

# Asymptotic distribution of the Yule-Walker estimator for INAR( $p$ ) processes

Isabel Silva<sup>a,b</sup>, M. Eduarda Silva<sup>b,c</sup>

<sup>a</sup>*Departamento de Engenharia Civil, Faculdade de Engenharia - Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, PORTUGAL*

<sup>b</sup>*Departamento de Matemática Aplicada, Faculdade de Ciências - Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, PORTUGAL*

<sup>c</sup>*UI&D Matemática e Aplicações, Universidade de Aveiro, Aveiro, PORTUGAL*

---

## Abstract

The INteger-valued AutoRegressive (INAR) processes were introduced in the literature by Al-Osh and Alzaid (1987) and McKenzie (1988) for modelling correlated series of counts. These processes have been considered as the discrete counter part of AR processes, but their highly nonlinear characteristics lead to some statistically challenging problems, namely in parameter estimation. Several estimation procedures have been proposed in the literature, mainly for processes of first order. For some of these estimators the asymptotic properties as well as finite sample properties have been obtained and studied. This paper considers Yule-Walker parameter estimation for the  $p$ th-order integer-valued autoregressive, INAR( $p$ ), process. In particular, the asymptotic distribution of the Yule-Walker estimator is obtained and it is shown that this estimator is asymptotically normally distributed, unbiased and consistent.

*Key words:* INAR Process, Autocovariance distribution, Yule-Walker Estimation, Delta Method.

---

## 1 Introduction

Recently, there has been a growing interest in modelling non-negative integer-valued time series and, specially, time series of counts. Several models have been proposed and, in particular, the INteger-valued AutoRegressive, INAR,

---

*Email addresses:* [ims@fe.up.pt](mailto:ims@fe.up.pt) (Isabel Silva), [mesilva@fc.up.pt](mailto:mesilva@fc.up.pt) (M. Eduarda Silva).

model has been the subject of study in several papers. The  $p$ th-order integer-valued autoregressive, INAR( $p$ ), process is defined as follows (Latour, 1998). A discrete time non-negative integer-valued stochastic process,  $\{X_t\}$ , is said to be an INAR( $p$ ) process if it satisfies the following equation

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \cdots + \alpha_p * X_{t-p} + e_t, \quad (1)$$

where

- (1)  $\{e_t\}$ , designated the innovation process, is a sequence of independent and identically distributed (i.i.d) non-negative integer-valued random variables with  $E[e_t] = \mu_e$ ,  $\text{Var}[e_t] = \sigma_e^2$ ,  $E[e_t^3] = \gamma_e$  and  $E[e_t^4] = \kappa_e$ ;
- (2) the symbol  $*$  represents the thinning operation (Steutel and Van Harn, 1979; Gauthier and Latour, 1994), defined by

$$\alpha_i * X_{t-i} = \sum_{j=1}^{X_{t-i}} Y_{i,j}, \quad i = 1, \dots, p,$$

where  $\{Y_{i,j}\}$ , designated the counting series, is a set of i.i.d. non-negative integer-valued random variables such that  $E[Y_{i,j}] = \alpha_i$ ,  $\text{Var}[Y_{i,j}] = \sigma_i^2$ ,  $E[Y_{i,j}^3] = \gamma_i$  and  $E[Y_{i,j}^4] = \kappa_i$ . All the counting series are assumed independent of  $\{e_t\}$ ;

- (3)  $0 \leq \alpha_i < 1$ ,  $i = 1, \dots, p-1$ , and  $0 < \alpha_p < 1$ .

Alternatively, an INAR( $p$ ) process can be represented as a  $p$ -dimensional INAR(1) process (Franke and Subba Rao, 1995). Accordingly, by using the vector thinning operation, defined as a random vector whose  $i$ th component is given by (Franke and Subba Rao, 1995)

$$[\mathbf{A} * \mathbf{X}]_i = \sum_{j=1}^p a_{ij} * X_j, \quad i = 1, \dots, p,$$

where  $\mathbf{X} = [X_1 \cdots X_p]^T$  is a random vector,  $\mathbf{A}$  is a  $p \times p$  matrix with entries  $a_{ij}$  satisfying  $0 \leq a_{ij} \leq 1$ , for  $i, j = 1, \dots, p$ , and the counting series of all  $a_{ij} * X_j$ ,  $i, j = 1, \dots, p$ , are assumed independent, the INAR( $p$ ) process defined in (1) can be written as

$$\mathbf{X}_t = \mathbf{A} * \mathbf{X}_{t-1} + \mathbf{W}_t, \quad (2a)$$

$$X_t = \mathbf{H}\mathbf{X}_t, \quad (2b)$$

where  $\mathbf{H} = [1 \ 0 \ \cdots \ 0]$ ,  $\mathbf{X}_t = [X_t \ X_{t-1} \ \cdots \ X_{t-p+1}]^T$ ,  $\mathbf{W}_t = [e_t \ 0 \ \cdots \ 0]^T$ , for  $\{e_t\}$  a sequence of i.i.d. random variables, with  $E[e_t] = \mu_e$ ,  $\text{Var}[e_t] = \sigma_e^2$ ,  $E[e_t^3] = \gamma_e$ , and  $E[e_t^4] = \kappa_e < \infty$ , and

$$\mathbf{A} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{p-1} & \alpha_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

