DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN The Netherlands

ASYMPTOTIC BEHAVIOUR OF ESTIMATORS OF THE PARAMETERS OF NEARLY UNSTABLE INAR(1) MODELS

Márton Ispány, Gyula Pap, Martien C.A. van Zuijlen

Report No. 0120 (September 2001)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN
Toernooiveld
6525 ED Nijmegen
The Netherlands

Asymptotic Behaviour of Estimators of the Parameters of Nearly Unstable INAR(1) Models

Márton Ispány* Gyula Pap Martien C.A. van Zuijlen

October 19, 2001

Abstract

A sequence of first–order integer–valued autoregressive type (INAR(1)) processes is investigated, where the autoregressive type coefficients converge to 1. It is shown that the limiting distribution of the joint conditional least squares estimators for this coefficient and for the mean of the innovation is normal. Consequences for sequences of Galton–Watson branching processes with unobservable immigration, where the mean of the offspring distribution converges to 1 (which is the critical value), are discussed.

1 Introduction

In many practical situations one has to deal with non–negative integer–valued time series. Examples of such time series, known as counting processes, arise in several fields of medicine (see, e.g., Cardinal et.al. [5] and Franke and Seligmann [9]). To construct counting processes Al–Osh and Alzaid [1] proposed a particular class of models, the so–called INAR(1) model. Later Al–Osh and Alzaid [2], Du and Li [8] and Latour [12] generalized this model by introducing the INAR(p) and GINAR(p) models. These processes can be considered as discrete analogues of the scalar– and vector–valued AR(p) processes, because their correlation structure is similar.

The present paper deals with so–called nearly unstable INAR(1) models. It is, in fact, a sequence of INAR(1) models where the autoregressive type coefficient α_n is close to one, more precisely, $\alpha_n = 1 - \gamma_n/n$ with $\gamma_n \to \gamma$, where $\gamma \geqslant 0$. This parametrization has been suggested by Chan and Wei [6] for the usual AR(1) model. The main motivation of our investigation comes from econometrics, where the so–called 'unit root problem' plays an important role (see, e.g., the monograph of Tanaka [15]). We considered in [10] the conditional least squares estimate (CLSE)

*Institute of Mathematics and Informatics, University of Debrecen, Pf. 12, H–4010 Debrecen, Hungary

This research has been supported by the Hungarian Scientific Research Fund under Grant No. OTKA–T032361/2000.

for α_n assuming that the mean μ_ε of the innovation is known. In this paper we do not suppose that μ_ε is known, and we show asymptotic normality of the joint CLSE of α_n and μ_ε.

To define the INAR(1) model let us recall the definition of the $\alpha\circ$ operator which is due to Steutel and van Harn [14].

Definition 1.1 *Let X be a non–negative integer–valued random variable. Let $(Y_j)_{j\in\mathbb{N}}$ be a sequence of independent and identically distributed (i.i.d.) Bernoulli random variables with mean α. We assume that the sequence $(Y_j)_{j\in\mathbb{N}}$ is independent of X. The non–negative integer–valued random variable $\alpha\circ X$ is defined by*

$$\alpha\circ X := \begin{cases} \sum\limits_{j=1}^{X} Y_j, & X>0,\\ 0, & X=0. \end{cases}$$

The sequence $(Y_j)_{j\in\mathbb{N}}$ is called a ***counting sequence****.*

Let $(\varepsilon_k)_{k\in\mathbb{N}}$ is an i.i.d. sequence of non–negative integer–valued random variables with mean μ_ε and variance σ_ε^2. The zero start INAR(1) time series model is defined as

$$X_k = \begin{cases} \alpha\circ X_{k-1}+\varepsilon_k, & k=1,2,\dots,\\ 0, & k=0, \end{cases}$$

where the counting sequences $(Y_j)_{j\in\mathbb{N}}$ involved in $\alpha\circ X_{k-1}$ for $k=1,2,\dots$ are mutually independent and independent of $(\varepsilon_k)_{k\in\mathbb{N}}$. We suppose that $\mu_\varepsilon>0$ (otherwise $X_k=0$ for all $k\in\mathbb{N}$).

