} \tag{5.1}$$

for all $t > 1$. We can write

$$\begin{aligned} \mu(s_t - \mu) &= \mu(a s_{t-1} + b u_{t-1}), \\ \mu s_t - \mu^2 &= u_t - \sigma^2 - \mu^2, \\ u_t &= \mu s_t + \sigma^2, \end{aligned} \tag{5.2}$$

and substitution of (5.2) into (5.1) gives

$$s_t = (a + b\mu) s_{t-1} + b\sigma^2 + \mu.$$

We also have

$$s = (a + b\mu) s + b\sigma^2 + \mu.$$

Now let $z_t = s_t - s = \mathrm{E}_{\mathcal{F}_0}[X_t] - \mathrm{E}[X_t]$. Straightforward computation yields

$$\begin{aligned} z_t &= (a + b\mu) z_{t-1} \\ &= (a + b\mu)^t z_0 \\ &= (a + b\mu)^t (\mathrm{E}_{\mathcal{F}_0}[X_0] - \mathrm{E}[X_0]) \\ &= (a + b\mu)^t (X_0 - \mathrm{E}[X_0]) \\ &\le 2(a + b\mu)^t\, \mathrm{E}[X_0] \\ &= 2(a + b\mu)^t \frac{\mu + b\sigma^2}{1 - (a + b\mu)}. \end{aligned}$$

Thus, $\{X_t\}$ is a θ-weakly dependent process, with θ_r given by

$$\theta_r = (a + b\mu)^r\, \mathrm{E}[|X_0 - \mathrm{E}[X_0]|] \le 2(a + b\mu)^t\, \mathrm{E}[|X_0|].$$

5.2. Basic general and asymptotic results

Definition 5.1. (*Asymptotic normality [5, p. 211], [29, p. 122].*) The sequence $\{\boldsymbol{X}_n\}$ of random k-vectors is asymptotically normal with 'mean vector' $\boldsymbol{\mu}$ and 'covariance matrix' $\boldsymbol{\Sigma}_n$ if

(i) $\boldsymbol{\Sigma}_n$ has no zero diagonal elements for all sufficiently large n, and

(ii) $\boldsymbol{\lambda}'\boldsymbol{X}$ is $\mathrm{AN}(\boldsymbol{\lambda}'\boldsymbol{\mu}_n, \boldsymbol{\lambda}'\boldsymbol{\Sigma}_n\boldsymbol{\lambda})$ for every $\boldsymbol{\lambda} \in \mathbb{R}^k$ such that $\boldsymbol{\lambda}'\boldsymbol{\Sigma}_n\boldsymbol{\lambda} > 0$ for all sufficiently large n.

Proposition 5.1. (Transformation of asymptotically normal vectors [5, p. 211], [29, p. 122].) *Suppose that $\boldsymbol{X}_n$ is $\mathrm{AN}(\boldsymbol{\mu}, c_n^2\boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a symmetric nonnegative-definite matrix and $c_n \to 0$ as $n \to \infty$. If $\boldsymbol{g}(\boldsymbol{X}) = (g_1(\boldsymbol{X}), \dots, g_m(\boldsymbol{X}))'$ is a mapping from $\mathbb{R}^k$ into $\mathbb{R}^m$ such that each $g_i(\cdot)$ is continuously differentiable in a neighborhood of $\boldsymbol{\mu}$, and if all diagonal elements of $\boldsymbol{D\Sigma D}'$ are nonzero, where $\boldsymbol{D}$ is the $m \times k$ matrix with (i, j)th entry $[(\partial g_i/\partial x_j)(\boldsymbol{\mu})]$, then $\boldsymbol{g}(\boldsymbol{X})$ is $\mathrm{AN}(\boldsymbol{g}(\boldsymbol{\mu}), c_n^2\boldsymbol{D\Sigma D}')$.*

Let us introduce some new notation. For the process $\{X_t\}$, we write $\mathrm{E}[X_t] = \mu_X$ (as before) and $\mathrm{cov}(X_t, X_{t+h}) = \gamma(h)$ for $h \in \mathbb{Z}$. We consider the following estimators:

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t \quad \text{(as before)},$$

$$\tilde{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \mu_X)(X_{t+h} - \mu_X), \qquad h = 0, 1, 2.$$

In practice, μ_X is replaced by its strongly convergent estimator $\bar{X}$.