Furthermore, the model (2a) can be expressed as

$$\mathbf{X}_t \stackrel{d}{=} \mathbf{A}^k * \mathbf{X}_{t-k} + \sum_{j=0}^{k-1} \mathbf{A}^j * \mathbf{W}_{t-j}, \quad (3)$$

where  $\stackrel{d}{=}$  stands for equality in distribution. Then, equations (2) can be written as

$$\mathbf{X}_t \stackrel{d}{=} \sum_{j=0}^{\infty} \mathbf{A}^j * \mathbf{W}_{t-j}, \quad (4a)$$

$$X_t \stackrel{d}{=} \mathbf{H}\mathbf{X}_t, \quad (4b)$$

where  $\mathbf{H} = [1 \ 0 \ \cdots \ 0]$ , provided the spectral radius of the matrix  $\mathbf{A}$  is less than one, that is,  $\rho(\mathbf{A}) < 1$ .

The existence and stationarity condition for the INAR( $p$ ) processes is that the roots of  $z^p - \alpha_1 z^{p-1} - \cdots - \alpha_{p-1} z - \alpha_p = 0$  lie inside the unit circle (Du and Li, 1991) or, equivalently, that  $\sum_{j=1}^p \alpha_j < 1$ , (Latour, 1997, 1998). Probabilistic characteristics of the INAR models, in terms of second and third-order moments and cumulants, have been obtained by Silva and Oliveira (2004, 2005). In particular, it is found that the autocovariance function,  $R(\cdot)$ , satisfies a set of Yule-Walker type difference equations, which can be written in scalar and vectorial form as

$$\begin{cases} R(0) = V_p + \sum_{i=1}^p \alpha_i R(i), \\ R(k) = \sum_{i=1}^p \alpha_i R(i-k), \end{cases} \iff \mathbf{R}_p \boldsymbol{\alpha} = \begin{bmatrix} -V_p & 0 & \cdots & 0 \end{bmatrix}^T,$$

respectively, where

$$\mathbf{R}_p = \begin{bmatrix} R(0) & R(1) & \cdots & R(p) \\ R(1) & R(0) & \cdots & R(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ R(p) & R(p-1) & \cdots & R(0) \end{bmatrix}, \quad \boldsymbol{\alpha} = \begin{bmatrix} -1 \\ \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix},$$

and  $V_p$  is defined by

$$V_p = \sigma_e^2 + \mu_X \sum_{i=1}^p \sigma_i^2,$$

where  $\sigma_e^2 = \text{Var}[e_t]$ ,  $\mu_X = \text{E}[X_t] = \mu_e / (1 - \sum_{i=1}^p \alpha_i)$ , and  $\sigma_i^2$  is the variance of the counting series,  $Y_{j,i}$  ( $j = 1, \dots, X_{t-i}$ ;  $i = 1, \dots, p$ ), of the  $i$ th thinning operation,  $\alpha_i * X_{t-i}$ .

Several estimation methods for the INAR( $p$ ) model have been proposed in the literature both in the time domain, namely, Yule-Walker and Conditional Least Squares (Du and Li, 1991; Latour, 1998), Generalized Method of Moments (Brännäs, 1994, 1995), Conditional Maximum Likelihood (Franke and Seligmann, 1993; Franke and Subba Rao, 1995) and in the frequency domain, the Whittle Criterion (Silva and Oliveira, 2004, 2005). Methods based on higher-order statistics (moments and cumulants) have also been considered (I. Silva, 2005). The small sample properties of these estimators have been studied empirically by Monte Carlo methods (I. Silva, 2005) but little is known about the asymptotic properties. It must be said that these depend on higher order moments and cumulants which are very difficult to obtain.

Here the asymptotic distribution of the sample autocovariance for INAR( $p$ ) processes is obtained and it is shown that the Yule-Walker estimator is asymptotically normally distributed, unbiased and consistent. This result generalizes the work of Park and Oh (1997), who derived the asymptotic distribution of the Yule-Walker estimator for an alternative parametrization of the Poisson INAR(1) process with binomial thinning operation.

