DIFFERENCE EQUATIONS FOR THE HIGHER ORDER MOMENTS AND CUMULANTS OF THE INAR(p) MODEL

By Maria Eduarda Silva and Vera Lúcia Oliveira

Universidade do Porto

First Version received May 2001

Abstract. Here we obtain difference equations for the higher order moments and cumulants of a time series $\{X_t\}$ satisfying an INAR(p) model. These equations are similar to the difference equations for the higher order moments and cumulants of the bilinear time series model. We obtain the spectral and bispectral density functions for the INAR(p) process in state–space form, thus characterizing it in the frequency domain. We consider a frequency domain method – the Whittle criterion – to estimate the parameters of the INAR(p) model and illustrate it with the series of the number of epilepsy seizures of a patient.

Keywords. INAR models; moments; cumulants; Yule–Walker equations; estimation; frequency domain; Whittle's criterion.

1. INTRODUCTION

Recently, there has been a growing interest in modelling discrete time-stationary processes with discrete marginal distributions. The usual linear models for time series, the well-known autoregressive moving-average (ARMA) models, are suitable for modelling stationary-dependent sequences under the assumption of Gaussianity. However, this assumption is inappropriate for modelling counting processes. Thus, motivated by the need for modelling correlated series of counts, several models for integer-valued time series were proposed in the literature. One of these models is the integer autoregressive (INAR) model proposed by Al-Osh and Alzaid (1987), Alzaid and Al-Osh (1990), and Du and Li (1991). These models are based on the Steutal and Van Harn (1979) thinning operation, a binomial operator, and have Poisson marginal distribution. A generalization of the INAR models, the GINAR models, have been proposed by Gauthier and Latour (1994), Latour (1997, 1998), which are based on a generalized Steutal and Van Harn thinning operator defined as follows.

Consider a non-negative integer-valued random variable X and $\alpha \in [0, 1]$, then the generalized thinning operation, denoted by '$*$', is defined as

$$\alpha * X = \sum_{k=1}^{X} Y_k \tag{1}$$

0143-9782/05/01 17–36 JOURNAL OF TIME SERIES ANALYSIS Vol. 26, No. 1

where $\{Y_k\}$, $k = 1,\ldots,X$ is a sequence of independently and identically distributed non-negative integer-valued random variables, independent of X, with finite mean α and variance σ^2. This sequence is called the counting series of $\alpha * X$. Note that in the Steutal and Van Harn definition, $\{Y_k\}$ is a sequence of Bernoulli random variables. For an account of the properties of the thinning operation, see Gauthier and Latour (1994), and Silva and Oliveira (2000a,b).

Now, $\{X_t\}$ a discrete time, positive integer-valued stochastic process, is said to be an INAR(p) process if it satisfies the following equation:

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \cdots + \alpha_p * X_{t-p} + e_t \tag{2}$$

where

1. $\{e_t\}$ is a sequence of independently and identically distributed integer-valued random variables, with $E(e_t) = \mu_e$, $V(e_t) = \sigma_e^2$ and $E(e_t^3) = \gamma_e$,
2. All counting series of $\alpha_i * X_{t-i}$, $i = 1,\ldots,p$, $\{Y_{i,k}\}$, $k = 1,\ldots,X_{t-i}$ are mutually independent, and independent of $\{e_t\}$, and such that $E(Y_{i,k}) = \alpha_i$, $V(Y_{i,k}) = \sigma_i^2$ and $E(Y_{i,k}^3) = \gamma_i$,
3. $0 < \alpha_i < 1$, $i = 1,\ldots,p$.

This definition has been proposed independently by Du and Li (1991) and Gauthier and Latour (1994) and is different from that of Alzaid and Al-Osh (1990) which assumes that the conditional distribution of the vector $(\alpha_1 * X_t, \alpha_2 * X_t,\ldots,\alpha_p * X_t)$ given $X_t = x_t$ is multinomial with parameters $(\alpha_1, \alpha_2,\ldots,\alpha_p, x_t)$ and is independent of past history of the process. This means that, given $X_t = x_t$ the random variable $\alpha_1 * X_t$ is independent of X_{t-k} and its survivals $\alpha_j * X_{t-k}$ for $i, j = 1, 2,\ldots,p$ and $k > 0$. The two different formulations imply different second-order structure for the processes: the Du and Li formulation implies that the autocorrelation function of the process is the same as that of an AR(p) whilst that of Alzaid and Al-Osh formulation is the same as an ARMA(p, $p - 1$).

The INAR(p) process defined in (2) can be viewed as an extension (or a special type) of a branching process with immigration; see Dion *et al.* (1995) for a complete account of this relationship.

Now, model (2) can be written in state–space form as follows

$$\begin{aligned} \mathbf{X}_t &= \mathbf{A} * \mathbf{X}_{t-1} + \mathbf{C}e_t \\ X_t &= \mathbf{H}\mathbf{X}_t \end{aligned} \tag{3}$$

where $\mathbf{X}_t = (X_t, X_{t-1},\ldots,X_{t-p+1})^T$, $\mathbf{A}$ is a $p \times p$ matrix of the form

$$\mathbf{A} = \begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_{p-1} & \alpha_p \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix} \tag{4}$$

the elements of which a_{ij}, $i = 1,\ldots,p$, $j = 1,\ldots,p$, satisfy $0 \le a_{i,j} \le 1$, $\mathbf{C} = [1 \quad 0 \quad \ldots \quad 0]^T$, $\mathbf{H} = [1 \quad 0 \quad \ldots \quad 0]$ and '$*$' is the vectorial thinning operation defined as follows. For a integer-valued random vector $\mathbf{X} = (X_1,\ldots,X_p)^T$ we define $\mathbf{A} * \mathbf{X}$, as the integer-valued vector the ith component of which, $i = 1,\ldots,p$ is given by Franke and Subba Rao (1995)

$$(\boldsymbol{A} * \boldsymbol{X})_i = \left(\sum_{j=1}^{p} a_{ij} * X_j\right) \tag{5}$$

where we assume that all the counting series of all $a_{ij} * X_j$ are independent, $i = 1,\ldots,p$, $j = 1,\ldots,p$.

The existence and the stationarity conditions were obtained by Du and Li (1991) for the Poisson INAR(p) process, defined with Bernoulli thinning operation and by Latour (1997, 1998) for the generalized INAR(p) process defined by (3). The stationary condition is that $\rho(\mathbf{A}) < 1$, where $\rho(\mathbf{A})$ is the spectral radius of matrix $\mathbf{A}$. Moreover, Franke and Subba Rao (1995) prove that if $\{e_t\}$ satisfies $0 < P(e_t = 0) < 1$ then any solution of (3) is an irreducible aperiodic chain in N_0^p.

We obtain difference equations for the higher order moments and cumulants of the INAR(p) process, defined in (2). These equations may be used to obtain moment estimates for the parameters. We also obtain the spectral and bispectral density functions, thus characterizing the process in the frequency domain. The knowledge of the spectral density function suggests estimating the parameters of the model using frequency domain methods, namely the Whittle criterion. This turns out to be especially useful for the INAR(p) model since the likelihood function is, in this case, quite difficult to obtain.

2. SECOND- AND THIRD-ORDER MOMENTS FOR THE INAR(p) PROCESS

Let $\{X_t\}$ be a stationary time series for which moments up to the order k exist. If the time series $\{X_t\}$ is kth-order stationary, then the kth-order joint moment of $X_{t_1},\ldots,X_{t_k}$, or, equivalently, of X_t, $X_{t+s_1},\ldots,X_{t+s_{k-1}}$, is a function of $k-1$ variables which we denote by $\mu(s_1,\ldots,s_{k-1}) = E(X_t X_{t+s_1},\ldots,X_{t+s_{k-1}})$, and $\mu = E(X_t)$. The corresponding moments about the mean will be denoted by $R(s_1,\ldots,s_{k-1})$. We define $\boldsymbol{\mu} = E(\mathbf{X}_t)$, $\boldsymbol{\mu}(s) = E(\mathbf{X}_t \mathbf{X}_{t+s}^T)$, $\boldsymbol{\mu}(s_1, s_2) = E(X_t \mathbf{X}_{t+s_1} \mathbf{X}_{t+s_2}^T)$ and $\mathbf{R}(s) = E((\mathbf{X}_t - \boldsymbol{\mu})(\mathbf{X}_{t+s} - \boldsymbol{\mu})^T)$, for $\mathbf{X}_t = (X_t, X_{t-1},\ldots,X_{t-p+1})^T$. In this section, we obtain the second- and third-order moments of a third-order stationary INAR(p) process, $\{X_t\}$, satisfying (3). Explicit expressions for the raw moments $\mu_k = E(X_t^k)$, $k \ge 1$ and Yule–Walker difference equations for the lagged second- and third-order moments (and cumulants) of the INAR(p) model are obtained. The scalar expressions for the difference equations are recovered using results from the calculus of Kronecker product matrices.