5.3. A central limit theorem

Results of [7] can be used to obtain asymptotic distributions. As an example, we can show that if $X_1, \dots, X_n$ are observations from model (1.3), then $\sqrt{n}\,\bar{X}$ is asymptotically normal. Indeed, if in (3.6) we let $p = 2 + \delta$, the sufficient conditions $\mathrm{E}[|X_t|^{2+\delta}] < \infty$ and $\sum_i i^{2/\delta}\theta_i < \infty$ are obviously satisfied.

Let $\boldsymbol{V} = (X_t, X_t^2, X_t X_{t-1}, X_t X_{t-2})'$ and let $\boldsymbol{\lambda} = (a, b, c, d)' \in \mathbb{R}^4$. We would like to give a central limit theorem for $\boldsymbol{\lambda}'\boldsymbol{V}$. Bardet *et al.* in [4] proved a heredity lemma giving conditions ensuring that a function $h\colon \mathbb{R}^k \to \mathbb{R}$ applied to a weakly dependent time series produces another weakly dependent time series. Their proof can easily be adapted, giving Lemma 5.1, in which we use the following norm:

$$\|(u_1, \dots, u_k)\|_\diamond = |u_1| + \cdots + |u_k|.$$

Lemma 5.1. *Assume that $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$ is a k-vectorial stationary time series such that there exists a $p > 2$ satisfying $\|\boldsymbol{X}_0\|_p < \infty$. Let $\{Y_t\}_{t\in\mathbb{Z}}$ be the stationary time series defined by $Y_t = h(\boldsymbol{X}_t)$, $t \in \mathbb{Z}$, where $h\colon \mathbb{R}^k \to \mathbb{R}$ is such that $|h(\boldsymbol{x})| \le c\|\boldsymbol{x}\|_\diamond^2$ and*

$$|h(\boldsymbol{x}) - h(\boldsymbol{x}')| \le c\|\boldsymbol{x} - \boldsymbol{x}'\|_\diamond(\|\boldsymbol{x}\|_\diamond + \|\boldsymbol{x}'\|_\diamond + 1)$$

for $\boldsymbol{x}$, $\boldsymbol{x}' \in \mathbb{R}^k$, and $c > 0$. If $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$ is a θ-weakly dependent time series then $\{Y_t\}_{t\in\mathbb{Z}}$ is a $\tilde{\theta}$-weakly dependent time series such that, for all $r \in \mathbb{N}$, $\tilde{\theta}_r = \mathrm{const.}\,\theta_r^{(p-2)/(p-1)}$, where the constant is greater than 0.

Let $\boldsymbol{U} = (X_t, X_{t-1}, X_{t-2})$. Let us define $h\colon \mathbb{R}^3 \to \mathbb{R}$ by $h(x, y, z) = ax + bx^2 + cxy + dxz$, for constants a, b, c, and d in $\mathbb{R}$. Obviously, $|h(\boldsymbol{x})| \le \mathrm{const.}\|\boldsymbol{x}\|_\diamond^2$. Hence,

$$\begin{aligned} |h(x, y, z) - h(x', y', z')| &\le |a|\,|x - x'| + |b|\,|x^2 - x'^2| + |c|\,|xy - x'y'| + |d|\,|xz - x'z'| \\ &\le C\|(x - x', y - y', z - z')\|_\diamond(\|(x, y, z)\|_\diamond + \|(x', z', z')\|_\diamond + 1), \end{aligned}$$

with $C = \max(|a|, |b|, |c|, |d|)$. We deduce that $\{Y_t\} = \{h(X_t, X_{t-1}, X_{t-2})\}$ is a $\tilde{\theta}$-weakly dependent time series, with $\tilde{\theta}_t \le \mathrm{const.}\,\theta_t^{(p-2)/(p-1)}$. Let $\boldsymbol{U}_n = (u_1, u_2, u_3, u_4)'$ be the four-vector $(\bar{X}, \tilde{\gamma}(0), \tilde{\gamma}(1), \tilde{\gamma}(2))'$. The vector $\sqrt{n}\boldsymbol{U}_n$ has a normal distribution. Parameters of this distribution can be obtained from classical results.