In Section 2.1 the asymptotic properties of the sample mean and sample autocovariance matrix are obtained and in Section 2.2 expressions for the asymptotic covariance matrix of the Yule-Walker estimator are provided.

## 2 Yule-Walker Estimation of the INAR( $p$ ) model

Given a realization  $\{X_1, \dots, X_N\}$  from an INAR( $p$ ) process, the Yule-Walker estimator  $\hat{\alpha}$  of  $\alpha$  is obtained by solving the following system of linear equations,

$$\hat{\mathbf{R}}_{p-1} \hat{\alpha} = \hat{\mathbf{r}}_p \Leftrightarrow \begin{bmatrix} \hat{R}(0) & \hat{R}(1) & \hat{R}(2) & \cdots & \hat{R}(p-1) \\ \hat{R}(1) & \hat{R}(0) & \hat{R}(1) & \cdots & \hat{R}(p-2) \\ \hat{R}(2) & \hat{R}(1) & \hat{R}(0) & \cdots & \hat{R}(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{R}(p-1) & \hat{R}(p-2) & \hat{R}(p-3) & \cdots & \hat{R}(0) \end{bmatrix} \begin{bmatrix} \hat{\alpha}_1 \\ \vdots \\ \vdots \\ \vdots \\ \hat{\alpha}_p \end{bmatrix} = \begin{bmatrix} \hat{R}(1) \\ \hat{R}(2) \\ \hat{R}(3) \\ \vdots \\ \hat{R}(p) \end{bmatrix},$$

with  $\hat{R}(k) = \frac{1}{N} \sum_{t=1}^{N-k} (X_t - \bar{X})(X_{t+k} - \bar{X})$ ,  $k \in \mathbb{Z}$ , and  $\bar{X} = \frac{1}{N} \sum_{t=1}^N X_t$ .

The estimators for  $\mu_e$  and  $\sigma_e^2$  are, respectively,

$$\hat{\mu}_e = \bar{X} \left( 1 - \sum_{i=1}^p \hat{\alpha}_i \right), \quad \hat{\sigma}_e^2 = \hat{V}_p - \bar{X} \sum_{i=1}^p \hat{\sigma}_i^2,$$

where

$$\hat{V}_p = \hat{R}(0) - \sum_{i=1}^p \hat{\alpha}_i \hat{R}(i)$$

and  $\hat{\sigma}_i^2$  is an estimator of the variance of the counting series for the  $i$ -th thinning operation,  $\alpha_i * X_{t-i}$ ,  $i = 1, \dots, p$ . The estimation of  $\hat{\sigma}_i^2$  depends on the distribution of the counting series. For instance, in the case of the binomial thinning operation (when the counting series are Bernoulli distributed),  $\hat{\sigma}_i^2 = \hat{\alpha}_i(1 - \hat{\alpha}_i)$ , for  $i = 1, \dots, p$ .

Now, in order to obtain the asymptotic distribution of the Yule-Walker estimator, the asymptotic properties of the sample covariance matrix are needed. These properties are obtained in 2.1 and the asymptotic distribution of the Yule-Walker estimator follows in 2.2.

### 2.1 Asymptotic distribution of the autocovariance function

In the following theorem the asymptotic multivariate normality of the sample mean,  $\bar{X}$ , and of the sample autocovariance function,  $\hat{R}(k)$ , is established.

The details of the proof are omitted since it follows closely Brockwell and Davis (1991, Chap. 7) by considering an auxiliary function

$$R^*(k) = \frac{1}{N} \sum_{t=1}^N (X_t - \mu_X)(X_{t+k} - \mu_X), \quad k \in \mathbb{N}_0,$$

obtaining its asymptotic properties and proving that  $\hat{R}(\cdot)$  has the same asymptotic properties as the auxiliary function,  $R^*(\cdot)$ .

**Theorem 1** *If  $\{X_t\}$  is an INAR( $p$ ) process with representation (4), where  $\sum_{j=0}^{\infty} |\mathbf{A}^j| < \infty$ , and if  $R(\cdot)$  is the autocovariance function of  $\{X_t\}$ , then for any non-negative integer  $h$ ,*

$$\begin{bmatrix} \bar{X} \\ \hat{R}(0) \\ \vdots \\ \hat{R}(h) \end{bmatrix} \text{ is AN } \left( \begin{bmatrix} \mu_X \\ R(0) \\ \vdots \\ R(h) \end{bmatrix}, N^{-1} \mathbf{V}_R \right)$$