It is easy to show (see [10]), that

$$\lim_{k\to\infty}\mathsf{E}X_k=\frac{\mu_\varepsilon}{1-\alpha},\qquad \lim_{k\to\infty}\mathsf{Var}X_k=\frac{\sigma_\varepsilon^2+\alpha\mu_\varepsilon}{1-lpha^2},\qquad \text{for all }\ \alpha\in[0,1),$$

and that $\lim_{k\to\infty}\mathsf{E}X_k=\lim_{k\to\infty}\mathsf{Var}X_k=\infty$ if $\alpha=1$. The case $\alpha\in[0,1)$ is called *stable* or *asymptotically stationary*, while the case $\alpha=1$ is called *unstable*.

Let $\mathcal{F}_k$ be the σ–algebra generated by the random variables $X_1,\dots,X_k$. Clearly $\mathsf{E}(X_k\mid\mathcal{F}_{k-1})=\alpha X_{k-1}+\mu_\varepsilon$, thus the conditional least squares estimator (CLSE) $\widehat{\alpha}$ of α based on the observations $(X_k)_{1\leqslant k\leqslant n}$ (assuming that μ_ε is known) can be obtained by minimizing the sum of squares

$$\sum_{k=1}^{n}(X_k-\alpha X_{k-1}-\mu_\varepsilon)^2 \tag{1}$$

with respect to α, and it has the form

$$\widehat{\alpha}_n=\frac{\sum_{k=1}^n X_{k-1}(X_k-\mu_\varepsilon)}{\sum_{k=1}^n (X_{k-1})^2}.$$

In the stable case under the assumption $\mathsf{E}\varepsilon_1^3<+\infty$ we have

$$n^{1/2}(\widehat{\alpha}_n-\alpha)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\sigma_{\alpha,\varepsilon}^2),\qquad \sigma_{\alpha,\varepsilon}^2=\frac{\alpha(1-\alpha)\mathsf{E}Z_0^3+\sigma_\varepsilon^2\mathsf{E}Z_0^2}{(\mathsf{E}Z_0^2)^2},$$

where $(Z_k)_{k\in\mathbb{Z}}$ is a stationary solution of the INAR(1) model

$$Z_k = \alpha \circ Z_{k-1} + \varepsilon_k, \qquad k \in \mathbb{Z},$$

see Klimko and Nelson [11].

Let us consider now a nearly unstable sequence of INAR(1) models

$$X_k^{(n)} = \begin{cases} \alpha_n \circ X_{k-1}^{(n)} + \varepsilon_k^{(n)}, & k = 1, 2, \dots, \\ 0, & k = 0, \end{cases} \qquad n = 1, 2, \dots,$$

where the autoregressive type coefficient has the form $\alpha_n = 1 - \gamma_n/n$ with $\gamma_n \to \gamma$ such that $\gamma \geqslant 0$. In [10] the authors have proved that $(\widehat{\alpha}_n)_{n\in\mathbb{N}}$ is asymptotically normal, namely,

$$n^{3/2}(\widehat{\alpha}_n - \alpha_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2_{\gamma,\varepsilon}).$$

In this case it suffices to assume $\mathsf{E}\varepsilon_1^2 < +\infty$. We draw the attention to the normalizing factor $n^{3/2}$, which is different from the stable case.

By minimizing the sum of squares (1) with respect to α_n and μ_ε, we obtain the joint conditional least squares estimator $(\widetilde{\alpha}_n, \widetilde{\mu}_{\varepsilon,n})$ of the vector $(\alpha_n, \mu_\varepsilon)$ based on the observations $(X_k^{(n)})_{1\leqslant k\leqslant n}$:

$$\widetilde{\alpha}_n = \frac{\sum_{k=1}^n X_{k-1}^{(n)}(X_k^{(n)} - \overline{X}^{(n)})}{\sum_{k=1}^n (X_{k-1}^{(n)} - \overline{X}_*^{(n)})^2}, \qquad \widetilde{\mu}_{\varepsilon,n} = \overline{X}^{(n)} - \widetilde{\alpha}_n \overline{X}_*^{(n)},$$

where

$$\overline{X}^{(n)} := \frac{1}{n}\sum_{k=1}^n X_k^{(n)}, \qquad \overline{X}_*^{(n)} := \frac{1}{n}\sum_{k=1}^n X_{k-1}^{(n)}.$$

In Section 3 we show that $(\widetilde{\alpha}_n, \widetilde{\mu}_{\varepsilon,n})_{n in \mathbb{N}}$ is asymptotically normal, namely,

$$\begin{pmatrix} n^{3/2}(\widetilde{\alpha}_n - \alpha_n) \\ n^{1/2}(\widetilde{\mu}_{\varepsilon,n} - \mu_\varepsilon) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_{\gamma,\varepsilon}), \tag{2}$$

and the covariance matrix $\Sigma_{\gamma,\varepsilon}$ will be given explicitely.