We know that $\bar{X}$ is unbiased for μ_X (see [5, pp. 218–219]). Because $\gamma(k)$ decreases at a geometric rate as $k \to \infty$, it is clear that $\sigma_{11} = \sum_{h=-\infty}^{\infty}\gamma(r) < \infty$: we conclude that the asymptotic variance of $\bar{X}$ is $n^{-1}\sigma_{11}$.

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

With the notation used in Wei [33, p. 95], let the kth-order cumulant of X_t be denoted by $C_k(i_1, \dots, i_{k-1})$. We have

$$\begin{aligned} C_3(i, j) &= \mathrm{E}[(X_t - \mu_X)(X_{t+i} - \mu_X)(X_{t+j} - \mu_X)], \\ C_4(i, j, k) &= \mathrm{E}[(X_t - \mu_X)(X_{t+i} - \mu_X)(X_{t+j} - \mu_X)(X_{t+k} - \mu_X)] \\ &\quad - C_2(i)C_2(j-k) - C_2(j)C_2(k-i) - C_2(k)C_2(i-j). \end{aligned}$$

Anderson in [3, Chapter 8], assuming that

$$\sum_{r=-\infty}^{\infty} |C_3(r-h, r)| < \infty \quad \text{and} \quad \sum_{r=-\infty}^{\infty} |C_4(h, -r, g-r)| < \infty,$$

provided results that are useful in computing the parameters of the asymptotic joint distribution of our estimators. These conditions are satisfied by the cumulants of the process considered here. In Appendix A we prove that these series are finite.

It is well known that $\tilde{\gamma}(h)$ is asymptotically unbiased for $\gamma(h)$, $h = 0, 1, 2$. For $i = 2, 3, 4$ and $j = 2, 3, 4$, let us denote by σ_{ij} the following expression:

$$\sigma_{ij} = \sum_{r=-\infty}^{\infty} [\gamma(r)\gamma(r+i-j) + \gamma(r-j+2)\gamma(r+i-2) + C_4(i-2, -r, j-2-r)].$$

Clearly we have $\sum_{r=-\infty}^{\infty} \gamma(r)^2 < \infty$, and, for $i = 2, 3, 4$ and $j = 2, 3, 4$, $n^{-1}\sigma_{ij}$ is the asymptotic covariance between $\tilde{\gamma}(i-2)$ and $\tilde{\gamma}(j-2)$. The covariance between $\bar{X}$ and $\tilde{\gamma}(h)$, $h = 0, 1, 2$, is also required. We only need to compute $\mathrm{E}[(\bar{X} - \mu_X)\tilde{\gamma}(h)]$ to obtain the asymptotic covariances. Simple computations yield:

$$\operatorname{cov}(\bar{X}, \tilde{\gamma}(h)) = \begin{cases} \dfrac{1}{n}\sigma_{12}^{(n)} := \dfrac{1}{n} \displaystyle\sum_{r=-n+1}^{n-1} \left(1 - \frac{|r|}{n}\right) C_3(r, r) & \text{if } h = 0, \\ \dfrac{1}{n}\sigma_{13}^{(n)} := \dfrac{1}{n} \displaystyle\sum_{r=-n+1}^{-1} \left(1 - \frac{|r|}{n}\right) C_3(r, r+1) + \sum_{r=1}^{n-1} \left(1 - \frac{|r|}{n}\right) C_3(r-1, r) & \text{if } h = 1, \\ \dfrac{1}{n}\sigma_{14}^{(n)} := \dfrac{1}{n} \displaystyle\sum_{r=-n+1}^{-2} \left(1 - \frac{|r|}{n}\right) C_3(r, r+2) + \left(1 - \frac{2}{n}\right) C_3(-1, 1) & \\ \qquad\qquad + \displaystyle\sum_{r=2}^{n-1} \left(1 - \frac{|r|}{n}\right) C_3(r-2, r) & \text{if } h = 2. \end{cases}$$