where

$$\mathbf{V}_R = \begin{bmatrix} [V_{11}]_{1 \times 1} & [V_{12}]_{1 \times (h+1)} \\ [V_{12}]_{1 \times (h+1)}^T & [V_{22}]_{(h+1) \times (h+1)} \end{bmatrix}$$

is an  $(h + 2) \times (h + 2)$  matrix for any non-negative integer  $h$ , such that

$$\begin{aligned} V_{11} &= \lim_{N \rightarrow \infty} N \text{Var}(\bar{X}) \\ &= \sum_{h=-\infty}^{\infty} R(h), \end{aligned} \quad (5a)$$

$$\begin{aligned} [V_{12}]_{k+1} &= \lim_{N \rightarrow \infty} N \text{Cov}(\bar{X}, R^*(k)), \quad k = 0, 1, \dots, h, \\ &= \sum_{h=-\infty}^{\infty} C_X(h, k + h), \end{aligned} \quad (5b)$$

$$\begin{aligned} [V_{22}]_{k+1, j+1} &= \lim_{N \rightarrow \infty} N \text{Cov}(R^*(k), R^*(j)), \quad k, j = 0, 1, \dots, h, \\ &= \sum_{h=-\infty}^{\infty} R(h)R(h + j - k) + R(h + j)R(h - k) \\ &\quad + C_Y(h, k, j + h), \end{aligned} \quad (5c)$$

where  $C_X(\cdot, \cdot)$  is the third-order cumulant of  $X_t$  and  $C_Y(\cdot, \cdot, \cdot)$  is the fourth-order cumulant of the process  $Y_t = X_t - \mu_X$  (I. Silva, 2005).

**PROOF.** Note that for  $0 \leq k \leq h$ ,

$$\begin{aligned} \sqrt{N}(R^*(k) - \hat{R}(k)) &= \frac{1}{\sqrt{N}} \sum_{t=N-k+1}^N (X_t - \mu_X)(X_{t+k} - \mu_X) \\ &\quad + \frac{1}{\sqrt{N}} \left( -\mu_X \sum_{t=1}^{N-k} X_t - \mu_X \sum_{t=1}^{N-k} X_{t+k} - (N-k)\mu_X^2 \right. \\ &\quad \left. + \bar{X} \sum_{t=1}^{N-k} X_t + \bar{X} \sum_{t=1}^{N-k} X_{t+k} - (N-k)\bar{X}^2 \right) \\ &= \frac{1}{\sqrt{N}} \sum_{t=N-k+1}^N (X_t - \mu_X)(X_{t+k} - \mu_X) \\ &\quad + \frac{N}{\sqrt{N}} (\bar{X} - \mu_X) \left( \frac{1}{N} \sum_{t=1}^{N-k} X_t + \frac{1}{N} \sum_{t=1}^{N-k} X_{t+k} \right. \\ &\quad \left. - \left(1 - \frac{k}{N}\right) (\bar{X} + \mu_X) \right). \end{aligned}$$

The first term,  $\frac{1}{\sqrt{N}} \sum_{t=N-k+1}^N (X_t - \mu_X)(X_{t+k} - \mu_X)$ , is  $o_p(1)$ , since

$$\frac{1}{\sqrt{N}} \mathbb{E} \left[ \left| \sum_{t=N-k+1}^N (X_t - \mu_X)(X_{t+k} - \mu_X) \right| \right] \leq \frac{1}{\sqrt{N}} k R(0) \xrightarrow{N \rightarrow \infty} 0.$$

The last term is also  $o_p(1)$ , since

$$\sqrt{N}(\bar{X} - \mu_X) \text{ is } O_p(1),$$

as  $N \rightarrow \infty$ , because  $\sqrt{N}(\bar{X} - \mu_X) \xrightarrow{d} Y$ , where  $Y \sim \mathcal{N}(0, V_{11})$ , and  $\xrightarrow{d}$  stands for convergence in distribution. Note that by the weak law of large numbers (Brockwell and Davis, 1991, p. 208),

$$\left( \frac{1}{N} \sum_{t=1}^{N-k} X_t + \frac{1}{N} \sum_{t=1}^{N-k} X_{t+k} - \left(1 - \frac{k}{N}\right) (\bar{X} + \mu_X) \right) \xrightarrow{P} 0.$$

This remark leads to the following result

$$\sqrt{N}(R^*(k) - \hat{R}(k)) \text{ is } o_p(1),$$

as  $N \rightarrow \infty$ . Then the proposition follows from the fact that

$$\begin{bmatrix} \bar{X} \\ R^*(0) \\ \vdots \\ R^*(h) \end{bmatrix} \text{ is } AN \left( \begin{bmatrix} \mu_X \\ R(0) \\ \vdots \\ R(h) \end{bmatrix}, N^{-1} \mathbf{V}_R \right)$$

and Proposition 6.3.3 of Brockwell and Davis (1991, p. 205).  $\square$

Note that this result is in agreement with the asymptotic distribution obtained by Park and Oh (1997) for the case of Poisson INAR(1) models with binomial thinning operation.