It is easy to observe that the INAR(1) process is a special case of the Galton–Watson branching process with immigration if the offspring distribution is a Bernoulli distribution (see, e.g., Franke and Seligmann [9]). We recall that a Galton–Watson process is said to be subcritical, critical or supercritical if the expectation of the offspring distribution is less than 1, equals 1 or greater than 1, respectively. The result (2) can be reformulated as follows.

Corollary 1.2 *Consider a sequence of Galton–Watson branching processes with Bernoulli offspring distribution with parameter* $\alpha_n = 1-\gamma_n/n$, $\gamma_n \to \gamma$ *where* $\gamma \geqslant 0$, *and (unobservable) immigration with expectation* $\mu_\varepsilon > 0$ *and variance* $\sigma_\varepsilon^2 < \infty$.

Then the joint conditional least squares estimator of α_n *and* μ_ε *is asymptotically normal.*

We remark that the asymtotic normality in the sub–critical case with general offspring distribution and observed immigration is proved by Venkataraman and Nanthi [16]. The rate of convergence is $n^{1/2}$ in this case. We conjeture that our result can be extended for Galton–Watson processes with a more general offspring distribution.

We note that Sriram [13] considered a nearly critical sequence of Galton–Watson branching processes with a general offspring distribution. However, the immigration was supposed to be observable. That is the reason why Sriram [13] investigated the limiting behaviour of another joint estimator for the offspring mean and for the mean of the immigration distribution.

2 Preliminaries

We shall need a simple lemma, which gives a sufficient condition for convergence to a functional of a continuous process. The proof is based on the Continuous Mapping Theorem (see Billingsley [4, Theorem 5.5]), and it can be found in Arató, Pap and Zuijlen [3].

For measurable mappings $\Phi, \Phi_n : \mathbb{D}(\mathbb{R}_+, \mathbb{R}^k) \to \mathbb{D}(\mathbb{R}_+, \mathbb{R}^\ell)$, $n = 1, 2, \ldots$ we shall write $\Phi_n \rightsquigarrow \Phi$ if $\|\Phi_n(x_n) - \Phi(x)\|_\infty \to 0$ for all $x, x_n \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$ with $\|x_n - x\|_\infty \to 0$, where $\|\cdot\|_\infty$ denotes the supremum norm.

Lemma 2.1 *Let* $\Phi, \Phi_n : \mathbb{D}(\mathbb{R}_+, \mathbb{R}^k) \to \mathbb{D}(\mathbb{R}_+, \mathbb{R}^\ell)$, $n = 1, 2, \ldots$ *be measurable mappings such that* $\Phi_n \rightsquigarrow \Phi$. *Let* Z, Z_n, $n = 1, 2, \ldots$ *be stochastic processes with values in* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$ *such that* $Z_n \xrightarrow{\mathcal{D}} Z$ *in* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$ *and almost all trajectories of* Z *are continuous. Then,* $\Phi_n(Z_n) \xrightarrow{\mathcal{D}} \Phi(Z)$ *in* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^\ell)$.

Let
$$M_k^{(n)} := X_k^{(n)} - \alpha_n X_{k-1}^{(n)} - \mu_\varepsilon.$$

Let us introduce the random step functions

$$X^{(n)}(t) := X_{[nt]}^{(n)}, \qquad M^{(n)}(t) := \sum_{k=1}^{[nt]} M_k^{(n)}, \qquad t \geqslant 0.$$

In [10] we have shown that

$$\big(\widetilde{M}^{(n)}, \widetilde{X}^{(n)}\big) := \left(\frac{M^{(n)}}{\sqrt{n}}, \frac{X^{(n)} - \mathsf{E}X^{(n)}}{\sqrt{n}}\right) \xrightarrow{\mathcal{D}} (M, X) \tag{3}$$

in the Skorokhod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$, where $(M(t))_{t \geqslant 0}$ is a time–changed Wiener process, namely, $M(t) = W(T_M(t))$ with

$$T_M(t) := \int_0^t \varrho_{\gamma,\varepsilon}(u)\, \mathrm{d}u, \qquad \varrho_{\gamma,\varepsilon}(t) := \sigma_\varepsilon^2 + \mu_\varepsilon(1 - \mathrm{e}^{-\gamma u}),$$