We conclude that

$$\lim_{n\to\infty} \sigma_{1j}^{(n)} = \sum_{r=-\infty}^{\infty} C_3(r-j+2, r), \qquad j = 2, 3, 4.$$

The vector $\boldsymbol{U}_n$ has an asymptotic normal distribution with mean $\boldsymbol{\mu}_{\boldsymbol{U}} = (\mu_X, \gamma(0), \gamma(1), \gamma(2))$ and covariance matrix $n^{-1}\boldsymbol{\Sigma}_{\boldsymbol{U}}$, where $\boldsymbol{\Sigma}_{\boldsymbol{U}}$ is the matrix whose (i, j)th entry is σ_{ij}. Using the definition of the estimators, in order to have

$$\boldsymbol{g}(\mu_X, \gamma(0), \gamma(1), \gamma(2)) = (\mu, a, b)$$

we let

$$g_1(u_1, u_2, u_3, u_4) = \frac{(u_3 - u_4)(u_1 + 1)u_1 + u_2u_4 - u_3^2}{u_3(u_1 + 1)},$$
$$g_2(u_1, u_2, u_3, u_4) = \frac{u_4(u_1 + u_2 + 1) - u_3^2}{u_3(u_1 + 1)},$$
$$g_3(u_1, u_2, u_3, u_4) = \frac{u_3^2 - u_2u_4}{(u_3 - u_4)(u_1 + 1)u_1 + u_2u_4 - u_3^2}.$$

The matrix $\boldsymbol{D}'$ of Proposition 5.1 is

$$\begin{bmatrix} \dfrac{(u_3 - u_4)(u_1 + 1)^2 - u_2u_4 + u_3^2}{D_1(u_1 + 1)} & \dfrac{u_3^2 - u_2u_4}{D_1(u_1 + 1)} & \dfrac{(u_3^2 - u_2u_4)(2u_1 + 1)(u_4 - u_3)}{D_2^2} \\ \dfrac{u_4}{D_1} & \dfrac{u_4}{D_1} & \dfrac{u_1u_4(u_1 + 1)(u_3 - u_4)}{D_2^2} \\ \dfrac{u_4(u_1(u_1 + 1) - u_2) - u_3^2}{u_3D_1} & -\dfrac{u_3^2 + u_4(u_1 + u_2 + 1)}{D_1u_3} & \dfrac{u_1(u_1 + 1)(u_3^2 - (2u_3 - u_2)u_4)}{D_2^2} \\ -\dfrac{u_1(u_1 + 1) - u_2}{D_1} & \dfrac{u_1 + u_2 + 1}{D_1} & \dfrac{u_1u_3(u_1 + 1)(u_3 - u_2)}{D_2^2} \end{bmatrix},$$

where $D_1 = u_3(u_1 + 1)$, $D_2 = (u_3 - u_4)(u_1 + 1)u_1 + u_2u_4 - u_3^2$, $u_1 = \mu_x$, $u_2 = \gamma(0)$, $u_3 = \gamma(1)$, and $v_4 = \gamma(2)$.

6. Applications

In this section we give two examples to illustrate the fact that the model described in this article can be used to represent series encountered in epidemiology. Two real series of length 143 are considered: the first (series 1) is the weekly number of cases of *E. coli* O157:H7 infections and the second (series 2) the weekly number of meningitis cases. Both series start in January 1990 (see Figure 1). The data set comes from the Infectious Disease Services of the Public Health Department in Roberval, Canada.

There is an important correlation at lag 1 for both series. Thus, assuming that they were generated according to a model satisfying (1.3), the parameters are estimated as suggested in Section 4. Results are given in Table 1.