## 2.2 Asymptotic distribution of the Yule-Walker estimator

In this section, the asymptotic properties of the sample covariance matrix is used to obtain the asymptotic distribution of the Yule-Walker estimator. Hence, consider the  $p \times p$  Toeplitz sample autocovariance matrix,  $\hat{\mathbf{R}}_{p-1}$ , and the vector of sample autocovariance function with  $p$  elements,  $\hat{\mathbf{r}}_p$ . Let

$$\hat{\mathbf{V}}_{Rr} = \begin{bmatrix} \text{vec}(\hat{\mathbf{R}}_{p-1}) \\ \hat{\mathbf{r}}_p \end{bmatrix} \quad (6)$$

be a vector with  $p(p+1)$  elements, where the  $\text{vec}$  operator is the column-stacking operator of a matrix (Neudecker, 1969), that is,

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{A}_{\cdot 1} \\ \mathbf{A}_{\cdot 2} \\ \vdots \\ \mathbf{A}_{\cdot q} \end{bmatrix}$$

is a  $pq \times 1$  vector where  $\mathbf{A}_{\cdot i}$  is the  $i$ -th column of the  $p \times q$  matrix  $\mathbf{A}$ .

Note that for  $p = 1$ ,

$$\hat{\mathbf{V}}_{Rr} = \left[ \hat{R}(0) \ \hat{R}(1) \right]^T$$

and for  $p > 1$ , the  $i$ -th element of  $\hat{\mathbf{V}}_{Rr}$  can be written as

$$\hat{\mathbf{V}}_{Rr}(i) = \begin{cases} \hat{R}(|(i-1) \bmod p - [(i-1)/p]|), & \text{if } i \leq p^2, \\ \hat{R}(|i \bmod p|), & \text{if } i > p^2, \end{cases}$$

where  $[a]$  represent the integer part of  $a \in \mathbb{R}$ .

Then, using the previous results it is found that

$$\hat{\mathbf{V}}_{Rr} = \begin{bmatrix} \text{vec}(\hat{\mathbf{R}}_{p-1}) \\ \hat{\mathbf{r}}_p \end{bmatrix} \text{ is } AN \left( \begin{bmatrix} \text{vec}(\mathbf{R}_{p-1}) \\ \mathbf{r}_p \end{bmatrix}, \frac{1}{N} \boldsymbol{\Sigma}_{Rr} \right),$$

where  $\boldsymbol{\Sigma}_{Rr}$  is the  $p(p+1) \times p(p+1)$  covariance matrix of  $\hat{\mathbf{V}}_{Rr}$ ,

$$\boldsymbol{\Sigma}_{Rr}(i, j) = \text{Cov}(\hat{\mathbf{V}}_{Rr}(i), \hat{\mathbf{V}}_{Rr}(j)), \quad (7)$$

defined as in (5c), for  $i, j = 1, \dots, p(p+1)$ .

Now, let

$$\mathbf{D}^T = - \begin{bmatrix} R(1)\mathbf{I}_p & \cdots & R(p)\mathbf{I}_p \end{bmatrix} \left( (\mathbf{R}_{p-1}^{-1})^T \otimes \mathbf{R}_{p-1}^{-1} \right) \begin{bmatrix} \mathbf{I}_{p^2} & \mathbf{0}_{p^2 \times p} \\ \mathbf{0}_{p \times p^2} & \mathbf{R}_{p-1}^{-1} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{p \times p^2} & \mathbf{R}_{p-1}^{-1} \end{bmatrix}, \quad (8)$$

where  $\mathbf{I}_n$  is the  $n \times n$  identity matrix,  $\mathbf{0}_{n \times m}$  the  $n \times m$  matrix of zeros and  $\otimes$  the Kronecker product (Graham, 1981) defined as follows. Let  $\mathbf{A}$  and  $\mathbf{B}$  be  $p \times q$  and  $m \times n$  matrices, respectively. Then  $\mathbf{A} \otimes \mathbf{B}$  is a  $pm \times qn$  matrix given by

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1q}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2q}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}\mathbf{B} & a_{p2}\mathbf{B} & \cdots & a_{pq}\mathbf{B} \end{bmatrix}.$$

The asymptotic distribution of the Yule-Walker estimator of an INAR( $p$ ) process is given in the following theorem.