and $(W(t))_{t\geqslant 0}$ is a standard Wiener process, and

$$X(t) := \int_0^t \mathrm{e}^{-\gamma(t-s)}\,\mathrm{d}M(s), \qquad t \geqslant 0$$

is a continuous zero mean Gaussian martingale (which is an Ornstein–Uhlenbeck type process driven by M). The main idea was first to prove that $\widetilde{M}^{(n)} \xrightarrow{\mathcal{D}} M$ by the help of the Martingale Central Limit Theorem, and then to show that $\widetilde{X}^{(n)}$ is a measurable function of $\widetilde{M}^{(n)}$, namely, $\big(\widetilde{M}^{(n)}, \widetilde{X}^{(n)}\big) = \Phi_n\big(\widetilde{M}^{(n)}\big)$ with $\Phi_n : \mathbb{D}(\mathbb{R}_+,\mathbb{R}) \to \mathbb{D}(\mathbb{R}_+,\mathbb{R}^2)$,

$$\Phi_n(x)(t) = \left(x(t),\, x\left(\frac{[nt]}{n}\right) - \gamma_n^* \int_0^{[nt]/n} \mathrm{e}^{-\gamma_n^*([nt]/n-s)} x(s)\,\mathrm{d}s \right),$$

where $\gamma_n^* := -n\log\alpha_n \to \gamma$. Clearly $\Phi_n \rightsquigarrow \Phi$, where

$$\Phi(x)(t) = \left(x(t),\, x(t) - \gamma \int_0^t \mathrm{e}^{-\gamma(t-s)} x(s)\,\mathrm{d}s \right).$$

By Lemma 2.1, $(\widetilde{M}^{(n)}, \widetilde{X}^{(n)}) \xrightarrow{\mathcal{D}} (M, X)$, since Itô's formula yields

$$\int_0^t \mathrm{e}^{-\gamma(t-s)}\,\mathrm{d}M(s) = M(t) - \gamma \int_0^t \mathrm{e}^{-\gamma(t-s)} M(s)\,\mathrm{d}s,$$

hence $(M, X) = \Phi(X)$.

Moreover, based on (3), we proved in [10] that

$$n^{3/2}(\widehat{\alpha}_n - \alpha_n) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mu_X(t)\,\mathrm{d}M(t)}{\int_0^1 \mu_X(t)^2\,\mathrm{d}t} \stackrel{\mathcal{D}}{=} \mathcal{N}(0, \sigma^2_{\gamma,\varepsilon}),$$

where

$$\mu_X(t) := \mu_\varepsilon \int_0^t \mathrm{e}^{-\gamma u}\mathrm{d}u = \begin{cases} \frac{\mu_\varepsilon}{\gamma}(1-\mathrm{e}^{-\gamma t}), & \gamma > 0, \\ \mu_\varepsilon t, & \gamma = 0, \end{cases}$$

$$\sigma^2_{\gamma,\varepsilon} := \frac{\int_0^1 \mu_X(t)^2 \varrho_{\gamma,\varepsilon}(t)\,\mathrm{d}t}{\left(\int_0^1 \mu_X(t)^2\,\mathrm{d}t\right)^2}.$$

Introducing

$$\mu_X^{(n)}(t) := \frac{1}{n}\mathsf{E}X^{(n)}(t) = \frac{1}{n}\mathsf{E}X^{(n)}_{[nt]},$$

it is easy to show (see [10]) that $\mu_X^{(n)} \to \mu_X$ locally uniformly on $\mathbb{R}_+$, hence also in $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$.

We remark that Sriram [13] proved a limit theorem for the process $n^{-1}X^{(n)}$ for a nearly critical sequence of Galton–Watson branching processes with a general offspring distribution. However, the result of Sriram [13] is not applicable for a nearly critical sequence of branching processes with Bernoulli offspring distribution, since the variance $\alpha(1-\alpha)$ of the Bernoulli distribution tends to 0 as α tends to its critical value 1. In fact, (3) implies that in this case we have $n^{-1}X^{(n)} \xrightarrow{\mathcal{D}} \mu_X$ in the Skorokhod space $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$, but this limiting relationship is not sufficient for deriving the limiting behaviour of the sequence $(\widetilde{\alpha}_n, \widetilde{\mu}_{\varepsilon,n})$.

3 Joint Estimator

The main result of the paper is that the joint conditional least squares estimator $(\widetilde{\alpha}_n, \widetilde{\mu}_{\varepsilon,n})$ of the vector $(\alpha_n, \mu_\varepsilon)$ for a nearly unstable sequence of INAR(1) models is asymptotically normal.