In the second case, we notice that the value of $\hat{b}$ seems to be small. It would be useful to determine if this coefficient is significant or not. This may be seen as questioning the usefulness of the bilinear component in the model. By using the asymptotic results given in Section 5, we could compute a confidence interval. However, many computations are involved in producing such a confidence interval, and in practice we prefer a simpler approach. In fact, the standard error of the parameter can also be estimated by bootstrap (see [12] for details). The latter approach is much more appropriate, and is indeed recommended if n is not very large.

A GINAR(1) process is a submodel of the bilinear model of (1.3). In a bootstrap framework, one approach would thus be to model these two time series using a GINAR model as in Latour [20] and assess if there is any gain in adding the bilinear component. (Estimation results are presented in Table 2). Under the hypothesis that the appropriate model is a GINAR(1) process, by bootstrap we can estimate the distribution of $\hat{b}$ if we proceed to estimate all of the parameters in the full model.

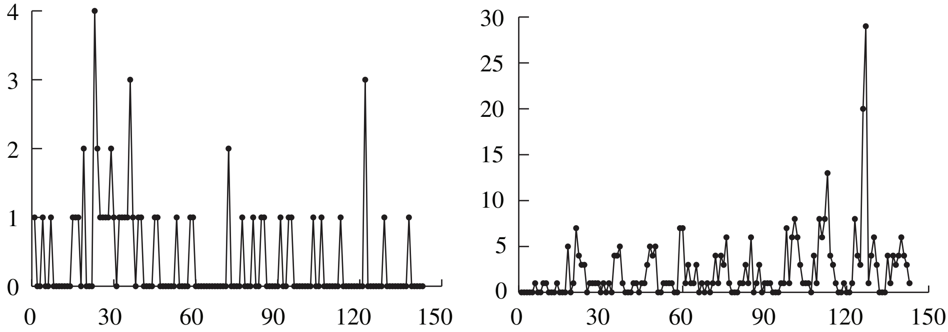

FIGURE 1: Time series data on infectious diseases: we plot X_t against t for cases of meningitis (*left*) and *E. coli* infection (*right*).

TABLE 1: Estimation of the parameters under the hypothesis that both series satisfy (1.3).

	$\hat{a}$	$\hat{b}$	$\hat{\mu}$
Series 1	0.145	0.196	1.198
Series 2	0.171	0.031	0.295

TABLE 2: Least-squares estimates of the parameters under the hypothesis that both series are generated by a GINAR(1) process.

	$\hat{a}$	$\hat{\mu}$
Series 1	0.441	1.311
Series 2	0.210	0.290

TABLE 3: Quantiles of the empirical distribution of $\hat{b}$.

	5%	10%	90%	95%
Series 1	-0.214	-0.148	0.121	0.147
Series 2	-1.045	-0.609	0.586	0.858

Under the hypothesis that the true model is a GINAR(1) process, the estimation of the parameters leads to a model that can be written as

$$X_t = \hat{a} \circ X_{t-1} + e_t \tag{6.1}$$

with $\hat{a} \approx 0.441$ for the first series and $\hat{a} \approx 0.210$ for the second series.

We simulated 4000 series of length 143 using (6.1). We computed the value of the estimators defined in Section 4 and determined the empirical distribution of $\hat{b}$. In Table 3 some quantiles are given. There we see that the percentage probability of observing a value as high as 0.196 for $\hat{b}$ is less than 5% for the first time series. In the second case, by a similar argument, the value $\hat{b} = 0.0311$ is not significant. A closer investigation may be performed to identify a better model in the GINAR family to describe the second series. We refer the reader to Latour [20] for a more complete example of fitting a GINAR model.

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

7. Conclusion

The model defined in this article is an effort to supply to practitioners another tool specifically designed for the analysis of integer-valued time series. Some other results are obviously needed in order to offer a complete tool box for the analysis of integer-valued time series that exhibit bilinear behavior.

Appendix A.