**Theorem 2** *Let  $\{X_t\}$  be an INAR( $p$ ) process, satisfying (1), and  $\hat{\boldsymbol{\alpha}}$  the Yule-Walker estimator of  $\boldsymbol{\alpha}$ . Then*

$$N^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \text{ is } AN(\mathbf{0}_p, \mathbf{D}^T \boldsymbol{\Sigma}_{Rr} \mathbf{D}),$$

where  $\mathbf{0}_n$  is a vector of  $n$  zeros,  $\boldsymbol{\Sigma}_{Rr}$  is given in (7) and  $\mathbf{D}^T$  is defined by (8).

**PROOF.** The Delta method (van der Vaart, 1998) is used to demonstrate the result. Thus, let  $\varphi$  be the function from  $\mathcal{D}_\varphi \subset \mathbb{R}^{p(p+1)}$  into  $\mathbb{R}^p$  defined by

$$\varphi(\mathbf{X}) = (\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1} (\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_2),$$

where  $\mathbf{X}$  is a vector with  $p(p+1)$  elements,

$$\mathbf{C}_1 = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{0}_p^T \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} \mathbf{0}_p \\ 1 \end{bmatrix},$$

are a  $(p+1) \times p$  matrix and a vector with  $p+1$  elements, respectively,  $\mathbf{0}_p$  is a vector of  $p$  zeros and  $\text{unvec}_{n \times m}$  is the inverse of the  $\text{vec}$  operator, that is,  $\text{unvec}_{p \times q}(\mathbf{X})$ , for a vector  $\mathbf{X}$  of  $pq \times 1$  elements is defined as a  $p \times q$  matrix, such that  $\text{unvec}_{p \times q}(\text{vec}(\mathbf{A})) = \mathbf{A}$  (Swami and Giannakis, 1994).

Note that for  $\hat{\mathbf{V}}_{Rr}$  defined in (6)

$$\varphi(\hat{\mathbf{V}}_{Rr}) = \hat{\mathbf{R}}_{p-1}^{-1} \hat{\mathbf{r}}_p = \hat{\boldsymbol{\alpha}}.$$

Then, by the Delta method,

$$\hat{\boldsymbol{\alpha}} \text{ is } AN \left( \boldsymbol{\alpha}, \frac{1}{N} \mathbf{D}^T \boldsymbol{\Sigma}_{Rr} \mathbf{D} \right),$$

where

$$\mathbf{D} = \left( \frac{\partial \varphi}{\partial \mathbf{X}} \right) \Big|_{\mathbf{X}=\hat{\mathbf{V}}_{Rr}} = \left( \frac{\partial \varphi}{\partial \mathbf{X}^T} \right)^T \Big|_{\mathbf{X}=\hat{\mathbf{V}}_{Rr}}$$

is the  $p(p+1) \times p$  derivative matrix of the function  $\varphi$ , defined as follows. Let  $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be a vector valued function with vector variable. The  $n \times m$  matrix derivative of  $f$  is defined as (Rao and Rao, 1998, p. 225)

$$\frac{\partial \mathbf{f}}{\partial \mathbf{X}} \triangleq \left( \frac{\partial \mathbf{f}}{\partial \mathbf{X}^T} \right)^T \triangleq \begin{bmatrix} \frac{\partial \mathbf{f}_1}{\partial \mathbf{X}_1} & \cdots & \frac{\partial \mathbf{f}_1}{\partial \mathbf{X}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{f}_m}{\partial \mathbf{X}_1} & \cdots & \frac{\partial \mathbf{f}_m}{\partial \mathbf{X}_n} \end{bmatrix}^T,$$

in order to meet the needs of the Jacobian matrix.

Note that  $\varphi(\mathbf{X})$  can be written as the product of two matrices,  $\varphi(\mathbf{X}) = \mathbf{M}_1^{-1}(\mathbf{X}) \mathbf{M}_2(\mathbf{X})$ . Then, by using the product rule (Rao and Rao, 1998, p.

234), it is found that

$$\begin{aligned} \left( \frac{\partial \varphi}{\partial \mathbf{X}^T} \right)^T &= \left( \frac{\partial (\mathbf{M}_1^{-1} \mathbf{M}_2)}{\partial \mathbf{X}^T} \right)^T \\ &= \left( (\mathbf{M}_2^T \otimes \mathbf{I}_p) \frac{\partial \text{vec}(\mathbf{M}_1^{-1})}{\partial \mathbf{X}^T} + (\mathbf{I}_1 \otimes \mathbf{M}_1^{-1}) \frac{\partial \text{vec}(\mathbf{M}_2)}{\partial \mathbf{X}^T} \right)^T. \end{aligned}$$