Theorem 3.1 *Consider a sequence of INAR(1) models with parameters $\alpha_n = 1-\gamma_n/n$ such that $\gamma_n \to \gamma$ with $\gamma \geqslant 0$, and suppose that $\mu_\varepsilon > 0$ and $\sigma_\varepsilon^2 < \infty$. Then*

$$\begin{pmatrix} n^{3/2}(\widetilde{\alpha}_n - \alpha_n) \\ n^{1/2}(\widetilde{\mu}_{\varepsilon,n} - \mu_\varepsilon) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \dfrac{\int_0^1 \mu_X(t)\,\mathrm{d}M(t) - \overline{\mu}_{X,1} M(1)}{\overline{\mu}_{X,2} - (\overline{\mu}_{X,1})^2} \\ \dfrac{\overline{\mu}_{X,2} M(1) - \overline{\mu}_{X,1}\int_0^1 \mu_X(t)\,\mathrm{d}M(t)}{\overline{\mu}_{X,2} - (\overline{\mu}_{X,1})^2} \end{pmatrix} \stackrel{\mathcal{D}}{=} \mathcal{N}(0, \Sigma_{\gamma,\varepsilon}),$$

where $\overline{\mu}_{X,1} := \int_0^1 \mu_X(t)\,\mathrm{d}t$, $\overline{\mu}_{X,2} := \int_0^1 (\mu_X(t))^2\,\mathrm{d}t$, *and*

$$\Sigma_{\gamma,\varepsilon} = \left(\frac{\sigma_{\gamma,\varepsilon}^{(i,j)}}{\big(\overline{\mu}_{X,2} - (\overline{\mu}_{X,1})^2\big)^2} \right)_{1\leqslant i,j\leqslant 2}$$

with

$$\begin{aligned}
\sigma_{\gamma,\varepsilon}^{(1,1)} &= \int_0^1 (\mu_X(t) - \overline{\mu}_{X,1})^2 \varrho_{\gamma,\varepsilon}(t)\,\mathrm{d}t, \\
\sigma_{\gamma,\varepsilon}^{(1,2)} &= -\int_0^1 (\mu_X(t) - \overline{\mu}_{X,1})(\overline{\mu}_{X,1}\mu_X(t) - \overline{\mu}_{X,2}) \varrho_{ma,\varepsilon}(t)\,\mathrm{d}t, \\
\sigma_{\gamma,\varepsilon}^{(2,2)} &= \int_0^1 (\overline{\mu}_{X,1}\mu_X(t) - \overline{\mu}_{X,2})^2 \varrho_{\gamma,\varepsilon}(t)\,\mathrm{d}t.
\end{aligned}$$

Proof. We have

$$\widetilde{\alpha}_n = \frac{\sum_{k=1}^n X_{k-1}^{(n)} X_k^{(n)} - n^{-1}\sum_{k=1}^n X_{k-1}^{(n)} \sum_{k=1}^n X_k^{(n)}}{\sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2 - n^{-1}\big(\sum_{k=1}^n X_{k-1}^{(n)}\big)^2},$$

hence $X_k^{(n)} - \alpha_n X_{k-1}^{(n)} = M_k^{(n)} + \mu_\varepsilon$ implies

$$\widetilde{\alpha}_n - \alpha_n = \frac{\sum_{k=1}^n X_{k-1}^{(n)} M_k^{(n)} - n^{-1} \sum_{k=1}^n X_{k-1}^{(n)} \sum_{k=1}^n M_k^{(n)}}{\sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2 - n^{-1}\Big(\sum_{k=1}^n X_{k-1}^{(n)}\Big)^2} = \frac{U_n}{V_n},$$

where

$$\begin{aligned}
U_n &:= \int_0^1 X^{(n)}(t)\,\mathrm{d}M^{(n)}(t) - M^{(n)}(1)\int_0^1 X^{(n)}(t)\,\mathrm{d}t,\\
V_n &:= n\int_0^1 \big(X^{(n)}(t)\big)^2 \mathrm{d}t - n\Big(\int_0^1 X^{(n)}(t)\,\mathrm{d}t\Big)^2.
\end{aligned}$$

Applying $X^{(n)}(t) = n\mu_X^{(n)}(t) + n^{1/2}\widetilde{X}^{(n)}(t)$ and $M^{(n)}(t) = n^{1/2}\widetilde{M}^{(n)}(t)$, we obtain

$$\begin{aligned}
U_n &= U_{n,1} n^{3/2} + U_{n,2} n,\\
V_n &= V_{n,1} n^3 + V_{n,2} n^{5/2} + V_{n,3} n^2,
\end{aligned}$$

where

$$\begin{aligned}
U_{n,1} &:= \int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t) - \widetilde{M}^{(n)}(1)\int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}t,\\
U_{n,2} &:= \int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t) - \widetilde{M}^{(n)}(1)\int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}t,\\
V_{n,1} &:= \int_0^1 \big(\mu_X^{(n)}(t)\big)^2 \mathrm{d}t - \bigg(\int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}t\bigg)^2,\\
V_{n,2} &:= 2\int_0^1 \mu_X^{(n)}(t)\widetilde{X}^{(n)}(t)\,\mathrm{d}t - 2\int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}t\int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}t,\\
V_{n,3} &:= \int_0^1 \big(\widetilde{X}^{(n)}(t)\big)^2 \mathrm{d}t - \bigg(\int_0^1 \widetilde{X}^{(n)}(t)\ \ ddt\bigg)^2.
\end{aligned}$$

Next we investigate

$$\begin{aligned}
\widetilde{\mu}_{\varepsilon,n} &= \overline{X}^{(n)} - \overline{X}_*^{(n)} \frac{\sum_{k=1}^n X_{k-1}^{(n)} X_k^{(n)} - n^{-1}\sum_{k=1}^n X_{k-1}^{(n)} \sum_{k=1}^n X_k^{(n)}}{\sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2 - n^{-1}\big(\sum_{k=1}^n X_{k-1}^{(n)}\big)^2}\\
&= \frac{n^{-1}\sum_{k=1}^n X_k^{(n)} \sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2 - n^{-1}\sum_{k=1}^n X_{k-1}^{(n)} \sum_{k=1}^n X_{k-1}^{(n)} X_k^{(n)}}{\sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2 - n^{-1}\big(\sum_{k=1}^n X_{k-1}^{(n)}\big)^2}.
\end{aligned}$$

Clearly we have

$$\widetilde{\mu}_{\varepsilon,n} - \mu_\varepsilon = \frac{W_n}{V_n},$$

where

$$W_n := n^{-1}\sum_{k=1}^n \big(X_k^{(n)} - \mu_\varepsilon\big)\sum_{k=1}^n \big(X_{k-1}^{(n)}\big)^2 - n^{-1}\sum_{k=1}^n X_{k-1}^{(n)} \sum_{k=1}^n X_{k-1}^{(n)}\big(X_k^{(n)} - \mu_\varepsilon\big).$$

By $X_k^{(n)} - \mu_\varepsilon = \alpha_n X_{k-1}^{(n)} + M_k^{(n)}$, we can write W_n in the form

$$\begin{aligned}
W_n &= n^{-1} \sum_{k=1}^{n} M_k^{(n)} \sum_{k=1}^{n} \big(X_{k-1}^{(n)}\big)^2 - n^{-1} \sum_{k=1}^{n} X_{k-1}^{(n)} \sum_{k=1}^{n} X_{k-1}^{(n)} M_k^{(n)} \\
&= M^{(n)}(1) \int_0^1 \big(X^{(n)}(t)\big)^2 \,\mathrm{d}t - \int_0^1 X^{(n)}(t)\,\mathrm{d}t \int_0^1 X^{(n)}(t)\,\mathrm{d}M^{(n)}(t).
\end{aligned}$$

Applying again $X^{(n)}(t) = n\mu_X^{(n)}(t) + n^{1/2}\widetilde{X}^{(n)}(t)$ and $M^{(n)}(t) = n^{1/2}\widetilde{M}^{(n)}(t)$, we obtain that

$$W_n = W_{n,1}n^{5/2} + W_{n,2}n^2 + W_{n,3}n^{3/2},$$

where

$$\begin{aligned}
W_{n,1} &:= \widetilde{M}^{(n)}(1) \int_0^1 \big(\mu_X^{(n)}(t)\big)^2 \mathrm{d}t - \int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}t \int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t), \\
W_{n,2} &:= 2\widetilde{M}^{(n)}(1) \int_0^1 \mu_X^{(n)}(t)\widetilde{X}^{(n)}(t)\,\mathrm{d}t - \int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}t \int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t) \\
&\quad - \int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}t \int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t), \\
W_{n,3} &:= \widetilde{M}^{(n)}(1) \int_0^1 \big(\widetilde{X}^{(n)}(t)\big)^2 \mathrm{d}t - \int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}t \int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t).
\end{aligned}$$

We can notice that

$$Z_n := (U_{n,1}, U_{n,2}, V_{n,1}, V_{n,2}, V_{n,3}, W_{n,1}, W_{n,2}, W_{n,3})$$

can be expressed as a continuous function of the random vector

$$\begin{aligned}
I_n := \Bigg(&\widetilde{M}^{(n)}(1), \int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}t, \int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}t, \int_0^1 \big(\mu_X^{(n)}(t)\big)^2 \mathrm{d}t, \\
&\int_0^1 \big(\widetilde{X}^{(n)}(t)\big)^2 \mathrm{d}t, \int_0^1 \mu_X^{(n)}(t)\widetilde{X}^{(n)}(t)\,\mathrm{d}t, \int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t)\Bigg)
\end{aligned}$$

and the random variable

$$\int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t).$$

In [10] it is shown that there exist measurable functionals $\Phi, \Phi_n : \mathbb{D}(\mathbb{R}_+, \mathbb{R}) \to \mathbb{R}$, $n \in \mathbb{N}$, such that

$$\int_0^1 \mu_X^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t) = \Phi_n(\widetilde{M}^{(n)}),$$

and $\Phi_n \rightsquigarrow \Phi$ in the sense that $|\Phi_n(x_n) - \Phi(x)| \to 0$ for all $x, x_n \in \mathbb{D}(\mathbb{R}_+, \mathbb{R})$ with $\|x_n - x\|_\infty \to 0$. Hence we conclude the existence of measurable functionals $\Psi, \Psi_n : \mathbb{D}(\mathbb{R}_+, \mathbb{R}^3) \to \mathbb{R}^7$, $n \in \mathbb{N}$, such that $I_n = \Psi_n(\mu_X^{(n)}, \widetilde{M}^{(n)}, \widetilde{X}^{(n)})$, and $\Psi_n \rightsquigarrow \Psi$ in the sense that $\|\Psi_n(x_n) - \Psi(x)\| \to 0$ for all $x, x_n \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^3)$ with

$\|x_n - x\|_\infty \to 0$. Thus (3), $\mu_X^{(n)} \to \mu_X$ in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$, and an appropriate analogue of Lemma 2.1 imply $I_n \xrightarrow{\mathcal{D}} I$ with

$$I := \bigg(M(1),\, \int_0^1 \mu_X(t)\,\mathrm{d}t,\, \int_0^1 X(t)\,\mathrm{d}t,\, \int_0^1 \big(\mu_X(t)\big)^2 \mathrm{d}t,\\ \int_0^1 \big(X(t)\big)^2 \mathrm{d}t,\, \int_0^1 \mu_X(t) X(t)\,\mathrm{d}t,\, \int_0^1 \mu_X(t)\,\mathrm{d}M(t) \bigg).$$

In [10] we have shown that

$$\int_0^1 \widetilde{X}^{(n)}(t)\,\mathrm{d}\widetilde{M}^{(n)}(t) = A_n + B_n,$$

where

$$A_n := \frac{1}{2}\big(\widetilde{X}^{(n)}(1)\big)^2 + \frac{(1+\alpha_n)\gamma_n}{2} \int_0^1 \big(\widetilde{X}^{(n)}(t)\big)^2 \mathrm{d}t,$$

$$B_n := \frac{1}{2n} \sum_{k=1}^n \big(M_k^{(n)}\big)^2 \xrightarrow{\mathcal{D}} \frac{1}{2} T_M(1).$$

Consequently, applying Slutsky's theorem and its corollary in Chow and Teicher [7, 8.1], we obtain $Z_n \xrightarrow{\mathcal{D}} Z$ with

$$Z := \big(U^{(1)}, U^{(2)}, V^{(1)}, V^{(2)}, V^{(3)}, W^{(1)}, W^{(2)}, W^{(3)}\big),$$

where

$$U^{(1)} = \int_0^1 \mu_X(t)\,\mathrm{d}M(t) - M(1) \int_0^1 \mu_X(t)\,\mathrm{d}t,$$

$$V^{(1)} = \int_0^1 \big(\mu_X(t)\big)^2 \mathrm{d}t - \bigg(\int_0^1 \mu_X(t)\,\mathrm{d}t \bigg)^2,$$

$$W^{(1)} = M(1) \int_0^1 \big(\mu_X(t)\big)^2 \mathrm{d}t - \int_0^1 \mu_X(t)\,\mathrm{d}t \int_0^1 \mu_X(t)\,\mathrm{d}M(t).$$

Again by Slutsky's argument we obtain

$$\Big(n^{3/2}(\widetilde{\alpha}_n - \alpha_n),\, n^{1/2}(\widetilde{\mu}_{\varepsilon,n} - \mu_\varepsilon) \Big) = \bigg(\frac{n^{-3/2} U_n}{n^{-3} V_n},\, \frac{n^{-5/2} W_n}{n^{-3} V_n} \bigg) \\ \xrightarrow{\mathcal{D}} \bigg(\frac{U^{(1)}}{V^{(1)}},\, \frac{W^{(1)}}{V^{(1)}} \bigg).$$

The covariance matrix $\Sigma_{\gamma,\varepsilon}$ of the limiting normal distribution can be calculated using $\mathrm{d}M(t) = \sqrt{\varrho_{\gamma,\varepsilon}(t)}\,\mathrm{d}W(t)$ (see [10]). This relationship implies

$$\bigg(M(1),\, \int_0^1 \mu_X(t)\,\mathrm{d}M(t) \bigg) \stackrel{\mathcal{D}}{=} \mathcal{N}(0, \Sigma)$$

with

$$\Sigma := \begin{pmatrix} \int_0^1 \varrho_{\gamma,\varepsilon}(t)\,\mathrm{d}t & -\int_0^1 \mu_X(t)\varrho_{\gamma,\varepsilon}(t)\,\mathrm{d}t \\ -\int_0^1 \mu_X(t)\varrho_{\gamma,\varepsilon}(t)\,\mathrm{d}t & \int_0^1 \big(\mu_X(t)\big)^2 \varrho_{\gamma,\varepsilon}(t)\,\mathrm{d}t \end{pmatrix}.$$

Now the formula for $\Sigma_{\gamma,\varepsilon}$ follows, since $U^{(1)}$ and $W^{(1)}$ are linear combinations of $M(1)$ and $\int_0^1 \mu_X(t)\,\mathrm{d}M(t)$. □

References

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

[2] Al–Osh, M.A. and Alzaid, A.A. (1990). An integer–valued pth–order autoregressive structure (INAR(p)) process. *J. Appl. Prob.* **27**, 314–324.

[3] Arató, M., Pap, G. and Zuijlen, M.v. (2001). Asymptotic inference for spatial autoregression and orthogonality of Ornstein–Uhlenbeck sheets. *Computers Math. Appl.* **42**, 219–229.

[4] Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sons, New York.

[5] Cardinal, M., Roy, R. and Lambert, J. (1999). On the application of integer–valued time series models for the analysis of disease incidence. *Stat. Medicine* **18**, 2025–2039.

[6] Chan, N.H. and Wei, C.Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15**, 1050–1063.

[7] Chow, Y.S. and Teicher, H. (1978). *Probability Theory.* Springer–Verlag, New York, Heidelberg, Berlin.

[8] Du, J.G. and Li, Y. (1991). The integer–valued autoregressive INAR(p) model. *J. Time Ser. Anal.* **12**, 129–142.

[9] Franke, J. and Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modeling epileptic seizure counts. in Subba Rao, T. (ed.), *Developments in Time Series Analysis*, Chapman and Hall, London, pp. 310–330.

[10] Ispány, M., Pap, G. and Zuijlen, M. v. (2001). Asymptotic inference for nearly unstable INAR(1) models. Report No. 0109, University of Nijmegen, The Netherlands. Submitted to the *Journal of Applied Probability*.

[11] Klimko, L.A. and Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* **6**, 629–642.

[12] LATOUR, A. (1997). The multivariate GINAR(p) process. *Adv. Appl. Prob.* **29**, 228–248.

[13] SRIRAM, T.N. (1994). Invalidity of bootstrap for critical branching processes with immigration. *Ann. Statist.* **22**, 1013–1023.

[14] STEUTEL, F. and VAN HARN, K. (1979). Discrete analogues of self–decomposability and stability. *Ann. Probab.* **7**, 893–99.

[15] TANAKA, K. (1996). *Time Series Analysis. Nonstationary and Noninvertible Distribution Theory.* John Wiley & Sons, New York.

[16] VENKATARAMAN, K.N. and NANTHI, K. (1982). A limit theorem on subcritical Galton–Watson process with immigration. *Ann. Probab.* **10**, 1069–1074.