Here we use the following notation: $\widehat{X}_t$ stands for $X_t - \mu_x$ and $x^\top = x \wedge T \vee (-T)$ is the 'truncated value' of x. A convenient value for T will be given later. Explicitly, we have

$$x^\top = \begin{cases} -T & \text{if } x < -T, \\ x & \text{if } -T \leq x \leq T, \\ T & \text{if } T < x. \end{cases}$$

The difference between the truncated and original values, $x^\top - x$, is denoted by $x^\perp$. We will show that

$$\kappa_4 = \sum_{i,j,k=-\infty}^{\infty} |C_4(i, i+j, i+j+k)| < \infty, \tag{A.1}$$

a consequence of which is that we have $\sum_{r=-\infty}^{\infty} |C_4(h, -r, g-r)| < \infty$.

The summation giving κ_4 is less than or equal to the sum of three terms that we denote by $\kappa_4^{(\ell)}$, $\ell = 1, 2, 3$. The first term is a sum over indices i such that $|i| \geq |j|, |k|$, the second is a sum over indices j such that $|j| \geq |i|, |k|$ and the third is a sum over indices k such that $|k| \geq |i|, |j|$. We will give the details for the first sum, $\kappa_4^{(1)}$, only.

We have

$$C_4(i, j, k) = \operatorname{cov}(X_0, X_i X_j X_k) - C_2(i)C_2(j-k) - C_2(j)C_2(k-i) - C_2(k)C_2(i-j).$$

Note that

$$X_i X_j X_k = X_i^\top X_j^\top X_k^\top + X_i^\top X_j^\top X_k^\perp + X_i^\top X_j^\perp X_k + X_i^\perp X_j X_k.$$

Let us have a closer look at $\operatorname{cov}(X_0, X_i X_j X_k)$, writing this term as

$$\begin{aligned} \operatorname{cov}(X_0, X_i X_j X_k) &= \operatorname{cov}(X_0, X_i^\top X_j^\top X_k^\top) \\ &\quad + \operatorname{cov}(X_0, X_i^\top X_j^\top X_k^\perp + X_i^\top X_j^\perp X_k + X_i^\perp X_j X_k) \\ &= \operatorname{cov}(X_0, X_i^\top X_j^\top X_k^\top) + R_T, \end{aligned} \tag{A.2}$$

where R_T is the second term in (A.2). By weak dependence, we have

$$\operatorname{cov}\left(X_0, \frac{1}{T^3} X_i^\top X_j^\top X_k^\top\right) \leq \frac{1}{T} 3\theta_{|i|} \quad \text{and} \quad \operatorname{cov}(X_0, X_i^\top X_j^\top X_k^\top) \leq 3\,T^2 \theta_{|i|},$$

from which we deduce that

$$\operatorname{cov}(X_0, X_i X_j X_k) \leq \operatorname{cov}(X_0, X_i^\top X_j^\top X_k^\top) + R_T.$$

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

By applying Hölder's inequality with $p = 4$ and $q = \frac{4}{3}$, we obtain

$$\begin{aligned} R_T &\leq \mathrm{E}[|\widehat{X}_0(X_i^\top X_j^\top X_k^\perp + X_i^\top X_j^\perp X_k + X_i^\perp X_j X_k)|] \\ &\leq \|\widehat{X}_0\|_4 \|X_i^\top X_j^\top X_k^\perp + X_i^\top X_j^\perp X_k + X_i^\perp X_j X_k\|_{4/3}. \end{aligned}$$

Applying Minkowski's inequality to the last term yields

$$\begin{aligned} &\|X_i^\top X_j^\top X_k^\perp + X_i^\top X_j^\perp X_k + X_i^\perp X_j X_k\|_{4/3} \\ &\qquad \leq \|X_i^\top X_j^\top X_k^\perp\|_{4/3} + \|X_i^\top X_j^\perp X_k\|_{4/3} + \|X_i^\perp X_j X_k\|_{4/3}. \end{aligned}$$

By applying Lyapunov's inequality to each term, we find that

$$R_T \leq \|\widehat{X}_0\|_4 (\|X_i^\top X_j^\top X_k^\perp\|_4 + \|X_i^\top X_j^\perp X_k\|_4 + \|X_i^\perp X_j X_k\|_4).$$

Because $|X_\ell^\top| \leq |X_\ell|$, we finally obtain

$$R_T \leq 6\|X_0\|_4^3 \|X_0^\perp\|_4.$$

Next, we show that $\|X_0^\perp\|_4^4 \leq 2^4 T^{4-p}\, \mathrm{E}[|X_0|^p]$. Note that

$$\|X_0^\perp\|_4^4 = \|X_0 - X_0^\top\|_4^4 = \mathrm{E}[|X_0 - X_0^\top|^4].$$

Clearly, $|X_0 - X_0^\top| = 0$ when $|X_0| \leq T$. Otherwise, we have $|X_0 - X_0^\top| \leq |X_0| + T \leq 2|X_0|$. We conclude that $|X_0 - X_0^\top| \leq 2|X_0| 1_{\{|X_0| \geq T\}}$. Thus,

$$\|X_0 - X_0^\top\|_4^4 \leq 2^4\, \mathrm{E}[|X_0|^4 1_{\{|X_0| \geq T\}}].$$

Now apply Hölder's inequality with $p^* = p/4$ and $q^* = p/(p-4)$. The last expectation can be bounded by

$$2^4\, \mathrm{E}[|X_0|^p]^{4/p}\, \mathrm{E}[1_{\{|X_0| \geq T\}}]^{1-4/p} = 2^4\, \mathrm{E}[|X_0|^p]^{4/p}\, \mathrm{P}\{|X_0| \geq T\}^{1-4/p}.$$

Applying Markov's inequality yields

$$\|X_0 - X_0^\top\|_4^4 \leq 2^4 \|X_0\|_p^4 \bigg(\frac{1}{T^p}\, \mathrm{E}[|X_0|^p]\bigg)^{1-4/p} = 2^4 \|X_0\|_p^p\, T^{4-p}.$$

We can thus assert that

$$|\operatorname{cov}(X_0, X_i X_j X_k)| \leq 3T^2 \theta_{|i|} + 6\|X_0\|_4^3 \|X_0\|_p^{p/4} \times 2\, T^{1-p/4}.$$

If we take $T^{1+p/4} = \theta_{|i|}^{-1}$, the orders of the two terms are therefore the same and

$$|\operatorname{cov}(X_0, X_i X_j X_k)| \leq \text{const.}\, \theta_{|i|}^{(-1+p/4)/(1+p/4)} = \text{const.}\, \theta_{|i|}^{(p-4)/(p+4)}.$$

Then, in $\kappa_4^{(1)}$, j and k take on $2|i| + 1$ values, implying that

$$\kappa_4^{(1)} \leq \text{const.} \sum_{r=0}^{\infty} (2r+1)^2 \theta_r^{(p-4)/(p+4)} < \infty.$$

The other terms are easier to control. For example, if $|j| \geq |i|, |k|$ then $\operatorname{cov}(X_0 X_i, X_{i+j} X_{i+j+k})$ is approximated by $\operatorname{cov}(X_0 X_i, X_{i+j}^\top X_{i+j+k}^\top)$. Thus, as claimed above, the series in (A.1) converges.

https://doi.org/10.1239/aap/1151337085 Published online by Cambridge University Press

Acknowledgements

The authors would like to thank Pierre Boivin from the Infectious Disease Services of the Public Health Department, Roberval, Canada, who provided the data set. We would also like to thank Professor David R. Brillinger, for his encouragement in studying these integer-valued models.

References

[1] Al-Osh, M. A. and Alzaid, A. A. (1991). Binomial autoregressive moving average models. *Commun. Statist. Stoch. Models* **7,** 261–282.
[2] Alzaid, A. and Al-Osh, M. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *J. Appl. Prob.* **27,** 314–324.
[3] Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. John Wiley, New York.
[4] Bardet, J.-M., Doukhan, P. and León, J. R. (2005). Uniform limit theorems for the periodogram of a weakly dependent time series and their applications to Whittle's estimate. Submitted.
[5] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods,* 2nd edn. Springer, Berlin.
[6] Davis, R. A., Dunsmuir, W. T. M. and Streett, S. B. (2003). Observation-driven models for Poisson counts. *Biometrika* **90,** 777–790.
[7] Dedecker, J. and Doukhan, P. (2003). A new covariance inequality and applications. *Stoch. Process. Appl.* **106,** 63–80.
[8] Dewald, L. S., Lewis, P. A. W. and McKenzie, E. (1989). A bivariate first-order autoregressive time series model in exponential variables (BEAR (1)). *Manag. Sci.* **35,** 1236–1246.
[9] Dion, J.-P., Gauthier, G. and Latour, A. (1995). Branching processes with immigration and integer-valued times series. *Serdica* **21,** 123–136.
[10] Doukhan, P. and Louhichi, S. (2001). Functional estimation of a density under a new weak dependence condition. *Scand. J. Statist.* **28,** 325–341.
[11] Du, J. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.* **12,** 129–142.
[12] Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.
[13] Ferland, R., Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process. To appear in *J. Time Ser. Anal.*
[14] Gauthier, G. and Latour, A. (1994). Convergence forte des estimateurs des paramètres d'un processus GENAR(p). *Ann. Sci. Math. Québec* **18,** 49–71.
[15] Granger, C. W. J. and Andersen, A. (1978). Non-linear time series modelling. In *Applied Time Series Analysis* (Proc. 1st Symp., Tulsa, OK, 1976), Academic Press, New York, pp. 25–38.
[16] Jacobs, P. A. and Lewis, P. A. W. (1978). Discrete time series generated by mixtures. I: Correlational and runs properties. *J. R. Statist. Soc. B* **40,** 94–105.
[17] Jacobs, P. A. and Lewis, P. A. W. (1978). Discrete time series generated by mixtures. II: Asymptotic properties. *J. R. Statist. Soc. B* **40,** 222–228.
[18] Jacobs, P. A. and Lewis, P. A. W. (1983). Stationary discrete autoregressive-moving average time series generated by mixtures. *J. Time Ser. Anal.* **4,** 19–36.
[19] Latour, A. (1997). The multivariate GINAR(p) process. *Adv. Appl. Prob.* **29,** 228–248.
[20] Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *J. Time Ser. Anal.* **19,** 439–455.
[21] Lewis, P. A. W. and McKenzie, E. (1991). Minification processes and their transformations. *J. Appl. Prob.* **28,** 45–57.
[22] MacDonald, I. L. and Zucchini, W. (1997). *Hidden Markov and Other Models for Discrete-Valued Time Series*. Chapman and Hall, Boca Raton, FL.
[23] McKenzie, E. (1986). Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. *Adv. Appl. Prob.* **18,** 679–705.
[24] McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Adv. Appl. Prob.* **20,** 822–835.
[25] Park, Y. and Kim, M. (1997). Some basic and asymptotic properties in INMA(q) processes. *J. Korean Statist. Soc.* **26,** 155–170.
[26] Pham, D. T. (1985). Bilinear Markovian representation and bilinear models. *Stoch. Process. Appl.* **20,** 295–306.
[27] Pham, D. T. (1986). The mixing property of bilinear and generalised random coefficient autoregressive models. *Stoch. Process. Appl.* **23,** 291–300.
[28] Rydberg, T. H. and Shephard, N. (2000). BIN models for trade-by-trade data. Modelling the number of trades in a fixed interval of time. Tech. Rep. 0740, Econometric Society. Available at http://ideas.repec.org/p/ecm/wc2000/0740.html.

[29] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics* (Wiley Ser. Prob. Math. Stat.). John Wiley, New York.
[30] Streett, S. B. (2000). Some observation driven models for time series. Doctoral Thesis, Department of Statistics, Colorado State University.
[31] Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach*. Clarendon Press, Oxford.
[32] Wang, Z.-K. (1978). *Stochastic Processes*. Scientific Press, Beijing.
[33] Wei, W. W. S. (1990). *Time Series Analysis: Univariate and Multivariate Methods*. Addison-Wesley, Boston, MA.