Now, using the matrix derivative of an inverse matrix,

$$\frac{\partial \text{vec}(\mathbf{M}_1^{-1})}{\partial \mathbf{X}^T} = - \left( (\mathbf{M}_1^{-1})^T \otimes \mathbf{M}_1^{-1} \right) \frac{\partial \text{vec}(\mathbf{M}_1)}{\partial \mathbf{X}^T},$$

and then, replacing  $\mathbf{M}_1$  by  $\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1$ , and using the properties of the  $\text{vec}$  and  $\text{unvec}_{n \times m}$  operators (Neudecker, 1969) and the derivative matrix rule of  $\mathbf{f} : \mathbb{R}^k \rightarrow \mathbb{R}$  such that  $\mathbf{f} = \mathbf{A}\mathbf{X}$ , where  $\mathbf{A}$  is a  $n \times k$  constant matrix (Rao and Rao, 1998, p. 233), it is found that

$$\begin{aligned} \frac{\partial \text{vec}(\mathbf{M}_1^{-1})}{\partial \mathbf{X}^T} &= - \left( ((\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1})^T \otimes (\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1} \right) \\ &\quad \frac{\partial \text{vec}(\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)}{\partial \mathbf{X}^T} \\ &= - \left( ((\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1})^T \otimes (\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1} \right) \\ &\quad \frac{\partial (\mathbf{C}_1^T \otimes \mathbf{I}_p) \mathbf{X}}{\partial \mathbf{X}^T} \\ &= - \left( ((\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1})^T \otimes (\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1} \right) \\ &\quad (\mathbf{C}_1^T \otimes \mathbf{I}_p). \end{aligned}$$

In a similar manner

$$\frac{\partial \text{vec}(\mathbf{M}_2)}{\partial \mathbf{X}^T} = \frac{\partial \text{vec}(\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_2)}{\partial \mathbf{X}^T} = \frac{\partial (\mathbf{C}_2^T \otimes \mathbf{I}_p) \mathbf{X}}{\partial \mathbf{X}^T} = (\mathbf{C}_2^T \otimes \mathbf{I}_p).$$

Then,

$$\begin{aligned} \left( \frac{\partial \varphi}{\partial \mathbf{X}^T} \right)^T &= \left( -((\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_2)^T \otimes \mathbf{I}_p) \right. \\ &\quad \left( ((\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1})^T \otimes (\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1} \right) (\mathbf{C}_1^T \otimes \mathbf{I}_p) \\ &\quad \left. + (\mathbf{I}_1 \otimes (\text{unvec}_{p \times p+1}(\mathbf{X}) \mathbf{C}_1)^{-1}) (\mathbf{C}_2^T \otimes \mathbf{I}_p) \right)^T. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbf{D}^T &= -(\mathbf{r}_p^T \otimes \mathbf{I}_p) ((\mathbf{R}_{p-1}^{-1})^T \otimes \mathbf{R}_{p-1}^{-1}) (\mathbf{C}_1^T \otimes \mathbf{I}_p) \\ &\quad + (\mathbf{I}_1 \otimes \mathbf{R}_{p-1}^{-1}) (\mathbf{C}_2^T \otimes \mathbf{I}_p) \end{aligned}$$

$$= - \begin{bmatrix} R(1)\mathbf{I}_p \cdots R(p)\mathbf{I}_p \\ \mathbf{0}_{p \times p^2} \mathbf{R}_{p-1}^{-1} \end{bmatrix} \left( (\mathbf{R}_{p-1}^{-1})^T \otimes \mathbf{R}_{p-1}^{-1} \right) \begin{bmatrix} \mathbf{I}_{p^2} \mathbf{0}_{p^2 \times p} \\ \mathbf{0}_{p \times p^2} \mathbf{R}_{p-1}^{-1} \end{bmatrix},$$

and the theorem follows.  $\square$

For the particular case of INAR(1) processes with Poisson,  $\mathcal{P}(\lambda)$ , innovations and binomial thinning operation, it is found that  $\mathbf{R}_0 = \begin{bmatrix} R(0) \end{bmatrix} = \begin{bmatrix} \frac{\lambda}{(1-\alpha)} \end{bmatrix}$ ,