The Kronecker product of two square matrices $\mathbf{M}_{m\times m} = [M_{ij}]$ and $\mathbf{N}_{n\times n} = [N_{ij}]$, in that order, is defined by

$$\begin{aligned}\mathbf{M}\otimes\mathbf{N} &= [M_{ij}\mathbf{N}]_{mn\times mn}\\ &= \begin{bmatrix} M_{11}\mathbf{N} & M_{12}\mathbf{N} & \dots & M_{1m}\mathbf{N}\\ M_{21}\mathbf{N} & M_{22}\mathbf{N} & \dots & M_{2m}\mathbf{N}\\ \vdots & \vdots & & \vdots\\ M_{m1}\mathbf{N} & M_{m2}\mathbf{N} & \dots & M_{mm}\mathbf{N}\end{bmatrix}\end{aligned} \tag{6}$$

Let $\mathbf{L}_{l\times l} = [L_{ij}]$ be a $l\times l$ matrix. We define vec$\{\mathbf{L}\}$ as the vector obtained by stacking the columns of the square matrix $\mathbf{L}$, that is to say,

$$\mathrm{vec}\{\mathbf{L}\} = \begin{bmatrix} L_{\cdot 1}\\ L_{\cdot 2}\\ \vdots\\ L_{\cdot l}\end{bmatrix}$$

where $L_{\cdot j}$, $j = 1,\dots,l$, denotes the jth column of $\mathbf{L}$. For an account of the Kronecker product properties, see Searle (1982) and Graham (1981).

2.1. *Difference equations in terms of moments*

Let $\{X_t\}$ be a stationary process satisfying (2). We obtain equations for the raw moments and difference equations for the lagged moments of the process in its state space form, (3). Under the above conditions it follows easily that

$$\boldsymbol{\mu} = E(\mathbf{X}_t) = (\mathbf{I}_p - \mathbf{A})^{-1}\mathbf{C}\mu_e \tag{7}$$

the scalar form of which is

$$\mu = E(X_t) = \frac{\mu_e}{1-\sum_{i=1}^{p}\alpha_i} \tag{8}$$

Moreover,

$$\begin{aligned}\boldsymbol{\mu}(0) &= E(\mathbf{X}_t\mathbf{X}_t^T)\\ &= \mathbf{A}E[\mathbf{X}_{t-1}\mathbf{X}_{t-1}^T]\mathbf{A}^T + \mathbf{C}_1 + E[e_t\mathbf{C}\mathbf{X}_{t-1}^T]\mathbf{A}^T\\ &\quad + \mathbf{A}E[\mathbf{X}_{t-1}e_t\mathbf{C}^T] + \mathrm{E}[e_t^2\mathbf{C}\mathbf{C}^T]\\ &= \mathbf{A}\boldsymbol{\mu}(0)\mathbf{A}^T + \mathbf{C}_1 + \mu_e\mathbf{C}\boldsymbol{\mu}^T\mathbf{A}^T + \mu_e\mathbf{A}\boldsymbol{\mu}\mathbf{C}^T + (\mu_e^2+\sigma_e^2)\mathbf{C}\mathbf{C}^T\end{aligned} \tag{9}$$

The matrix $\mathbf{C}_1$ is a diagonal matrix with elements

$$\mathbf{C}_1(j,j) = \mu\sum_{i=1}^{p}\sigma_i^2, \quad j = 1,\dots,p,$$

resulting from the properties of the vectorial thinning operation.

Applying the vec operator and its properties to (9) we can write

$$\begin{aligned}\mathrm{vec}\{\boldsymbol{\mu}(0)\} = (\mathbf{I}_{p^2} - \mathbf{A}\otimes\mathbf{A})^{-1}[&\mathrm{vec}\{\mathbf{C}_1\} + \mu_e(\mathbf{A}\otimes\mathbf{C} + \mathbf{C}\otimes\mathbf{A})\boldsymbol{\mu} \\ &+ (\mu_e + \sigma_e^2)\mathrm{vec}\{\mathbf{C}\mathbf{C}^T\}]\end{aligned} \tag{10}$$

Moreover,

$$\boldsymbol{\mu}(s) = E[\mathbf{X}_t\mathbf{X}_{t+s}^T] = \boldsymbol{\mu}(s-1)\mathbf{A}^T + \mu_e\boldsymbol{\mu}\mathbf{C}^T, \quad s \geq 1 \tag{11}$$

The expressions for the second-order raw moments show that the process is second-order stationary if $\rho(\mathbf{A}\otimes\mathbf{A}) < 1$, which, in view of the form of $\mathbf{A}$, is equivalent to $\rho(\mathbf{A}) < 1$.

Now we obtain the difference equations for the third-order lagged moments.

$$\boldsymbol{\mu}(0,s) = E[X_t\mathbf{X}_t\mathbf{X}_{t+s}^T] = \boldsymbol{\mu}(0,s-1)A^T + \mu_e\boldsymbol{\mu}(0)\mathbf{C}\mathbf{C}^T$$

$$\begin{aligned}\boldsymbol{\mu}(s,s) &= E[X_t\mathbf{X}_{t+s}\mathbf{X}_{t+s}^T] \\ &= \mathbf{A}\boldsymbol{\mu}(s-1,s-1)\mathbf{A}^T + \mu_e\mathbf{C}\mathbf{C}^T\boldsymbol{\mu}(s-1)\mathbf{A}^T + \mu_e\mathbf{A}\boldsymbol{\mu}(s-1)\mathbf{C}\mathbf{C}^T \\ &\quad + \mathbf{C}_2^{(s-1)} + \mu(\mu_e^2 + \sigma_e^2)\mathbf{C}\mathbf{C}^T, \quad s \geq 1\end{aligned} \tag{12}$$

$$\begin{aligned}\boldsymbol{\mu}(s,\tau) &= E[X_t\mathbf{X}_{t+s}\mathbf{X}_{t+\tau}^T] \\ &= \boldsymbol{\mu}(s,\tau-1)\mathbf{A}^T + \mu_e\boldsymbol{\mu}^T(s)\mathbf{C}\mathbf{C}^T, \quad s \geq 1,\ \tau \geq s+1\end{aligned} \tag{13}$$

The matrix $\mathbf{C}_2^{(s)}$ is a diagonal matrix with elements

$$\mathbf{C}_2^{(s)}(j,j) = \sum_{i=1}^{p}\sigma_i^2\mu(s-j), \quad j = 1,\ldots,p,$$

resulting from the properties of the vectorial thinning operation.

In scalar form the above difference equations are written as follows:

$$\mu(s) = \sum_{j=1}^{p}\alpha_j\mu(s-j) + \mu_e\mu \tag{14}$$

$$\mu(0,s) = \sum_{j=1}^{p}\alpha_j\mu(0,s-j) + \mu_e\mu(0) \tag{15}$$

$$\mu(s,s) = \sum_{j=1}^{p}\sum_{k=1}^{p}\alpha_j\alpha_k\mu(s-j,s-k) + \sum_{j=1}^{p}\sigma_i^2\mu(s-j) + 2\mu_e\sum_{j=1}^{p}\alpha_j\mu(s-j) + \mu(\mu_e^2+\sigma_e^2) \tag{16}$$

$$\mu(s,\tau) = \sum_{j=1}^{p}\alpha_j\mu(s,\tau-j) + \mu_e\mu(s). \tag{17}$$

2.2. *Difference equations in terms of the cumulants*

Let $\{X_t\}$ be a stationary time series for which moments up to the order k exist and let $C_k(t_1, t_2,\ldots,t_k)$ denote the kth-order joint cumulant of $X_{t_1},\ldots,X_{t_k}$, which can be written as $\mathrm{cum}(X_{t_1},\ldots,X_{t_k})$. However, $C_k(t_1, t_2,\ldots,t_k)$ is a function of $k-1$ variables which we, henceforth shall denote by $C_k(s_1,\ldots,s_{k-1})$. Leonov and Shiryaev (1959) derived relations between joint moments and joint cumulants of a stationary time series. In this section we give, without proof, explicit relations between moments and cumulants that will be needed later for the derivation of the difference equations. It is easily shown that

$$\mathrm{cum}(X_t) = \mathrm{E}(X_t) = \mu \tag{18}$$

$$C_2(s) = \mu(s) - \mu^2 = R(s) \tag{19}$$

$$C_3(s_1,s_2) = \mu(s_1,s_2) - \{\mu(s_1)+\mu(s_2)+\mu(s_2-s_1)\}\mu + 2\mu^3 = R(s_1,s_2) \tag{20}$$

From equations (19) and (20) we see that the second- and third-order cumulants are the same as the second- and third-order moments about the mean. If the time series $\{X_t\}$ is a real-valued stationary time series, the third-order cumulants satisfy the following set of symmetry relations over and above the well-known conditions $C_2(s) = C_2(-s)$. For $s \geq 0$ and $\tau \geq 0$ we have,

$$\begin{aligned} C_3(s,s+\tau) &= C_3(s+\tau,s) = C_3(-s,\tau) = C_3(\tau,-s) \\ &= C_3(-\tau,-\tau-s) = C_3(-s-\tau,-\tau) \end{aligned} \tag{21}$$

In view of the above symmetry relations, $C_3(s_1, s_2)$ is completely defined over the entire plane by its values for $s_1 \geq 0$ and $0 \leq s_2 \leq s_1$.

Using the above relations between moments and cumulants, now we obtain difference equations for the second- and third-order cumulants. Since the second order cumulants are just the autocovariance function we obtain easily

$$\mathbf{R}(0) = E[\mathbf{X}_t\mathbf{X}_t^T] - E(\mathbf{X}_t)E(\mathbf{X}_t^T) = \mathbf{A}R(0)\mathbf{A}^T + \mathbf{C}_1 + \sigma_e^2\mathbf{C}\mathbf{C}^t \tag{22}$$

Applying the operator vec to (22) we obtain

$$\mathrm{vec}\{R(0)\} = (\mathbf{A}\otimes\mathbf{A})\mathrm{vec}\{R(0)\} + \mathrm{vec}\{\mathbf{C}_1 + \sigma_e^2\mathbf{C}\mathbf{C}^T\} \tag{23}$$

This is a first-order difference equation in vec$\{\mathbf{R}(0)\}$ which, under the condition

$$\rho(\mathbf{A}\otimes\mathbf{A}) < 1 \tag{24}$$

as a convergent solution of the form

$$\mathrm{vec}\{\mathbf{R}(0)\} = (\mathbf{I}_{p^2} - \mathbf{A}\otimes\mathbf{A})^{-1}[\mathrm{vec}\{\mathbf{C}_1\} + \mathrm{vec}\{\sigma_e^2\mathbf{C}\mathbf{C}^T\}] \tag{25}$$

Now we obtain $\mathbf{R}(s)$, $s \geq 1$ as follows.

$$\mathbf{R}(s) = E[\mathbf{X}_{t-k}\mathbf{X}_t] - E[\mathbf{X}_{t-k}]E[\mathbf{X}_t] = \mathbf{R}(s-1)\mathbf{A}^T = \mathbf{R}(0)[\mathbf{A}^T]^s \tag{26}$$

We can easily write the expression for $R(s)$ in scalar form as follows:

$$R(s) = \sum_{i=1}^{p} \alpha_i R(s-i) \tag{27}$$

This expression shows that the INAR(p) process has the same second-order correlation structure as the AR(p) process. For a detailed comparison between the INAR and AR processes, see section 4 of Latour (1997) and section 5 of Dion *et al.* (1995).

The third-order cumulants are obtained using the relations between the moments and cumulants, given by equations (18) to (20). Thus we have the following expressions:

$$C_3(0,s) = \sum_{i=1}^{p} \alpha_i \mathrm{C}_3(0,s-i), \quad s \geq 1 \tag{28}$$

$$C_3(s,s) = \sum_{i=1}^{p}\sum_{j=1}^{p} \alpha_i\alpha_j \mathrm{C}_3(s-i,s-j) + \sum_{i=1}^{p} \sigma_i^2 \mathrm{C}_2(s-i), \quad s \geq 1 \tag{29}$$

$$\mathrm{C}_3(s,\tau) = \sum_{i=1}^{p} \alpha_i \mathrm{C}_3(s,\tau-i), \quad s \geq 1, \quad \tau \geq s+1 \tag{30}$$

It is interesting to note that the third-order cumulants satisfy a set of Yule–Walker-type difference equations, similar to those satisfied by the bilinear process and which reveal the nonlinear structure of the INAR process.

3. CUMULANTS OF RANDOM VECTORS

In this section we consider the joint cumulant functions of random vectors defined by Yow-Jen and Shaman, (1988). We begin by defining the joint cumulant of k random vectors. The definition implies a specific arrangement of scalar joint cumulants of the components of the vectors. Unlike the scalar joint cumulant, the joint cumulant of k random vectors is not invariant under permutations of its arguments, and we shall describe its symmetries.

Let $\mathbf{X}_i = (X_{i1}, X_{i2},\ldots,X_{in_i})$, $i = 1,\ldots,k$ be random vectors of size n_i. Then we define (see Yow-Jen and Shaman, 1988) $\mathrm{cum}(\mathbf{X}_1, \mathbf{X}_2,\ldots,\mathbf{X}_k)$ as the $n_1\ldots n_k \times 1$ vector the elements of which are the cumulants $\mathrm{cum}(X_{1j_1}, X_{2j_2},\ldots,X_{kj_k})$ ordered so that the indexes $j_1,\ldots,j_k$ appear in lexicographic order.

According to this definition, if $\mathbf{X} = (X_1,\ldots,X_p)^T$, $\mathbf{Y} = (Y_1,\ldots,Y_p)^T$, $\mathbf{Z} = (Z_1,\ldots,Z_p)^T$, are $p \times 1$ vectors we may write

$$\mathrm{cum}(\mathbf{X}) = \mathrm{cum}\left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}\right) = \begin{bmatrix} \mathrm{cum}(X_1) \\ \mathrm{cum}(X_2) \\ \vdots \\ \mathrm{cum}(X_p) \end{bmatrix}_{(p\times 1)} \tag{31}$$

$$\mathrm{cum}(\mathbf{X}, \mathbf{Y}) = \mathrm{cum}\left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}, \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix}\right) = \begin{bmatrix} \mathrm{cum}(X_1, Y_1) \\ \mathrm{cum}(X_1, Y_2) \\ \vdots \\ \mathrm{cum}(X_1, Y_p) \\ \vdots \\ \mathrm{cum}(X_p, Y_p) \end{bmatrix}_{(p^2\times 1)} \tag{32}$$

$$\mathrm{cum}(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) = \begin{bmatrix} \mathrm{cum}(X_1, Y_1, Z_1) \\ \mathrm{cum}(X_1, Y_1, Z_2) \\ \vdots \\ \mathrm{cum}(X_1, Y_1, Z_p) \\ \mathrm{cum}(X_1, Y_2, Z_1) \\ \vdots \\ \mathrm{cum}(X_p, Y_p, Z_p) \end{bmatrix}_{(p^3\times 1)} \tag{33}$$

Yow-Jen and Shaman (1988) prove that this definition is consistent with and verifies the following property:

LEMMA 1. *Cumulants of linear combinations of random vectors are linear combinations of cumulants of random vectors, that is*

$$\operatorname{cum}\left(\sum_{i_1=1}^{m_1}\mathbf{A}_{1i_1}\mathbf{X}_{i_1},\ldots,\sum_{i_k=1}^{m_k}\mathbf{A}_{ki_k}\mathbf{X}_{i_k}\right)=\sum_{i_1}\ldots\sum_{i_k}(\mathbf{A}_{1i_1}\otimes\cdots\otimes\mathbf{A}_{ki_k})\operatorname{cum}(\mathbf{X}_{i_1},\ldots,\mathbf{X}_{i_k}) \tag{34}$$

Similarly we can prove Lemma 2 (Silva and Oliveira, 2000a,b).

LEMMA 2. *Let* $\mathbf{X}$, $\mathbf{Y}$ *and* $\mathbf{Z}$ *be* $p\times 1$ *random vectors and* $\mathbf{A}$ *a* $p\times p$ *matrix with elements* $0\le a_{ij}\le 1$, $i, j = 1,\ldots,p$. *The counting series of* $\mathbf{A}*\mathbf{Z}$ *have means* μ_{ij} *and variances* σ_{ij}^2 *and if they are all independent then*

$$\operatorname{cum}(\mathbf{A}*\mathbf{X})=\mathbf{A}\ \operatorname{cum}(\mathbf{X}) \tag{35}$$

$$\operatorname{cum}(\mathbf{A}*\mathbf{X},\mathbf{A}*\mathbf{X})=(\mathbf{A}\otimes\mathbf{A})\operatorname{cum}(\mathbf{X},\mathbf{X})+\operatorname{vec}\{\mathbf{C}_1\} \tag{36}$$

$$\operatorname{cum}(\mathbf{X},\mathbf{A}*\mathbf{X},\mathbf{A}*\mathbf{X})=(\mathbf{I}_p\otimes\mathbf{A}\otimes\mathbf{A})\operatorname{cum}(\mathbf{X},\mathbf{X},\mathbf{X})+\mathbf{D}_1 \tag{37}$$

where the $p\times p$ *matrix* $\mathbf{C}_1$ *is a diagonal matrix with*

$$\mathbf{C}_1(j,j)=\sum_{i=1}^{p}\sigma_{ji}^2E(X_i)$$

and the matrix $\mathbf{D}_1$ *is a* $p^3\times 1$ *matrix the ith element of which is given by*

$$(\mathbf{D}_1)_i=\operatorname{vec}\left\{\mathbf{C}_2^{(i)}-E(X_i)\mathbf{C}_1\right\} \tag{38}$$

and $\mathbf{C}2^{(\mathrm{i})}$ *is a diagonal* $p\times p$ *matrix with*

$$\mathbf{C}_2^{(i)}(j,j)=\sum_{k=1}^{p}\sigma_{jk}^2\mathrm{E}(X_iX_k).$$

Moreover, if all the counting series of $\mathbf{A}*\mathbf{Y}$ *and* $\mathbf{A}*\mathbf{Z}$ *are independent and independent of* $\mathbf{X}$ *we have*

$$\operatorname{cum}(\mathbf{X},\mathbf{A}*\mathbf{Z})=(\mathbf{I}_p\otimes\mathbf{A})\operatorname{cum}(\mathbf{X},\mathbf{Z}) \tag{39}$$

$$\operatorname{cum}(\mathbf{X},\mathbf{X},\mathbf{A}*\mathbf{Y})=(\mathbf{I}_p\otimes\mathbf{I}_p\otimes\mathbf{A})\operatorname{cum}(\mathbf{X},\mathbf{X},\mathbf{Y}) \tag{40}$$

$$\operatorname{cum}(\mathbf{X},\mathbf{Y},\mathbf{A}*\mathbf{Z})=(\mathbf{I}_p\otimes\mathbf{I}_p\otimes\mathbf{A})\operatorname{cum}(\mathbf{X},\mathbf{Y},\mathbf{Z}) \tag{41}$$

$$\mathrm{cum}(\mathbf{X}, \mathbf{A} * \mathbf{Y}, \mathbf{A} * \mathbf{Y}) = (\mathbf{I}_p \otimes \mathbf{A} \otimes \mathbf{A})\mathrm{cum}(\mathbf{X}, \mathbf{Y}, \mathbf{Y}) + \mathbf{D}_2 \tag{42}$$

where the matrix $\mathbf{D}_2$ *is defined similarly to* $\mathbf{D}_1$, *substituting* X_i *by* Y_i *in* $\mathbf{C}_1$ *and* X_k *by* Y_k *in* $\mathbf{C}2^{(\mathrm{i})}(\mathrm{j},\mathrm{j})$.

Now, let $\mathbf{A}$ *be a* $p \times p$ *matrix as in* (4). *Then, the counting series of* $\mathbf{A} * \mathbf{X}$ *have means* α_j, *variances* σ_j^2 *and third-order raw moments* γ_j, $j = 1,\ldots,p$, *and we may write*

$$\mathrm{cum}(\mathbf{A} * \mathbf{X}, \mathbf{A} * \mathbf{X}, \mathbf{A} * \mathbf{X}) = (\mathbf{A} \otimes \mathbf{A} \otimes \mathbf{A})\mathrm{cum}(\mathbf{X}, \mathbf{X}, \mathbf{X}) + \mathbf{D}_3 \tag{43}$$

where

$$\mathbf{D}_3 = \begin{bmatrix} \mathbf{C}_3 + \mathbf{C}_4 + \mathbf{C}_5 \\ \mathbf{C}_2^{(1)} \\ \vdots \\ \mathbf{C}_2^{(p-1)} \end{bmatrix}_{(p^3 \times 1)} \tag{44}$$

with

$$\mathbf{C}_3 = \begin{bmatrix} \sum_{i=1}^{p} \alpha_i \sigma_i^2 [E(X_i^2) - E(X_i)^2] \\ \sum_{i=1}^{p} \sigma_i^2 [E(X_i X_1) - E(X_i)E(X_1)] \\ \vdots \\ \sum_{i=1}^{p} \sigma_i^2 [E(X_i X_{p-1}) - E(X_i)E(X_{p-1})] \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(p^2 \times 1)} \tag{45}$$

$\mathbf{C}_4$ *a* $p^2 \times 1$ *matrix such that*

$$\mathbf{C}_4(1,1) = \sum_{j=1}^{p} (\gamma_j - 3\alpha_j \sigma_j^2 - \alpha_j^3)\mathrm{E}(X_j), \quad \mathbf{C}_4(i,j) = 0,$$

for $(i, j) \neq (1, 1)$, $\mathbf{C}_5$ *a* $p^2 \times 1$ *matrix such that*

$$\mathbf{C}_5(1,1) = \sum_{i=1}^{p} \sum_{j \neq i} \sigma_i^2 \alpha_j [E(X_i X_j) - 3E(X_i)E(X_j)], \quad \mathbf{C}_5(i,j) = 0,$$

for $(i, j) \neq (1, 1)$, *and where, now,* $\mathbf{C}_2^{(\mathrm{i})}$ *is the* $p^2 \times 1$ *matrix with elements*

$$\mathbf{C}_2^{(i)}(1,1) = \sum_{k=1}^{p} \sigma_k^2 [E(X_i X_k) - E(X_i)E(X_k)] \quad \text{and} \quad \mathbf{C}_2^{(i)}(i,j) = 0$$

for $(i, j) \neq (1, 1)$.

We shall denote the joint cumulants by $\boldsymbol{\mu} = \mathrm{cum}(\mathbf{X}_t)$, a $p \times 1$ vector, $\mathbf{k}_2(s) = \mathrm{cum}(\mathbf{X}_t, \mathbf{X}_{t+s})$, a $p^2 \times 1$ vector and $\mathbf{k}_3(s, \tau) = \mathrm{cum}(\mathbf{X}_t, \mathbf{X}_{t+s}, \mathbf{X}_{t+\tau})$, a $p^3 \times 1$ vector. The symmetry relations satisfied by the vectorial cumulants defined above

are different from the symmetry relations satisfied by the cumulants in the univariate case, since now the arguments are vectors and therefore the order does matter. However, the cumulants defined in this section do satisfy a set of symmetry relations but we now require the notation of permutation matrices to describe them. We follow Yow-Jen and Shaman closely.

A permutation is a one-to-one transformation of a finite set into itself. Permutations which give a circular rearrangement of symbols permuted are called cyclic permutation or cycles. Let $\mathbf{I}_{(p,q)}$ be the $pq \times pq$ matrix partitioned into $p \times q$ submatrices such that the ijth submatrix has a 1 in position (j, i) and 0 elsewhere. The matrix $\mathbf{I}_{(p,q)}$ is called *the permuted identity matrix* by MacRae (1974) and is used to change the order of a Kronecker product. If A is an $m \times n$ matrix and B is a $p \times q$ matrix then

$$\mathbf{B} \otimes \mathbf{A} = \mathbf{I}_{(m,p)}(\mathbf{A} \otimes \mathbf{B})\mathbf{I}_{(q,n)} \tag{46}$$

In particular, if $n = q = 1$, i.e. $\mathbf{A}$ and $\mathbf{B}$ are column vectors,

$$\mathbf{B} \otimes \mathbf{A} = \mathbf{I}_{(m,p)}(\mathbf{A} \otimes \mathbf{B}) \tag{47}$$

Let $\mathbf{Y}$ be a $p \times 1$ vector. We now construct the permutation matrices using the permuted identity matrix $\mathbf{I}_{(p,p)}$. Let $\mathbf{J}_{11} = \mathbf{I}_p$, $\mathbf{J}_{12} = \mathbf{I}_{p^2}$, $\mathbf{J}_{22} = \mathbf{I}_{(p,p)}$, $\mathbf{J}_{13} = \mathbf{I}_{p^3}$, $\mathbf{J}_{23} = \mathbf{I}_{(p,p)} \otimes \mathbf{I}_p$ and $\mathbf{J}_{33} = \mathbf{I}_p \otimes \mathbf{I}_{(p,p)}$. Thus, for $k = 1, 2, 3$, $\mathbf{J}_{ik}$, $i = 1,\ldots,k$ is a matrix of dimensions $p^k \times p^k$ which interchanges the positions of the $(i-1)$th and ith arguments of cumulant cum($\mathbf{Y}_1,\ldots,\mathbf{Y}_k$) when $\mathbf{J}_{ik}$ cum($\mathbf{Y}_1,\ldots,\mathbf{Y}_k$) is formed. Then, the cumulants satisfy the following set of symmetries.

$$\mathbf{k}_2(-g) = \mathbf{I}_{(p,p)}\mathbf{k}_2(g) = \mathbf{J}_{22}\mathbf{k}_2(g) \tag{48}$$

$$\mathbf{k}_3(h, g) = \mathbf{J}_{33}\mathbf{k}_3(g, h) \tag{49}$$

$$\mathbf{k}_3(-g, h-g) = \mathbf{J}_{23}\mathbf{k}_3(g, h) \tag{50}$$

$$\mathbf{k}_3(-h, g-h) = \mathbf{J}_{23}\mathbf{J}_{33}\mathbf{k}_3(g, h) \tag{51}$$

$$\mathbf{k}_3(h-g, -g) = \mathbf{J}_{33}\mathbf{J}_{23}\mathbf{k}_3(g, h) \tag{52}$$

$$\mathbf{k}_3(g-h, -h) = \mathbf{J}_{23}\mathbf{J}_{33}\mathbf{J}_{23}\mathbf{k}_3(g, h) \tag{53}$$

In view of the above expressions it is sufficient to calculate $\mathbf{k}_2(g)$ for $g \geq 0$ and $\mathbf{k}_3(g, h)$ for $0 \leq g \leq h < \infty$.

Under matrix multiplication, the matrices $\mathbf{J}_{1k},\ldots,\mathbf{J}_{kk}$ generate a group of order $k!$ which is isomorphic to the symmetric group on k objects. We note that $\mathbf{J}_{23}\mathbf{J}_{33}\mathbf{J}_{23} = \mathbf{J}_{33}\mathbf{J}_{23}\mathbf{J}_{33}$, $\mathbf{J}_{ik}\mathbf{J}_{ik} = \mathbf{I}$, $\mathbf{J}_{23}\mathbf{J}_{33} = \mathbf{I}_{(p^2,\, p)}$ and $\mathbf{J}_{33}\mathbf{J}_{23} = \mathbf{I}_{(p,p^2)}$.

To derive cumulant functions of model INAR(p) we shall substitute using (3) and then apply the results above. We shall use the independence of e_t and X_s, $s < t$, and also the assumptions of (2). We find the cumulant functions up to third order and corresponding sufficient conditions for stationarity. We shall use repeatedly the fact that

$$\mathrm{cum}(\mathbf{X}_t, \mathbf{C}e_t) = \mathrm{cum}(\mathbf{A} * \mathbf{X}_{t-1} + \mathbf{C}e_t, \mathbf{C}e_t) = (\mathbf{C} \otimes \mathbf{C})\sigma_e^2$$

which follows from (3), results above and the given assumptions.

Now, we state the following.

THEOREM 1. *Let $\{X_t\}$ be a stationary process satisfying* (3). *Then*

$$\boldsymbol{\mu} = (\mathbf{I}_p - \mathbf{A})^{-1}\mathbf{C}\mu_e \tag{54}$$

$$\mathbf{k}_2(0) = [\mathbf{I}_{p^2} - (\mathbf{A} \otimes \mathbf{A})]^{-1}[\mathrm{vec}\{\mathbf{C}_1\} + (\mathbf{C} \otimes \mathbf{C})\sigma_e^2] \tag{55}$$

$$\mathbf{k}_2(s) = (\mathbf{I}_p \otimes \mathbf{A})\mathbf{k}_2(s-1) = (\mathbf{I}_p \otimes \mathbf{A}^s)\mathbf{k}_2(0) \tag{56}$$

$$\mathbf{k}_3(0,0) = [\mathbf{I}_{p^3} - (\mathbf{A} \otimes \mathbf{A} \otimes \mathbf{A})]^{-1}[\mathbf{F}_3 + (\mathbf{C} \otimes \mathbf{C} \otimes \mathbf{C})\gamma_e] \tag{57}$$

$$\mathbf{k}_3(0,s) = (\mathbf{I}_p \otimes \mathbf{I}_p \otimes \mathbf{A})\mathbf{k}_3(0,s-1) = (\mathbf{I}_p \otimes \mathbf{I}_p \otimes \mathbf{A}^s)\mathbf{k}_3(0,0) \tag{58}$$

$$\begin{aligned}\mathbf{k}_3(s,s) &= (\mathbf{I}_p \otimes \mathbf{A} \otimes \mathbf{A})\mathbf{k}_3(s-1,s-1) + \mathbf{F}_2^{(s-1)} \\ &= (\mathbf{I}_p \otimes \mathbf{A} \otimes \mathbf{A})^s\mathbf{k}_3(0,0) \\ &\quad + \sum_{j=0}^{s-1}(\mathbf{I}_p \otimes \mathbf{A} \otimes \mathbf{A})^j(\mathbf{I}_p \otimes \mathbf{H} \otimes \mathbf{F})(\mathbf{I}_p \otimes \mathbf{I}_p \otimes \mathbf{A})^{(s-1-j)}\mathbf{F}_2^{(0)}\end{aligned} \tag{59}$$

$$\begin{aligned}\mathbf{k}_3(s,\tau) &= (\mathbf{I}_p \otimes \mathbf{I}_p \otimes \mathbf{A})\mathbf{k}_3(s,\tau-1) \\ &= (\mathbf{I}_p \otimes \mathbf{A}^s \otimes \mathbf{A}^\tau)\mathbf{k}_3(0,0) + (\mathbf{I}_p \otimes \mathbf{I}_p \otimes \mathbf{A}^{\tau-s}) \\ &\quad \times \sum_{j=0}^{s-1}(\mathbf{I}_p \otimes \mathbf{A} \otimes \mathbf{A})^j(\mathbf{I}_p \otimes \mathbf{H} \otimes \mathbf{F})(\mathbf{I}_p \otimes \mathbf{I}_p \otimes \mathbf{A})^{(s-1-j)}\mathbf{F}_2^{(0)}\end{aligned} \tag{60}$$

where $\mathbf{C}_1$ has been defined before, $\mathbf{H}$ is such that $\mathbf{H}(1,1) = 1$ and $\mathbf{H}(i,j) = 0$, $i, j = 2,\ldots,p$, $\mathbf{F}$ is the $p \times p$ matrix such that $\mathbf{F}(1,\mathrm{j}) = \sigma_{\mathrm{j}}^2$ and $\mathbf{F}(i,j) = 0$, $i = 2,\ldots,p$, $j = 1,\ldots,p$,

$$\mathbf{F}_2^{(0)} = \begin{bmatrix} \mathbf{k}_2(0) \\ (\mathbf{I}_p \otimes \mathbf{A})^{-1}\mathbf{k}_2(0) \\ \vdots \\ (\mathbf{I}_p \otimes \mathbf{A})^{-p+1}\mathbf{k}_2(0) \end{bmatrix}_{(p^3\times 1)} \tag{61}$$

$$\mathbf{F}_3 = \begin{bmatrix} \mathbf{G}_3 + \mathbf{G}_4 + \mathbf{G}_5 \\ \mathbf{G}_2^{(1)} \\ \vdots \\ \mathbf{G}_2^{(p-1)} \end{bmatrix}_{(p^3\times 1)} \tag{62}$$

with

$$\mathbf{G}_3 = \begin{bmatrix} \sum_{i=1}^{p} \alpha_i \sigma_i^2 [E(X_{t-i}^2) - E(X_{t-i})^2] \\ \sum_{i=1}^{p} \sigma_i^2 [E(X_{t-i}X_{t-1}) - E(X_{t-i})E(X_{t-1})] \\ \vdots \\ \sum_{i=1}^{p} \sigma_i^2 [E(X_{t-i}X_{t-(p-1)}) - E(X_{t-i})E(X_{t-(p-1)})] \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(p^2\times 1)} \tag{63}$$

$\mathbf{G}_4$ *a* $p^2 \times 1$ *matrix such that*

$$\mathbf{G}_4(1,1) = \sum_{j=1}^{p} (\gamma_j - 3\alpha_j\sigma_j^2 - \alpha_j^3)E(X_{t-j}), \quad \mathbf{G}_4(i,j) = 0,$$

for $(i, j) \neq (1, 1)$, $\mathbf{G}_5$ *a* $p^2 \times 1$ *matrix such that*

$$\mathbf{G}_5(1,1) = \sum_{i=1}^{p} \sum_{j\neq i} \sigma_i^2 \alpha_j [E(X_{t-i}X_{t-j}) - 3E(X_{t-i})E(X_{t-j})], \quad \mathbf{G}_5(i,j) = 0,$$

for $(i, j) \neq (1, 1)$ *and where, now,* $\mathbf{G}_2^{(\mathrm{i})}$ *is the* $p^2 \times 1$ *matrix with elements*

$$\mathbf{G}_2^{(i)}(1,1) = \sum_{k=1}^{p} \sigma_k^2 [E(X_{t-i}X_{t-k}) - E(X_{t-i})E(X_{t-k})]$$

and $\mathbf{C}_2^{(\mathrm{i})}(\mathrm{i},\mathrm{j}) = 0$ *for* $(i, j) \neq (1, 1)$.

For the above expressions to hold we must have that $\rho(\mathbf{A}) < 1$ for the first-order stationarity, $\rho(\mathbf{A}\otimes\mathbf{A}) < 1$ for the second-order stationarity and $\rho(\mathbf{A}\otimes\mathbf{A}\otimes\mathbf{A}) < 1$ for the third-order stationarity of a INAR(p) process. However, in view of the matrix $\mathbf{A}$ the three conditions are equivalent. Thus an INAR(p) process is stationary up to third order if $\rho(\mathbf{A}) < 1$.

4. THE SPECTRAL AND BISPECTRAL DENSITY FUNCTIONS

Let $\mathbf{f_X}(\omega)$ denote the spectral density function of $\{\mathbf{X}_t\}$ satisfying (3). Then

$$\begin{aligned}\mathbf{f_X}(\omega) &= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\mathbf{R}(k)\mathrm{e}^{-\mathrm{i}\omega k}\\ &= \frac{1}{2\pi}\left\{\mathbf{R}(0)+\sum_{k=-\infty}^{-1}\mathbf{R}(k)\mathrm{e}^{-\mathrm{i}\omega k}+\sum_{k=1}^{\infty}\mathbf{R}(k)\mathrm{e}^{-\mathrm{i}\omega k}\right\}\end{aligned}$$

But since $\mathbf{R}(-s)=\mathbf{R}^T(s)$ for real-valued $\mathbf{X}_t$ and $\mathbf{R}(s)=\mathbf{R}(0)(\mathbf{A}^T)^s$, $s\geq 1$ we have that

$$\mathbf{f_X}(\omega)=\frac{1}{2\pi}\{\mathbf{R}(0)+\mathbf{M}^*\mathbf{R}(0)+\mathbf{R}(0)\mathbf{M}\} \tag{64}$$

where

$$\mathbf{M}=\mathbf{A}^T\left\{\mathbf{I}_p-\mathbf{A}^T\mathrm{e}^{-\mathrm{i}\omega}\right\}^{-1}\mathrm{e}^{-\mathrm{i}\omega} \tag{65}$$

and $\mathbf{M}^*$ is the complex conjugate transpose of $\mathbf{M}$.

In order to express the spectral density matrix in terms of the parameters of the INAR(p) model we use the matrix algebra results on Kronecker products (Graham, 1981; Searle, 1982). Then

$$\mathrm{vec}\{\mathbf{f_X}(\omega)\}=\frac{1}{2\pi}\{\mathrm{vec}\{\mathbf{R}(0)\}+\mathrm{vec}\{\mathbf{M}^*\mathbf{R}(0)\}+\mathrm{vec}\{\mathbf{R}(0)\mathbf{M}\}\} \tag{66}$$

where

$$\mathrm{vec}\{\mathbf{R}(0)\}=(\mathbf{I}_{p^2}-\mathbf{A}\otimes\mathbf{A})^{-1}\mathrm{vec}\{\mathbf{C}_1+\mathbf{C}\mathbf{C}^T\sigma_e^2\} \tag{67}$$

$$\mathrm{vec}\{\mathbf{M}^*\mathbf{R}(0)\}=(\mathbf{I}_p\otimes\mathbf{M}^*)\mathrm{vec}\{\mathbf{R}(0)\} \tag{68}$$

$$\mathrm{vec}\{\mathbf{R}(0)\mathbf{M}\}=\left(\mathbf{M}^T\otimes\mathbf{I}_p\right)\mathrm{vec}\{\mathbf{R}(0)\} \tag{69}$$

Here $\mathbf{I}_{p^2}$ is $p^2\times p^2$, $\mathbf{A}$, $\mathbf{I}_p$, $\mathbf{M}$, $\mathbf{C}$ are $p\times p$ matrices. Hence we may write

$$\begin{aligned}\mathrm{vec}\{\mathbf{f_X}(\omega)\}=\frac{1}{2\pi}&\left\{\mathbf{I}_{p^2}+\left(\mathbf{I}_p\otimes\mathbf{M}^*\right)\right.\\ &\left.+\left(\mathbf{M}^T\otimes\mathbf{I}_p\right)\right\}\left(\mathbf{I}_{p^2}-\mathbf{A}\otimes\mathbf{A}\right)^{-1}\mathrm{vec}\left\{\mathbf{C}_1+\mathbf{C}\mathbf{C}^T\sigma_e^2\right\}\end{aligned} \tag{70}$$

The spectral density function $f_X(\omega)$ of X_t is then given by the first element of $\mathrm{vec}\{\mathbf{f_X}(\omega)\}$:

$$f_X(\omega)=\frac{\sigma_e^2+\mu\sum_{i=1}^{p}\sigma_i^2}{2\pi|\alpha(\mathrm{e}^{-\mathrm{i}\omega})|^2} \tag{71}$$

where $\alpha(z)=1-\alpha_1 z-\cdots-\alpha_p z^p$.

The bispectral density function of a stationary process $\{X_t\}$, $f_X(\omega_1, \omega_2)$, is defined as the Fourier transform of the third-order cumulants of the process. Thus, using the cumulants, $\mathbf{k}_3(s, \tau)$ defined above, we can write the following.

$$\mathbf{f}_{\mathbf{X}}(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum_{s=-\infty}^{\infty} \sum_{\tau=-\infty}^{\infty} \mathbf{k}_3(s, \tau) \mathrm{e}^{-\mathrm{i}s\omega_1 - \mathrm{i}\tau\omega_2}, \quad -\pi \leq \omega_1, \omega_2 \leq \pi \tag{72}$$

Then we can say that the bispectrum, $f_X(\omega_1, \omega_2)$ of X_t is the first element of $\mathbf{f}_{\mathbf{X}}(\omega_1, \omega_2)$. Considering the symmetry relations satisfied by the cumulants, using (58), (59), (60) and after application of algebra (Silva and Oliveira, 2000a,b) can state the following result.

THEOREM 2. *Let* $\{\mathbf{X}_t\}$ *be a stationary process satisfying* (3). *Then, the bispectrum,* $\mathbf{F}_X(\omega_1, \omega_2)$ *of* X_t *is the first element of* $\mathbf{f}_{\mathbf{X}}(\omega_1, \omega_2)$, *given by*

$$\begin{aligned}
\mathbf{f}_X(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \{ & \mathbf{k}_3 + [\mathbf{J}_{33}\mathbf{g}(-\omega_1) + \mathbf{g}(-\omega_2) + \mathbf{J}_{23}\mathbf{J}_{33}\mathbf{g}(\omega_1 + \omega_2)]\mathbf{k}_3 \\
& + [\mathbf{h}(-\omega_1 - \omega_2) + \mathbf{J}_{23}\mathbf{h}(\omega_1) + \mathbf{J}_{33}\mathbf{J}_{23}\mathbf{h}(\omega_2)]\mathbf{k}_3 \\
& + [\mathbf{S}(-\omega_1 - \omega_2) + \mathbf{J}_{23}\mathbf{S}(\omega_1) + \mathbf{J}_{33}\mathbf{J}_{23}\mathbf{S}(\omega_2)]\mathbf{F}_2^{(0)} \\
& + [\mathbf{g}(-\omega_2) + \mathbf{J}_{33}\mathbf{g}(-\omega_1)][\mathbf{h}(-\omega_1 - \omega_2)\mathbf{k}_3 + \mathbf{S}(-\omega_1 - \omega_2)\mathbf{F}_2^{(0)}] \\
& + [\mathbf{J}_{23}\mathbf{g}(-\omega_2) + \mathbf{J}_{23}\mathbf{J}_{33}\mathbf{g}(\omega_1 + \omega_2)][\mathbf{h}(\omega_1)\mathbf{k}_3 + \mathbf{S}(\omega_1)\mathbf{F}_2^{(0)}] \\
& + [\mathbf{J}_{33}\mathbf{J}_{23}\mathbf{g}(-\omega_1) + \mathbf{J}_{23}\mathbf{J}_{33}\mathbf{J}_{23}\mathbf{g}(\omega_1 + \omega_2)][\mathbf{h}(\omega_2)\mathbf{k}_3 + \mathbf{S}(\omega_2)\mathbf{F}_2^{(0)}]\}
\end{aligned}$$

where $\mathbf{k}_3 = \mathbf{k}_3(0, 0)$ *given by* (57), $\mathbf{g}(\omega_j) = \mathbf{W}z_j[\mathbf{I}_{p^3} - \mathbf{W}z_j]^{-1}$, $\mathbf{h}(\omega_j) = \mathbf{U}z_j[\mathbf{I}_{p^3} - \mathbf{U}z_j]^{-1}$ *with* $z_j = e^{\mathrm{i}\omega_j}$, $j = 1, 2$,

$$\mathbf{U} = \mathbf{I}_p \otimes \mathbf{A} \otimes \mathbf{A}, \quad \mathbf{V} = \mathbf{I}_p \otimes \mathbf{H} \otimes \mathbf{F}, \quad \mathbf{W} = \mathbf{I}_p \otimes \mathbf{I}_p \otimes \mathbf{A}$$

and $\mathbf{S}(\omega)$ *a* $p^3 \times p^3$ *matrix such that*

$$\begin{aligned}
\mathrm{vec}\{\mathbf{S}(\omega_j)\} = \{ & e^{i\omega_j}[\mathbf{I}_{p^6} - (\mathbf{W}^T \otimes \mathbf{I}_{p^3})e^{i\omega_j}]^{-1} \\
& - ((\mathbf{W}^{-1})^T \otimes \mathbf{U})e^{i\omega_j}[\mathbf{I}_{p^6} - (\mathbf{I}_{p^3} \otimes \mathbf{U})e^{i\omega_j}]^{-1}\} \\
& [\mathbf{I}_{p^6} - ((\mathbf{W}^{-1})^T \otimes \mathbf{U})]^{-1}\mathrm{vec}\{\mathbf{V}\}.
\end{aligned}$$

The above expression for the bispectra shows that the INAR processes are nonlinear processes (Subba Rao and Gabr, 1984).

5. FREQUENCY DOMAIN ESTIMATION OF THE INAR(P) MODEL

Let $X_1,\ldots,X_n$ be n observations from the INAR(p) model given by (2), denote the parameters of the model by θ and their true values by $\theta_0 = (\alpha_1, \alpha_2,\ldots,\alpha_p, \lambda)$. We

now consider the problem of estimating the parameters of the model. There are two approaches for the estimation of the parameters of finite-parameter time series models: (i) the frequency-domain approach and (ii) the time-domain approach. The Whittle criterion comes under the former approach. This approach has been originally proposed by Whittle (1953) for Gaussian processes, further investigated by several authors (Walker, 1964; Hannan, 1973; Rice, 1979; Dzhaparidze and Yaglom, 1983) and since used in many situations (Fox and Taku, 1986; Sesay and Subba Rao, 1992; Subba Rao and Chandler, 1996; Silva and Oliveira, 2000a,b). The motivation is that the spectral density function of a model may be easy to obtain whereas an exact likelihood is not. In fact, for the INAR(p) processes we were able to obtain an explicit expression for the spectrum in terms of the parameters. However, that was not the case for the likelihood. In Section 5.1, we review the Whittle criterion.

5.1. *Whittle criterion*

The idea of representing the likelihood of a stochastic process via its spectral properties rather than directly was first suggested by Whittle (1953) for Gaussian models. If the time series is Gaussian, the maximization of the logarithm of the likelihood function of the sample is asymptotically equivalent to minimization of Whittle's criterion given by

$$\hat{L}_N(\boldsymbol{\theta}) = \frac{N}{4\pi}\int_{-\pi}^{\pi}\left\{\log f(\omega,\boldsymbol{\theta}) + \frac{I_N(\omega)}{f(\omega,\boldsymbol{\theta})}\right\}\mathrm{d}\omega \tag{73}$$

where $f(\omega, \theta)$ denotes the non-normalized spectral density function of the process and $I_N(\omega)$ is given by

$$I_N(\omega) = (2\pi N)^{-1}\left|\sum_{t=1}^{N} X_t \mathrm{e}^{\mathrm{i}\omega t}\right|^2.$$

Let $\hat{\boldsymbol{\theta}}$ represent the estimator obtained by minimizing (73), the so-called Whittle estimator. Under the Gaussianity assumption, Whittle (1953) and Walker (1964) have shown that the estimator obtained by minimizing (73) is weakly consistent and has an asymptotic normal distribution. If the series is not Gaussian, Chandler (1996) has shown that the use of Whittle's criterion is justified because this is in fact the criterion one would arrive at if one considers approximate likelihood functions of collections of sample Fourier coefficients. In this case, the Whittle's criterion has a quasi-likelihood interpretation. Results for the consistency and asymptotic normality of the Whittle estimators when the series is not Gaussian, were provided by Rice (1979). However, the asymptotic variance of $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ depends on the fourth-order spectra of the process, trispectrum. Thus the distributional results of the estimators are not very useful in practice since it is difficult to obtain expressions for the trispectrum.

5.2. *Application to the number of daily epileptic seizures of a patient*

The series of the number of daily epileptic seizures has been studied by Franke and Seligmann (1993) to assess the effectiveness of a certain treatment. These authors modelled the series using a SINAR(1) process and conditional maximum likelihood estimates. Later, Latour (1998) considered the first half of the time series of one of the patients, corresponding to the pre-treatment period. Here we consider the same set of data which is represented in Figure 1. The sample autocorrelation and partial autocorrelation functions are represented in Figure 2. These functions indicate as a possible model for the data an INAR(14) with nonzero coefficientes at lag 6 and lag 14 ($X_t = \alpha_6 * X_{t-6} + \alpha_{14} * X_{t-14} + e_t$, $e_t \sim Po(\lambda)$). To estimate the parameters using the minimization of Whittle criterion, (73) is replaced by

$$\hat{l}_N(\boldsymbol{\theta}) = \frac{1}{N}\sum_{j=1}^{[N/2]}\left\{\log f(\omega_j, \boldsymbol{\theta}) + \frac{I_N(\omega_j)}{f(\omega_j, \boldsymbol{\theta})}\right\} \tag{74}$$

Here, $f(\omega_j, \boldsymbol{\theta})$ is the spectral density function at the frequency point $\omega_j = 2\pi j/N$, $I_N(\omega_j)$ is the periodogram ordinate at the same frequency point and Σ is a summation over $j = \pm 1,\ldots,\pm [N/2]$, the point j=0 being excluded as a mean correction for X_t.

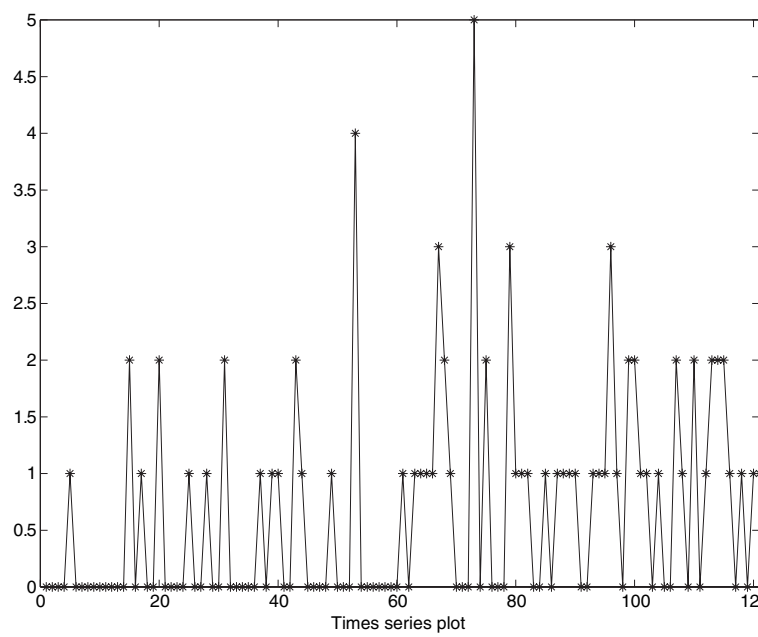
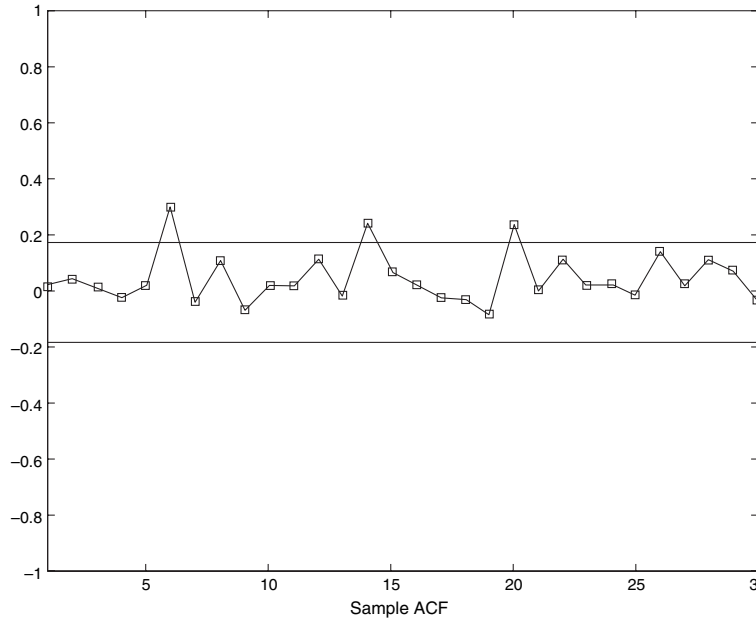


FIGURE 1. Number of daily epileptic seizures of a patient before treatment (Franke and Seligmann, 1993; Fig. 22.3).



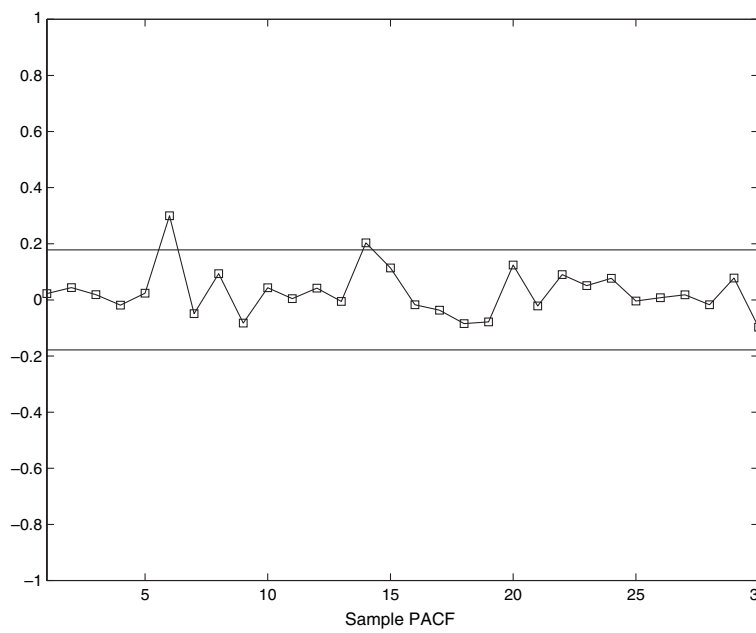


FIGURE 2. Sample autocorrelations and sample partial autocorrelations of the number of daily epileptic seizures.

The numerical minimization of equation (74) is achieved using numerical algorithms. Because of the nature of the model parameters, each α_i is a probability, we consider constrained minimization, imposing that $0 < \alpha_i < 1$, $i = 1,\ldots,p$. We use the minimization algorithm implemented in the Matlab function, '*constr*.' which implements a sequential quadratic programming method. The initial estimate needed to start the algorithm (see Optimization Toolbox User's Guide of MATLAB) is obtained by the method of moments from the Yule–Walker equations. Thus, the estimated parameters are: $\hat{\alpha}_6 = 0.2749$, $\hat{\alpha}_{14} = 0.1907$, $\hat{\lambda} = 0.4806$. The Ljung–Box statistic does not indicate correlated residuals. The estimates obtained here are similar to those obtained by Latour by the conditional least squares method (SAS ARIMA procedure).

ACKNOWLEDGEMENTS

The authors thank the anonymous referee for the comments that helped to improved the paper. This work has been partially funded by Centro de Matemática Aplicada do Porto, Praxis PCEX/Mat/47/96 and PRODEP – Medida 5.

NOTES

Corresponding author: Maria Eduarda Silva, CMAUP, Departamento de Matemática Aplicada, Faculdade de Ciências, Universidade do Porto, Rua das Taipas 135, 4050 600 Porto, Portugal. E-mail: mesilva@fc.up.pt

REFERENCES

Al-Osh, M. A. and Alzaid, A. A. (1987) First order integer-valued autoregressive (INAR(1)) processes. *J. Time Ser. Anal.* 8, 261–75.

Alzaid, A. A. and Al-Osh, M. A. (1990) An integer-valued pth order autoregressive structure (INAR(p)) process. *J. Appl. Prob.* 27, 314–24.

Chandler, R. (1996) A spectral method for estimating parameters in rainfall models. Research Report 142, Department of Statistical Science, UCL.

Dion, J-P., Gauthier, G. and Latour, A. (1995) Branching processes with immigration and integer valued time series. *Serdica Math.* 21, 123–36.

Dzhaparidze, K. O. and Yaglom, A. M. (1983) Spectrum parameter estimation in time series analysis. *Developments in Statistics*, Vol. 4, Ch. 1. Academic Press, New York.

Du, J. -G. and Li, Y. (1991) The integer-valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.* 12, 129–42.

Fox, R. and Taku, M. S. (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* 14, 517–32.

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* (ed. T. SUBBA RAO). London: Chapman and Hall.

FRANKE, J. and SUBBA RAO, T. (1995) Multivariate first order integer valued autoregressions. Technical Report, Maths Dep., UMIST.

GAUTHIER, G. and LATOUR, A. (1994) Convergence forte des estimateurs des paramètres d'un processus GENAR(p). *Ann. Sci. Math. Québec* 18, 37–59.

GRAHAM, A. (1981) *Kronecker Products and Matrix Calculus with Applications*. John Wiley and Sons, Chischester.

HANNAN, E. J. (1973) The asymptotic theory of linear time series models. *J. Appl. Prob.* 10, 130–45.

LEONOV, V. P. and SHIRYAEV, A. N. (1959) On a method of calculation of semi-invariants. *Theory Pro. Appl. (URSS)* 4, 319–29.

LATOUR, A. (1997) The multivariate GINAAR(p) process. *Adv. Appl. Prob.* 29, 228–47.

LATOUR, A. (1998) Existence and stochastic strucrure of a nonnegative integer-valued autoregressive process. *J. Time Ser. Anal.* 19, 439–55; 337–46.

MACRAE, E. C. (1974) Matrix derivates with an application to an adaptive linear decision problem. *Ann. Statist.* 2, 337–46.

RICE, J. (1979) On the estimation of the parameters of a power spectrum. *J. Mult. Anal.* 9, 378–92.

SEARLE, S. R. (1982) *Matrix Algebra Useful for Statistics*. John Wiley and Sons.

SESAY, S. A. O. and SUBBA RAO, T. (1992) Frequency domain estimation of bilinear time series models. *J. Time Ser. Anal.* 13, 521–45.

SILVA, M. E. and OLIVEIRA, V. L. (2000a) Difference equations for the higher order moments and cumulants of the INAR(1) model. Technical Report CMA/6/00, Universidade do Porto.

SILVA, M. E. and OLIVEIRA, V. L. (2000b) Difference equations for the higher order moments and cumulants of the INAR(p) model. Technical Report CMA/9/00, Universidade do Porto.

STEUTAL, F. W. and VAN HARN, K. (1979) Discrete analogues of self-decomposability and stability. *Ann. Prob.* 7, 893–9.

SUBBA RAO, T. and GABR, M. M. (1984) *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Lecture Notes in Statistics, Vol 24. Springer-Verlag, Berlin.

SUBBA RAO, T. and CHANDLER, R. E. (1996) A frequency domain approach for estimating parameters in point process models. In *Athens Conference on Applied Probability and Time Series* Vol II (eds P. ROBINSON and M. ROSENBLATT), Lecture Notes in Statistics No. 115, Springer-Verlag, pp. 392–405.

WALKER, A. M. (1964) Asymptotic properties of least squares estimates of parameters of the spectrum of a stationary non-deterministic series. *J. Austl. Math. Soc.* 4, 107–28.

WHITTLE, P. (1953) Estimation and information in stationary time series. *Ark. Mat.* 2, 423–34.

YOW-JEN, J. and SHAMAN, P. (1988) The third-order cumulant function and bispectral density of bilinear time series model BL(p, 0, p, 1). Technical Report, Department of Statistics, University of Pennsylvania, Philadelphia.