$$\begin{aligned} \boldsymbol{\Sigma}_{Rr} &= \begin{bmatrix} \text{Var}(\hat{R}(0)) & \text{Cov}(\hat{R}(0), \hat{R}(1)) \\ \text{Cov}(\hat{R}(0), \hat{R}(1)) & \text{Var}(\hat{R}(1)) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\lambda(2\lambda(1+\alpha^2)+(1-\alpha)(1+\alpha)^2)}{(1-\alpha)^2(1-\alpha^2)} & \frac{2\lambda\alpha(1+2\lambda-\alpha^2)}{(1-\alpha)^2(1-\alpha^2)} \\ \frac{2\lambda\alpha(1+2\lambda-\alpha^2)}{(1-\alpha)^2(1-\alpha^2)} & \frac{\lambda(\lambda(1+4\alpha^2-\alpha^4)+\alpha(1-\alpha)(1+\alpha)^2)}{(1-\alpha)^2(1-\alpha^2)} \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \mathbf{D}^T &= - \begin{bmatrix} R(1) \end{bmatrix} \left( \begin{bmatrix} 1/R(0) \end{bmatrix} \otimes \begin{bmatrix} 1/R(0) \end{bmatrix} \right) \begin{bmatrix} 1 & 0 \\ 0 & 1/R(0) \end{bmatrix} \\ &= \begin{bmatrix} -\frac{R(1)}{R(0)^2} & \frac{1}{R(0)} \end{bmatrix} = \begin{bmatrix} \frac{-\alpha(1-\alpha)}{\lambda} & \frac{1-\alpha}{\lambda} \end{bmatrix}. \end{aligned}$$

Hence,  $\mathbf{D}^T \boldsymbol{\Sigma}_{Rr} \mathbf{D} = (1 - \alpha^2) + \frac{\alpha(1 - \alpha)^2}{\lambda}$ . Therefore, Theorem 2 states that

$$N^{1/2} (\hat{\alpha} - \alpha) \text{ is } AN \left( 0, (1 - \alpha^2) + \frac{\alpha(1 - \alpha)^2}{\lambda} \right),$$

which is in agreement with the asymptotic distribution obtained by Park and Oh (1997) for the  $\alpha$  parameter.

## References

- Al-Osh, M.A. and Alzaid, A.A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, Vol. 8, pp. 261-275.
- Brännäs, K. (1994). Estimation and testing in integer-valued AR(1) models. Umeå Economic Studies, 335, Umeå University.
- Brännäs, K. (1995). Explanatory variables in the AR(1) count data model. Umeå Economic Studies, 381, Umeå University.
- Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd. edition, Springer-Verlag, New York.

- Du, Jin-Guan and Li, Yuan (1991). The integer-valued autoregressive (INAR( $p$ )) model. *Journal of Time Series Analysis*, Vol. 12, pp. 129-142.
- Franke, J. and Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In *Developments in Time Series Analysis* (ed. T. Subba Rao). Chapman and Hall, London. pp. 310-330
- Franke, J. and Subba Rao, T. (1995). Multivariate first order integer valued autoregressions. *Technical Report*, Math Dep., UMIST.
- Gauthier, G. and Latour, A. (1994). Convergence forte des estimateurs des paramètres d'un processus GENAR( $p$ ). *Annales des Sciences Mathématiques du Québec*, Vol. 18, pp. 49-71.
- Graham, A. (1981). *Kronecker products and matrix calculus: with applications*. Ellis Horwood, Chichester.
- Latour, A. (1997). The Multivariate GINAR( $p$ ) process. *Advances in Applied Probability*, Vol. 29, pp. 228-248.
- Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *Journal of Time Series Analysis*, Vol. 19, pp. 439-455.
- McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability*, Vol. 20, pp. 822-835.
- Neudecker, H. (1969). Some theorems on matrix differentiation with special reference to Kronecker matrix product. *Journal of the American Statistical Association*, Vol. 64, pp. 953-963.
- Park, Y. and Oh, C.W. (1997). Some asymptotic properties in INAR(1) processes with Poisson marginals. *Statistical Papers*, Vol. 38, pp. 287-302.
- Rao, C.R. and Rao, M.B. (1998). *Matrix algebra and its applications to statistics and econometrics*. World Scientific. Singapore.
- Silva, I. (2005). *Contributions to the analysis of discrete-valued time series*. PhD Thesis. Universidade do Porto, Portugal.
- Silva, M.E. and Oliveira, V.L. (2004). Difference equations for the higher order moments and cumulants of the INAR(1) model. *Journal of the Time Series Analysis*, Vol. 25, pp. 317-333.
- Silva, M.E. and Oliveira, V.L. (2005). Difference equations for the higher order moments and cumulants of the INAR( $p$ ) model. *Journal of the Time Series Analysis*, Vol. 26, pp. 17-36.
- Steutel, F. W. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, Vol. 7, pp. 893-899.
- Swami, A. and Giannakis, G. (1994). Multichannel ARMA processes. *IEEE Transactions on Signal Processing*, Vol. 42, pp. 898-913.
- van der Vaart, A.W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge.