FACTA UNIVERSITATIS (NIŠ)
Ser. Math. Inform. Vol. 30, No 3 (2015), 263–280

RANDOM COEFFICIENT BIVARIATE INAR(1) PROCESS

Predrag M. Popović

Abstract. A bivariate autoregressive model for time series of counts is presented. The model is composed of survival and innovation components. The dependence between series is achieved through innovation parts. The autoregression is modelled with the survival component which is based on the binomial thinning operator. The coefficients that figure here are random variables. Statistical properties of the model are presented. The existence, stationarity and ergodicity of the model is proved. We focus on the model where the innovation components follow bivariate Poisson distribution. We suggest conditional maximum likelihood and the method of moments for parameters estimation. Both methods are tested on simulated data sets.
Keywords: Bivariate model; Integer valued autoregressive model; Thinning operator

1. Introduction

Time series of counts can be met in many fields of science. For example, in biology the number of some species represents such a series, or in finance the traded volume of some security, etc. Thus, modelling them is a challenging task that draws attention of many researchers. Many of the models are based on the thinning operator and have ARMA-like structure. The thinning operator is introduced to support discreetness of data since the standard Box-Jenkins ARMA models have many shortages with modelling integer valued times series. Different marginal distributions and different thinning operators are used to capture the specificity of observed data. The first integer valued autoregressive model (INAR) for univariate case was presented by [1]. This model is based on the binomial thinning operator with Poisson marginal distribution. Some discussion on univariate models can be found in [17], [7], [8].

When there are serial and cross correlations between some events, multivariate models are needed. These kinds of events occur in many fields. For example, [14] investigated dependence in trading two securities, [3] presented a model for the number of guest nights in hotels and cottages, [11] observed the number of patients

Received November 20, 2014; Accepted February 27, 2015
2010 *Mathematics Subject Classification.* Primary 62M10

infected by different diseases, [9] modelled the number of committed crimes, etc. The multivariate models were introduced in [4] and [6]. The latest contribution to developing multivariate INAR models can be found in [10], [12], [13], [5] and [16], while multivariate models with random coefficients were developed in [15] and [9]. All these models are composed of a survival and an innovation component. The survival component is the autoregressive part of these models while the innovation component is designed to support arrival of new members of the process. While in [15] and [9] survival components are dependent in [10] and [12], dependence is achieved through the innovation component.

The model presented in this paper (BVDINAR(1)) comprises random coefficients but with dependent innovation components. Also, the survival component of the presented model is truncated. For this reason, the model represents some sort of generalization of the model presented in [10] and an extension of the univariate models with truncated survival component discussed in [18] and [2]. BVDINAR(1) model is suitable for dependent processes where autocorrelation is not present from time to time. To motivate the model we consider real data examples where we observe two correlated series of light criminal activities. The series are monthly counts of activities categorized as larceny and criminal mischief. Besides the fact that these two series are autocorrelated and dependent, their frequency is often influenced by some external factors, such as the number of policemen in the area or time of the year. So, the absence of the autocorrelation from time to time might be expected. Lag one cross-correlation is not always expected with these series, but dependency between innovation processes is quite realistic since the same factors push people to commit these two criminal acts.

The paper is organized in the following way. Section 2. introduces the general form of the model, discusses stationarity and proves existence of the model. Section 3. gives us joint distribution, conditional and unconditional expectation and variance. A special case of the model where the innovation components follows binomial Poisson distribution is discussed in Section 4. Section 5. presents the application of the model to real data. In the end, some concluding remarks are given.

2. General specifications of the model

In this section we define BVDINAR(1) model for nonnegative integer valued time series $\{X_{1,t}, X_{2,t}\}_{t\in\mathbb{Z}}$. Focusing on the general form of the model we prove its existence. Also, we state some properties of the thinning operator which figures in the model.

For construction of BVDINAR(1) model we will not specify marginal distribution of these processes. The processes are composed of two components, survival and innovation, and we just assume that the innovation parts follow some bivariate distribution. The model is defined in the following way:

$$X_{1,t} = \begin{cases} \alpha_1 \circ X_{1,t-1} + \varepsilon_{1,t}, & \text{w.p. } p_1, \\ \varepsilon_{1,t}, & \text{w.p. } 1-p_1, \end{cases} \tag{2.1}$$

$$X_{2,t} = \begin{cases} \alpha_2 \circ X_{2,t-1} + \varepsilon_{2,t} & \text{w.p. } p_2, \\ \varepsilon_{2,t}, & \text{w.p. } 1-p_2, \end{cases} \tag{2.2}$$

where $\alpha_1, \alpha_2 \in (0,1)$ and $p_1, p_2 \in [0,1]$. The thinning operator that figures here is defined as $\alpha_1 \circ X_{1,t} = \sum_{i=1}^{X_{1,t}} B_i$ where $\{B_i\}$ is a sequence of independent identically distributed Bernoulli(α_1) random variables. Binomial thinnings $\alpha_1 \circ X_{1,t}$ and $\alpha_2 \circ X_{2,t}$ are mutually independent. The random variables $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are in general mutually dependent since they follow binomial distribution, but they are independent of $(X_{1,s}, X_{2,s})$ for $s < t$.

Similarly as in [2] and [18] we introduce random variables α_{1t} and α_{2t} such that $P(\alpha_{1t} = \alpha_1) = 1 - P(\alpha_{1t} = 0) = p_1$ and $P(\alpha_{2t} = \alpha_2) = P(\alpha_{2t} = 0) = p_2$. Equations (2.1) and (2.2) can be written as

$$X_{1,t} = \alpha_{1t} \circ X_{1,t-1} + \varepsilon_{1,t}$$
$$X_{2,t} = \alpha_{2t} \circ X_{2,t-1} + \varepsilon_{2,t}$$

The processes defined in this way are integer valued autoregressive processes of order one with random coefficients. Also, the above equation can be written in a matrix form by introducing matrix $\boldsymbol{A_t} = \begin{bmatrix} \alpha_{1t} & 0 \\ 0 & \alpha_{2t} \end{bmatrix}$ and vectors $\boldsymbol{Z}_t = (X_{1,t}, X_{2,t})'$ and $\boldsymbol{e}_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$. So we get the model as

$$\boldsymbol{Z}_t = \boldsymbol{A}_t \circ \boldsymbol{Z}_{t-1} + \boldsymbol{e}_t. \tag{2.3}$$

where $\boldsymbol{A}_t \circ$ is defined as matrix multiplication but applies thinning operation instead of multiplication. The process defined with Equation (2.3) is Markov process of order one. Notice that $\boldsymbol{A}_t$ is a random matrix and $E\boldsymbol{A}_t = \boldsymbol{A}$ where $\boldsymbol{A} = \begin{bmatrix} \alpha_1 p_1 & 0 \\ 0 & \alpha_2 p_2 \end{bmatrix}$. It can be easily shown that eigenvalues of matrix $\boldsymbol{A}$ are inside the unit circle. Following the definition of the thinning operator Lemma 2.1 holds.

Lemma 2.1. *Thinning operator properties:*

1. $E(\boldsymbol{A}_t \circ \boldsymbol{Z}) = \boldsymbol{A} E\boldsymbol{Z}$
2. $E((\boldsymbol{A}_t \circ \boldsymbol{Z})(\boldsymbol{A}_t \circ \boldsymbol{Z})') = \boldsymbol{A} E(\boldsymbol{Z}\boldsymbol{Z}')\boldsymbol{A} + \begin{bmatrix} \alpha_1^2 p_1(1-p_1)EX^2 & 0 \\ 0 & \alpha_2^2 p_2(1-p_2)EY^2 \end{bmatrix}$
   $+ \begin{bmatrix} \alpha_1 p_1(1-\alpha_1)EX & 0 \\ 0 & \alpha_2 p_2(1-\alpha_2)EY \end{bmatrix}$
3. $\boldsymbol{A}_t \circ (\boldsymbol{A}_{t-1} \circ \boldsymbol{Z}) = \boldsymbol{A}_t \boldsymbol{A}_{t-1} \circ \boldsymbol{Z}$

where $\boldsymbol{Z} = (X, Y)'$ *is a two dimensional random vector.*

Theorem 2.1. *There exist a unique strictly stationary ergodic solution to equation (2.3).*

Proof. Applying equation (2.3) k-times and following properties 3 from Lemma(2.1), we can define process $\boldsymbol{Z}_t$ as

$$\boldsymbol{Z}_t = \prod_{i=0}^{k-1} \boldsymbol{A}_{t-i} \circ \boldsymbol{Z}_{t-k} + \sum_{i=1}^{k} \prod_{j=0}^{i-1} \boldsymbol{A}_{t-j} \circ \boldsymbol{e}_{t-i} + \boldsymbol{e}_t.$$

Denote with $\boldsymbol{Y}_t = \boldsymbol{Z}_t - \sum_{i=1}^{k} \prod_{j=0}^{i-1} \boldsymbol{A}_{t-j} \circ \boldsymbol{e}_{t-i} - \boldsymbol{e}_t$. Then

$$\begin{aligned} E\left(\boldsymbol{Y}_t \boldsymbol{Y}_t'\right) &= E\left(\prod_{i=0}^{k-1} \boldsymbol{A}_{t-i} \circ \boldsymbol{Z}_{t-k}\right)\left(\prod_{i=0}^{k-1} \boldsymbol{A}_{t-i} \circ \boldsymbol{Z}_{t-k}\right)' = \\ &= E((\boldsymbol{A}_{kt} \circ \boldsymbol{Z}_{t-k})\,(\boldsymbol{A}_{kt} \circ \boldsymbol{Z}_{t-k})') \end{aligned} \tag{2.4}$$

where matrix $\boldsymbol{A}_{kt}$ is obtained after multiplications of matrices A_t k-times and has the form $\boldsymbol{A}_{kt} = \begin{bmatrix} \alpha_{1kt} & 0 \\ 0 & \alpha_{2kt} \end{bmatrix}$. Elements are random variables distributed as $\alpha_{1kt} : \begin{pmatrix} \alpha_1^k & 0 \\ p_1^k & 1-p_1^k \end{pmatrix}$ and $\alpha_{2kt} : \begin{pmatrix} \alpha_2^k & 0 \\ p_2^k & 1-p_2^k \end{pmatrix}$. Following property 2 from Lemma (2.1) equation (2.4) becomes

$$\begin{aligned} E(\boldsymbol{Y}_t \boldsymbol{Y}_t') &= \begin{bmatrix} (\alpha_1 p_1)^k & 0 \\ 0 & (\alpha_2 p_2)^k \end{bmatrix} E(\boldsymbol{Z}_t \boldsymbol{Z}_t') \begin{bmatrix} (\alpha_1 p_1)^k & 0 \\ 0 & (\alpha_2 p_2)^k \end{bmatrix} + \\ &+ \begin{bmatrix} \alpha_1^{2k} p_1^k (1-p_1^k) E X_{1,t}^2 & 0 \\ 0 & \alpha_2^{2k} p_2^k (1-p_2^k) E X_{2,t}^2 \end{bmatrix} + \\ &+ \begin{bmatrix} (\alpha_1 p_1)^k (1-\alpha_1^k) E X_{1,t} & 0 \\ 0 & (\alpha_2 p_2)^k (1-\alpha_2^k) E X_{2,t} \end{bmatrix} \end{aligned}$$

Thus $E(\boldsymbol{Y}_t \boldsymbol{Y}_t')$ tends to zero matrix as k tends to infinity which proves the existence of the solution. We can conclude that the solution to equation (2.3) in mean square sense is of the form

$$\boldsymbol{Z}_t = \sum_{i=1}^{k} \prod_{j=0}^{i-1} \boldsymbol{A}_{t-j} \circ \boldsymbol{e}_{t-i} - \boldsymbol{e}_t. \tag{2.5}$$

Let us prove the uniqueness of the solution. Suppose that there exists another solution $\boldsymbol{Z}_t^*$. Since $\boldsymbol{Z}_t^*$ is also a solution to the equation, then $E(\boldsymbol{Z}_t) = E(\boldsymbol{Z}_t^*)$. Then

$$
\begin{aligned}
& E((\boldsymbol{Z}_t - \boldsymbol{Z}_t^*)(\boldsymbol{Z}_t - \boldsymbol{Z}_t^*)') = E(\boldsymbol{A}_t \circ (\boldsymbol{Z}_{t-1} - \boldsymbol{Z}_{t-1}^*)(\boldsymbol{A}_t \circ (\boldsymbol{Z}_{t-1} - \boldsymbol{Z}_{t-1}^*))') = \\
= \;& \boldsymbol{A}E((\boldsymbol{Z}_{t-1} - \boldsymbol{Z}_{t-1}^*)(\boldsymbol{Z}_{t-1} - \boldsymbol{Z}_{t-1}^*)')\boldsymbol{A} + \\
+ \;& \begin{bmatrix} \alpha_1^2 p_1(1-p_1)E(X_{1,t-1} - X_{1,t-1}^*)^2 & 0 \\ 0 & \alpha_2^2 p_2(1-p_2)E(X_{2,t-1} - X_{2,t-1}^*)^2 \end{bmatrix} + \\
+ \;& \begin{bmatrix} \alpha_1 p_1(1-\alpha_1)E(X_{1,t-1} - X_{1,t-1}^*) & 0 \\ 0 & \alpha_2 p_2(1-\alpha_2)E(X_{2,t-1} - X_{2,t-1}^*) \end{bmatrix}.
\end{aligned}
$$

From this matrix equality we have, componentwise:

$$
\begin{aligned}
E(X_{1,t} - X_{1,t}^*) &= \alpha_1^2 p_1^2 E(X_{1,t-1} - X_{1,t-1}^*)^2 + \alpha_1^2 p_1(1-p_1)E(X_{1,t-1} - X_{1,t-1}^*)^2 \\
&+ \alpha_1 p_1(1-\alpha_1)E(X_{1,t-1} - X_{1,t-1}^*) \leq \\
&\leq \alpha_1 p_1^2 E(X_{1,t-1} - X_{1,t-1}^*)^2 + \alpha_1 p_1(1-p_1)E(X_{1,t-1} - X_{1,t-1}^*)^2 \\
&= \alpha_1 p_1 E(X_{1,t-1} - X_{1,t-1}^*)^2
\end{aligned}
$$

and

$$
E(X_{1,t} - X_{1,t}^*)(X_{2,t} - X_{2,t}^*) = \alpha_1\alpha_2 E(X_{1,t-1} - X_{1,t-1}^*)(X_{2,t-1} - X_{2,t-1}^*)
$$

which implies the uniqueness of the solution $\boldsymbol{Z}_t$.

The solution (2.5) has the same functional form for each t and therefore it must be stationary. This implies stationarity of $\boldsymbol{Z}_t$. Further, as $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ are independent for $i \neq j$ and random matrices $\boldsymbol{A}_i$ and $\boldsymbol{A}_j$ are independent for $i \neq j$, $\{\boldsymbol{e}_t, \boldsymbol{A}_t\}$ is a sequence of independent identically distributed random vectors and thus ergodic. As σ-field $\mathcal{G}_t$ generated by $(\boldsymbol{Z}_t, \boldsymbol{Z}_{t-1}, ...)$ is a subset of σ field $\mathcal{F}_t$ generated by $(\boldsymbol{e}_t, \boldsymbol{A}_t, \boldsymbol{e}_{t-1}, \boldsymbol{A}_{t-1}, ...)$ for any t, we have that $\boldsymbol{Z}_t$ is also ergodic. $\square$

Remark: For some special cases this model reduces to some known models.

- For parameters $p_1 = 1$ and $p_2 = 1$ the process reduces to the bivariate process discussed in [10].
- If we suppose independencies between processes $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, processes (2.1) and (2.2) are two univariate integer valued processes discussed in [2].

3. Properties of the Model

Following the definition of BVDINAR(1) model we continue the research by deriving properties of the model. In this section we discuss marginal as well as joint probability distribution of the model. Also, we investigate the conditional probability distribution function and conditional and unconditional moments. The conditional moments are derived for k-steps ahead and some discussion on their asymptotic properties is given.

We investigate the distribution of the process through the probability generating function. First, we state marginal distribution for process $X_{1,t}$ which is given with

$$\begin{aligned}\Phi_{X_{1,t}}(s) &= p_1\Phi_{\alpha_1\circ X_{1,t-1}}\Phi_{\varepsilon_{1,t}}(s) + (1-p_1)\Phi_{\varepsilon_{1,t}}(s) = \ldots = \\ &= p_1^t\Phi_{X_{1,0}}(1-\alpha_1^t(1-s))\Phi_{\varepsilon_{1,1}} + (1-p_1)\sum_{i=0}^{t-1} p_1^i\Phi_{\varepsilon_{1,t-i}}(s) = \\ &= (p_1^t\Phi_{X_{1,0}}(1-\alpha_1^t(1-s)) + (1-p_1^t))\Phi_{\varepsilon_{1,t}}(s) \longrightarrow \Phi_{\varepsilon_{1,t}}(s)\end{aligned}$$

as t tends to infinity, and similarly for $X_{2,t}$

$$\Phi_{X_{2,t}}(s) = (p_2^t\Phi_{X_{2,0}}(1-\alpha_2^t(1-s)) + (1-p_2^t))\Phi_{\varepsilon_{2,t}}(s) \longrightarrow \Phi_{\varepsilon_{2,t}}(s)$$

The probability generating function of our bivariate model is given with the following equation.

$$\begin{aligned}\Phi_{X_{1,t},X_{2,t}}(s_1,s_2) &= p_1^tp_2^t\Phi_{X_{1,0},X_{2,0}}(1-\alpha_1^t(1-s_1), 1-\alpha_2^t(1-s_2))\cdot \\ &\cdot \prod_{i=0}^{t-1}\Phi_{\varepsilon_{1,t-i},\varepsilon_{2,t-i}}(1-\alpha_1^i(1-s_1), 1-\alpha_2^i(1-s_2)) + \\ &+ \sum_{i=1}^{t}(p_1p_2)^{i-1}\Big[p_1(1-p_2)\Phi_{X_{1,t-i}}(1-\alpha_1^i(1-s_1)) + \\ &+ (1-p_1)p_2\Phi_{X_{2,t-i}}(1-\alpha_2^i(1-s_2)) + (1-p_1)(1-p_2)\Big]\cdot \\ &\cdot \prod_{j=1}^{i}\Phi_{\varepsilon_{1,t-j+1},\varepsilon_{2,t-j+i}}(1-\alpha_1^{i-1}(1-s_1), 1-\alpha_2^{i-1}(1-s_2)).\end{aligned}$$

Since the model is constructed by assuming the bivariate distribution of the innovation processes, let us denote with μ_{ε_1} and μ_{ε_2} expected values of two innovation processes respectively with σ_{ε_1} and σ_{ε_2} their variances. The covariance between the innovation processes denote with ϕ. Then we have

$$E(X_{i,t}) = \frac{\mu_{\varepsilon_i}}{1-\alpha_ip_i}$$

$$Var(X_{i,t}) = \frac{\alpha_ip_i(1-\alpha_i)(1-\alpha_ip_i)\mu_{\varepsilon_i} + \alpha_i^2p_i(1-p_i)\mu_{\varepsilon_i}^2 + (1-\alpha_ip_i)^2\sigma_{\varepsilon_i}^2}{(1-\alpha_i^2p_i)(1-\alpha_ip_i)}.$$

Covariance of each process is given with

$$Cov(X_{i,t+k}, X_{i,t}) = \alpha_i^kp_i^kVar(X_{i,t}) \tag{3.1}$$

so covariance tends to zero as $k \longrightarrow \infty$. Covariances between these two processes is equal to

$$Cov(X_{1,t}, X_{2,t}) = \frac{Cov(\varepsilon_{1,t}, \varepsilon_{2,t})}{1-\alpha_1\alpha_2p_1p_2}.$$

Thus, the covariance of the processes is proportional to the covariance of the innovation processes. We also derive k-steps cross-covariance which is

$$Cov(X_{i,t+k}, X_{j,t}) = \alpha_i^k p_i^k Cov(X_{i,t}, X_{j,t}). \tag{3.2}$$

so we can conclude from equations (3.1) and (3.2) that covariances tend to zero as k tends to infinity.

Following the definition of the processes $X_{1,t}$ and $X_{2,t}$ we derive k-steps conditional expectation as

$$\begin{aligned} E(X_{i,t+k}|X_{1,t}, X_{2,t}) &= \alpha_i p_i E(X_{i,t+k-1}|X_{1,t}, X_{2,t}) + \mu_{\varepsilon_i} = \dots = \\ &= (\alpha_i p_i)^k X_{it} + \mu_{\varepsilon_i} \frac{1-(\alpha_i p_i)^k}{1-\alpha_i p_i}. \end{aligned}$$

From the above equation we have that conditional expectation tends to unconditional as k tends to infinity. Before we derive k-steps conditional variance of the processes, we focus on conditional statistical measures of the thinning component. The following lemma gives some properties of the thinning operator.

Lemma 3.1. *For the binomial thinning operator and processes defined with (2.1) and (2.2) the following equations hold:*

1. $E(\alpha_i \circ X_{i,t+k}|X_{1,t}, X_{2,t}) = p_i E(\alpha_i \circ X_{i,t+k-1}|X_{1,t}, X_{2,t}) + E(\alpha_i \circ \varepsilon_{i,t+k})$
2. $E((\alpha_i \circ X_{i,t+k})^2|X_{1,t}, X_{2,t}) = p_i E((\alpha_i^2 \circ X_{i,t+k-1})^2|X_{1,t}, X_{2,t}) +$ $2\alpha_i p_i E(\alpha_i^2 \circ X_{i,t+k-1}|X_{1,t}, X_{2,t}) E(\varepsilon_{i,t+k}) + E(\alpha_i \circ \varepsilon_{i,t+k})^2$
3. $Var(\alpha_i \circ X_{i,t+k}|X_{1,t}, X_{2,t}) = p_i^k Var(\alpha_i^{k+1} \circ X_{i,t}|X_{1,t}, X_{2,t}) +$ $(1-p_i)\sum_{j=1}^{k} p_i^j (E(\alpha_i^{j+1} \circ X_{1,t+k-j}|X_{1,t}, X_{2,t}))^2 + \sum_{j=0}^{k-1} p_i^j Var(\alpha_i^{j+1} \circ \varepsilon_{i,t+k-j})$

Proof. The first two properties follow directly from the definition of the processes. They imply that $Var(\alpha_i \circ X_{i,t+k}|X_{1,t}, X_{2,t}) = p_i Var(\alpha_i^2 \circ X_{i,t+k-1}|X_{1,t}, X_{2,t}) + p_i(1-p_i)(E(\alpha_i^2 \circ X_{i,t+k-1}|X_{1,t}, X_{2,t}))^2 + Var(\alpha_i \circ \varepsilon_{i,t+k})$. Applying this conditional variance k-time we obtain property 3. □

Conclusions of Lemma (3.1) are building blocks for k-steps conditional variance which is given as

$$\begin{aligned} &Var(X_{i,t+k}|X_{1,t}, X_{2,t}) = p_i Var(\alpha_i \circ X_{i,t+k-1}|X_{1,t}, X_{2,t}) + \\ &+ p_i(1-p_i)(E(\alpha_i \circ X_{i,t+k-1}|X_{1,t}, X_{2,t}))^2 + Var(\varepsilon_{i,t+k}) = \dots = \alpha_i^k p_i^k (1-\alpha_i^k) X_{i,t} + \\ &+ \left((1-p_i)(\alpha_i^k p_i^k X_{i,t} + \frac{1-\alpha_i^k p_i^k}{1-\alpha_i p_i} \mu_{\varepsilon_i})^2 \alpha_i^2 p_i + \sigma_{\varepsilon_i}^2 - \mu_{\varepsilon_i} \right) + \frac{1-\alpha_i^{2k} p_i^k}{1-\alpha_i^2 p_i} + \mu_{\varepsilon_i} \frac{1-\alpha_i^k p_i^k}{1-\alpha_i p_i}. \end{aligned}$$

Once again we can notice that conditional variance tends to unconditional as k tends to infinity.

4. The model with Poisson marginal distribution

In this section we introduce an assumption about the distribution of the innovation processes. For the specified bivariate distribution, in subsection 4.1., we derive unconditional and k-steps ahead conditional expectation and variance. Estimation of the parameters is discussed in detail. Two methods for parameter estimation are suggested in subsection 4.2., conditional maximum likelihood and the method of moments. The former method is derived in subsection 4.2.1. where the conditional probability mass function of the model can also be found. The latter method is investigated in subsection 4.2.2. There is also a detailed discussion about asymptotic properties of the estimates. Finally, in subsection 4.3. both methods are tested on simulated data sets.

4.1. Model

Innovation processes $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ follow bivariate Poisson distribution with parameters $(\lambda_1, \lambda_2, \phi)$ where the probability distribution function is

$$(4.1)\qquad \begin{aligned} P(\varepsilon_{1,t} = u, \varepsilon_{2,t} = v) \quad &= \quad e^{-(\lambda_1+\lambda_2-\phi)} \frac{(\lambda_1-\phi)^u}{u!} \frac{(\lambda_2-\phi)^v}{v!} \cdot \\ &\cdot \quad \sum_{i=0}^{k} \binom{u}{i}\binom{v}{i} i! \left(\frac{\phi}{(\lambda_1-\phi)(\lambda_2-\phi)}\right)^i, \end{aligned}$$

$k = \min(u, v)$. While parameter λ_i determines the mean value and variance for processes $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, a correlation between these two processes is determined with parameter ϕ. Marginal distribution of innovation processes is Poisson with parameter λ_1 and λ_2 respectively. If we assume that $\phi = 0$ bivariate process reduces to two independent univariate processes.

The mean and variance of processes $X_{1,t}$ and $X_{2,t}$ are

$$(4.2)\qquad EX_{i,t} = \frac{\lambda_i}{1-\alpha_i p_i}$$

$$(4.3)\qquad Var(X_{i,t}) = \frac{(1-\alpha_i^2 p_i)(1-\alpha_i p_i)\lambda_i + \alpha_i^2 p_i(1-p_i)\lambda_i^2}{(1-\alpha_i^2 p_i)(1-\alpha_i p_i)^2}$$

k-steps covariance and cross-covariance are given by equations (3.1) and (3.2) and they do not depend on marginal distributions, while

$$(4.4)\qquad Cov(X_{1,t}, X_{2,t}) = \frac{\phi}{1-\alpha_1\alpha_2 p_1 p_2}.$$

Further, conditional expectation and conditional variance are given by the following equations, respectively

$$E(X_{i,t+k}|X_{1,t}, X_{2,t}) = (\alpha_i p_i)^k X_{i,t} + \lambda_i \frac{1-(\alpha_i p_i)^k}{1-\alpha_i p_i}$$

$$Var(X_{i,t+k}|X_{1,t},X_{2,t}) = (\alpha_i p_i)^k(1-\alpha_i^k)X_{i,t} + \alpha_i^2 p_i(1-p_i)\cdot$$
$$\cdot \quad \left((\alpha_i p_i)^k X_{i,t} + \lambda_i \frac{1-(\alpha_i p_i)^k}{1-\alpha_i p_i}\right)^2 \frac{1-\alpha_i^{2k}p_i^k}{1-\alpha_i^2 p_i} + \lambda_i \frac{1-(\alpha_i p_i)^k}{1-\alpha_i p_i}$$

4.2. Parameter Estimation

Our model is defined with seven parameters. In this subsection we state two methods for their estimation, conditional maximum likelihood (CML) and the method of moments (MM).

4.2.1. Conditional Maximum Likelihood

The conditional probability function of $(X_{1,t}, X_{2,t})$ conditioned on $(X_{1,t-1}, X_{2,t-1})$ is a weighted sum of conditional probabilities of components that define processes $X_{1,t}$, $X_{2,t}$.

$$\begin{aligned}
& P(X_{1,t+1}=x, X_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v) = \\
= \quad & p_1p_2P(\alpha_1\circ X_{1,t}+\varepsilon_{1,t+1}=x, \alpha_2\circ X_{2,t}+\varepsilon_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v) + \\
+ \quad & (1-p_1)p_2P(\varepsilon_{1,t+1}=x, \alpha_2\circ X_{2,t}+\varepsilon_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v) + \\
+ \quad & p_1(1-p_2)P(\alpha_1\circ X_{1,t}+\varepsilon_{1,t+1}=x, \varepsilon_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v) + \\
+ \quad & (1-p_1)(1-p_2)P(\varepsilon_{1,t+1}=x, \varepsilon_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v)
\end{aligned}$$

Following independency of $(\varepsilon_{1,n}, \varepsilon_{2,n})$ from $(X_{1,m}, X_{2,m})$ for $n > m$ and the fact that thinning $\alpha_1 \circ X_{1,t}$, when $X_{1,t} = u$ is known, is a binomial random variable with parameters (α_1, u) and it is independent of $X_{2,t}$ (and similarly for $\alpha_2 \circ X_{2,t}$) we obtain the following equations.

$$\begin{aligned}
& P(\alpha_1\circ X_{1,t}+\varepsilon_{1,t+1}=x, \alpha_2\circ X_{2,t}+\varepsilon_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v) = \\
= \quad & \sum_{m=s_1}^{x}\sum_{n=s_2}^{y} P(Bin(\alpha_1,u)=x-m)P(Bin(\alpha_2,v)=y-n)\cdot \\
& \cdot P(\varepsilon_{1,t+1}=m, \varepsilon_{2,t+1}=n)
\end{aligned}$$

where $s_1 = \max(x-u, 0)$ and $s_2 = \max(y-v, 0)$,

$$\begin{aligned}
& P(\varepsilon_{1,t+1}=x, \alpha_2\circ X_{2,t}+\varepsilon_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v) = \\
= \quad & \sum_{n=s_2}^{y} P(Bin(\alpha_2,v)=y-n)P(\varepsilon_{1,t+1}=x, \varepsilon_{2,t+1}=n)
\end{aligned}$$

and

$$\begin{aligned}
& P(\alpha_1\circ X_{1,t}+\varepsilon_{1,t+1}=x, \varepsilon_{2,t+1}=y|X_{1,t}=u, X_{2,t}=v) = \\
= \quad & \sum_{m=s_1}^{x} P(Bin(\alpha_1,u)=x-m)P(\varepsilon_{1,t+1}=m, \varepsilon_{2,t+1}=y)
\end{aligned}$$

Probability distribution function for innovation processes is given by equation (4.1).

To estimate the parameters we need to maximize the conditional log-likelihood function of the form

$$L = \sum_{i=1}^{n-1} \ln P(X_{1,i+1}, X_{2,i+1} | X_{1,i}, X_{2,i}, \boldsymbol{\theta})$$

where $\boldsymbol{\theta}$ is the vector of parameters. Due to the complexity of the function L, the minimization procedure is obtained numerically.

4.2.2. Method of moments

Suppose we have a random sample $(X_{1,j}, X_{2,j})_{j=\overline{1,n}}$. Let us introduce new parameters $u_1 = \alpha_1 p_1$ and $u_2 = \alpha_2 p_2$. This parameter is estimated by equation (3.1) so

$$\hat{u}_i = \frac{\gamma_{X_i X_i}(1)}{\gamma_{X_i}(0)}.$$

Now from equation (4.2) we have

$$\hat{\lambda}_i = (1 - \hat{u}_i)\overline{X}_i$$

where $\overline{X}_i$ is the sample mean, $\gamma_{X_i}(0)$ is the sample variance and $\gamma_{X_i X_i}(1)$ is the sample auto-covariance for lag 1 for series $\{X_{i,t}\}$. Further, from equation (4.3) we estimate parameter α_i as

$$\hat{\alpha}_i = \frac{\hat{\lambda}_i(1 - \hat{u}_i) - \hat{u}_i^2\hat{\lambda}_i^2 - (1 - \hat{u}_i)^2\gamma_{X_i}(0)}{\hat{u}_i(1 - \hat{u}_i)\hat{\lambda}_i - \hat{u}_i\hat{\lambda}_i^2 - \hat{u}_i(1 - \hat{u}_i)^2\gamma_{X_i}(0)}.$$

We estimate parameter p_i as $\hat{p}_i = \frac{\hat{u}_i}{\hat{\alpha}_i}$.

Finally, parameter ϕ is estimated from equation (4.4) as

$$\hat{\phi} = \frac{\gamma_{X_1 X_2}(0)}{1 - \hat{u}_1\hat{u}_2}.$$

Denote with $\boldsymbol{m}_n$ a vector of moments used for the estimation of the parameters (n in the index stands for the sample size), where m_j, $j = \overline{1,3}$ are the sample mean, variance and the auto-covariance function for series $X_{1,t}$, m_j, $j = \overline{4,6}$ the corresponding functions for series $X_{2,t}$ and m_7 the sample cross-correlation function. Consequently, $\overline{m}_j = 1/n \sum_{i=1}^{n} m_j(X_{1,i}, X_{2,i})$, $j = 1,7$. The vector of the parameters is $\boldsymbol{\theta} = (\lambda_1, \lambda_2, \alpha_1, \alpha_2, p_1, q_2, \phi)$. If we expand in a linear Taylor series the set of solved moment equations around the true values of parameter $\boldsymbol{\theta}_0$, we obtain the linear approximation

$$\mathbf{0} \approx [\boldsymbol{m}_n(\boldsymbol{\theta}_0)] + \boldsymbol{G}_n(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$$

from which we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \approx -[\boldsymbol{G}_n(\boldsymbol{\theta}_0)]^{-1} \sqrt{n}[\boldsymbol{m}_n(\boldsymbol{\theta}_0)]. \tag{4.5}$$

Since moment functions are all continues, according to Central Limit Theorem, $\sqrt{n}[\boldsymbol{m}_n(\boldsymbol{\theta}_0)]$ is normally distributed with mean $\mathbf{0}$ as n tends to infinity. Further, all functions in the moment equations are functionally independent, thus $\boldsymbol{G}_n(\boldsymbol{\theta}_0)$ converges to a nonsingular matrix of constants. All this implies normal distribution of the right side of equation (4.5) with mean $\mathbf{0}$ and asymptotic covariance matrix $\boldsymbol{\Phi}$.

The asymptotic covariance matrix for the estimated parameters is obtained as $\boldsymbol{\Phi} = \frac{1}{n}\left(\boldsymbol{G}_n^{'}(\boldsymbol{\theta})\boldsymbol{F}^{-1}\boldsymbol{G}_n(\boldsymbol{\theta})\right)^{-1}$. Matrix $\boldsymbol{F} = [F_{jk}]$ is a 7×7 matrix the elements of which are $F_{jk} = \frac{1}{n}\sum_{i=1}^{n}[(m_j(X_{1,i}.X_{2,i}) - \overline{m}_j)(m_k(X_{1,i}, X_{2,i}) - \overline{m}_k)]$, $\boldsymbol{G} = [G_{ij}]$ is 7×7 matrix the elements of which are partial derivatives of moment functions used to estimate the parameters.

$$G_{ij} = \frac{\partial m_i}{\partial \theta_j}$$

The elements of the matrix are given in Appendix.

4.3. Simulation results

In the previous two subsections we stated the methods for estimation of unknown parameters. To demonstrate the efficiency of these methods we preform tests on a simulated data sets. The data sets contains 100 samples of length 50, 100, 500 and 1000. We perform two tests with different parameters: a) $\alpha_1 = 0.6$, $\alpha_2 = 0.55$, $p = 0.55$, $q = 0.4$, $\lambda_1 = 5$, $\lambda_2 = 3$, $\phi = 1$; b) $\alpha_1 = 0.3$, $\alpha_2 = 0.2$, $p = 0.65$, $q = 0.6$, $\lambda_1 = 5$, $\lambda_2 = 3$, $\phi = 1$. Results are presented in Table 4.1 and Table 4.2. Table 4.1 and Table 4.2 contain the estimated values as well as the standard deviation of these estimates. The first column of both tables states the number of Monte Carlo replications. For the parameter estimation procedure we use R software.

We can notice that both methods converge to real values with an increase in the sample length. The CML method provides good results even for samples of length 100, while there are some deviations for samples of length 50, but the results are still not too away from true values. The MM method is quite unprecise for small samples. When parameters α_i take higher values, the performance of the method is slightly better but it still needs more than 100 elements per sample to achieve precise results. For samples of length 1000 MM the results are quite good. Due to the complexity of the model, computation times for samples of length 1000 for MM and CML are incomparable. Computation time for the first method is measured in seconds and for the second in hours.

Table 4.1: Estimated parameters with Method of moments (standard deviation are given underneath)

	α_1	α_2	p_1	p_2	λ_1	λ_2	ϕ
50	0.5986	0.4473	0.4963	0.488	5.0695	3.0345	1.1524
	0.178	0.3332	0.2456	0.3652	0.9019	0.6349	1.092
100	0.5971	0.4879	0.4941	0.4683	5.1058	2.9775	1.0263
	0.1106	0.2658	0.1545	0.618	0.5776	0.4359	0.8233
500	0.593	0.5059	0.5473	0.4523	4.9423	2.9246	1.1887
	0.0407	0.1044	0.0624	0.1366	0.2946	0.1953	0.4252
1000	0.5956	0.5193	0.5498	0.4288	4.9314	2.9277	1.1826
	0.0298	0.0615	0.0504	0.0927	0.2302	0.1391	0.3177
50	0.3794	0.3447	0.4679	0.3286	5.2949	3.2607	1.267
	0.2888	0.3592	0.351	0.3866	0.7692	0.5501	0.767
100	0.3692	0.3005	0.4865	0.3616	5.2166	3.1819	1.1993
	0.2419	0.3178	0.2975	0.3814	0.6407	0.3967	0.5342
500	0.3392	0.2368	0.5562	0.4844	5.0773	3.1136	1.1667
	0.0705	0.1879	0.141	0.3265	0.3022	0.1827	0.2404
1000	0.324	0.2078	0.5925	0.5932	5.0449	3.117	1.1451
	0.0437	0.1158	0.0893	0.2906	0.2023	0.1351	0.1661

Table 4.2: Estimated parameters with Maximum likelihood method (standard deviation are given underneath)

	α_1	α_2	p_1	p_2	λ_1	λ_2	ϕ
50	0.6013	0.5081	0.5396	0.472	4.955	2.9713	1.0303
	0.1269	0.3027	0.1646	0.3119	0.6323	0.4463	1.1042
100	0.5973	0.5645	0.5304	0.4204	4.9394	2.9868	1.0471
	0.0687	0.2132	0.1052	0.2324	0.4437	0.3459	0.6186
500	0.598	0.5529	0.548	0.3874	4.9278	2.9858	1.082
	0.0271	0.0809	0.0418	0.0866	0.1784	0.1305	0.2845
1000	0.5982	0.5504	0.5514	0.3925	4.9352	2.9826	1.069
	0.021	0.0502	0.0306	0.0608	0.1342	0.0995	0.2049
50	0.3426	0.2086	0.6218	0.7023	5.0648	3.1589	1.1625
	0.1566	0.2317	0.274	0.304	0.547	0.3648	0.5412
100	0.3342	0.1943	0.5986	0.7001	5.0943	3.1565	1.1116
	0.125	0.1754	0.248	0.2898	0.4813	0.3032	0.4312
500	0.327	0.2046	0.6251	0.6208	5.0766	3.1254	1.0842
	0.0478	0.1557	0.1096	0.3057	0.2361	0.1714	0.1818
1000	0.3099	0.207	0.6474	0.6189	5.0629	3.1269	1.066
	0.0385	0.0841	0.0762	0.2755	0.1622	0.1205	0.1438

5. Real data example

In order to demonstrate the practical aspect of the introduced model in this section we analyze some series of counts from real life. We will discuss the results and statistical properties of the observed series. The idea is to show that BVDINAR(1) model is adequate for time series of counts with certain features and not only for the observed two series. The obtained results will be compared with some other INAR models.

We analyze data downloaded from `http://www.forecastingprinciples.com/index.php?option=com_content&view=article&id=47&Itemid=250` where we focus on monthly counts of larceny (LAR) and criminal mischiefs (CMIS) in Pittsburgh from January 1990 till December 2001. These two acts can be classified as light criminal activities. LAR can be described as a nonviolent theft, while CMIS represents injures, damages, or destruction of any property of another or public property without consent. So the same social environment brings forth these two crime acts. On the other side, there is no evident lag 1 cross-correlations between these two series although they are correlated. The mean values for LAR and CMIS are 1.18 and 1.43, while the variances are 1.37 and 1.95, respectively. The correlation coefficient for the series is 0.31. There is some overdispersion with both series, but as it is not big, Poisson distribution should not be discard. There are 144 observations. Data series, autocorrelation functions and cross-correlation functions are given in Figure 5.1. Data series have very similar patterns, which supports the assumption that they are dependent. We can notice the presence of lag 1 autocorrelation in both series. Also, there is little or no cross-correlation.

Goodness of fit criteria that we use are the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the root mean square error (RMS). While the first two measure the quality of the assumed distribution for the series, the third one suggests one step ahead forecasting power. BVDINAR(1) model is compared with two similar models but with constant coefficients BVPOIBINAR(1) and BVNBIBINAR(1) both introduced in [10] and with BVFPIB model presented in [12] which suggests lag 1 cross-correlation between series. All these models assume dependencies between innovation processes where innovation processes are generated by bivariate Poisson distribution with BVPOIBINAR(1) and BVFPIB and by bivariate negative binomial with BVNBIBINAR(1). There is no intention to prove that one model is the best overall but only to show that in some examples with specific features one model is better than the other. The results are summarized in Table 5.1. To estimate the parameters of the models we use the ML method because of two reasons. First, we want to take into consideration the probability distribution of the processes and, second, the sample is not large enough to achieve MM precise estimates.

From Table 5.1 we can notice that the models based on bivariate Poisson distribution are more adequate then the one based on the bivariate negative binomial with respect to all three criteria. BVDINAR(1) model gives the best scores. There are some improvements with respect to BVPOIBINAR(1) model, although their BIC values are almost the same, which stems from the fact that BVDINAR(1) has two more parameters then BVPOIBINAR(1). Inclusion of lag 1 cross-correlation into the model worsens the results, which is shown by BVFPIB model. Its results are similar to the other two Poisson distribution-based models, but a bit worse. When we are talking about the values of the estimated parameters for BVDINAR(1), we can notice that low, but still present, autocorrelation with series LAR is captured with a low value of parameter p_1. According to the estimated parameters, the expected values for the processes are 1.192 and 1.426 and the variances 1.437 and 1.482 for LAR and CMIS, respectively, which is quite close to the true values. Also, ϕ different from zero implies that innovation processes and thus LAR and CMIS are dependent.

Table 5.1: Parameter estimates, AIC, BIC, RMS for larceny and criminal mischief (values in the brackets are standard errors of the estimates).

Model	CML estimates	AIC	BIC	RMS LAR	RMS CMIS
BVDINAR(1)	$\hat{\alpha}_1 = 0.977(0.273)$, $\hat{p}_1 = 0.183(0.094)$ $\hat{\alpha}_2 = 0.323(0.188)$, $\hat{p}_2 = 0.537(0.35)$ $\hat{\lambda}_1 = 0.979(0.106)$, $\hat{\lambda}_2 = 1.179(0.125)$ $\hat{\phi} = 0.286(0.094)$	854.6	875.38	1.139	1.360
BVPOIBINAR(1)	$\hat{\alpha}_1 = 0.135(0.074)$, $\hat{\lambda}_1 = 1.036(0.121)$ $\hat{\alpha}_2 = 0.121(0.065)$, $\hat{\lambda}_2 = 1.261(0.13)$ $\hat{\phi} = 0.337(0.093)$	860.8	875.66	1.142	1.367
BVNBIBINAR(1)	$\hat{\alpha}_1 = 0.089(0.057)$, $\hat{\lambda}_1 = 1.558(0.222)$ $\hat{\alpha}_2 = 0.002(0.0715)$, $\hat{\lambda}_2 = 2.253(0.367)$ $\hat{\beta} = 0.918(0.067)$	905.43	920.28	1.24	1.61
BVFPIB	$\hat{\alpha}_{11} = 0.271(0.109)$, $\hat{\alpha}_{12} = 0.002(0.079)$ $\hat{\alpha}_{21} = 0.113(0.089)$, $\hat{\alpha}_{22} = 0.159(0.089)$ $\hat{\lambda}_1 = 1.094(0.129)$, $\hat{\lambda}_2 = 1.165(0.121)$ $\hat{\phi} = 0.404(0.105)$	881.11	901.19	1.161	1.369

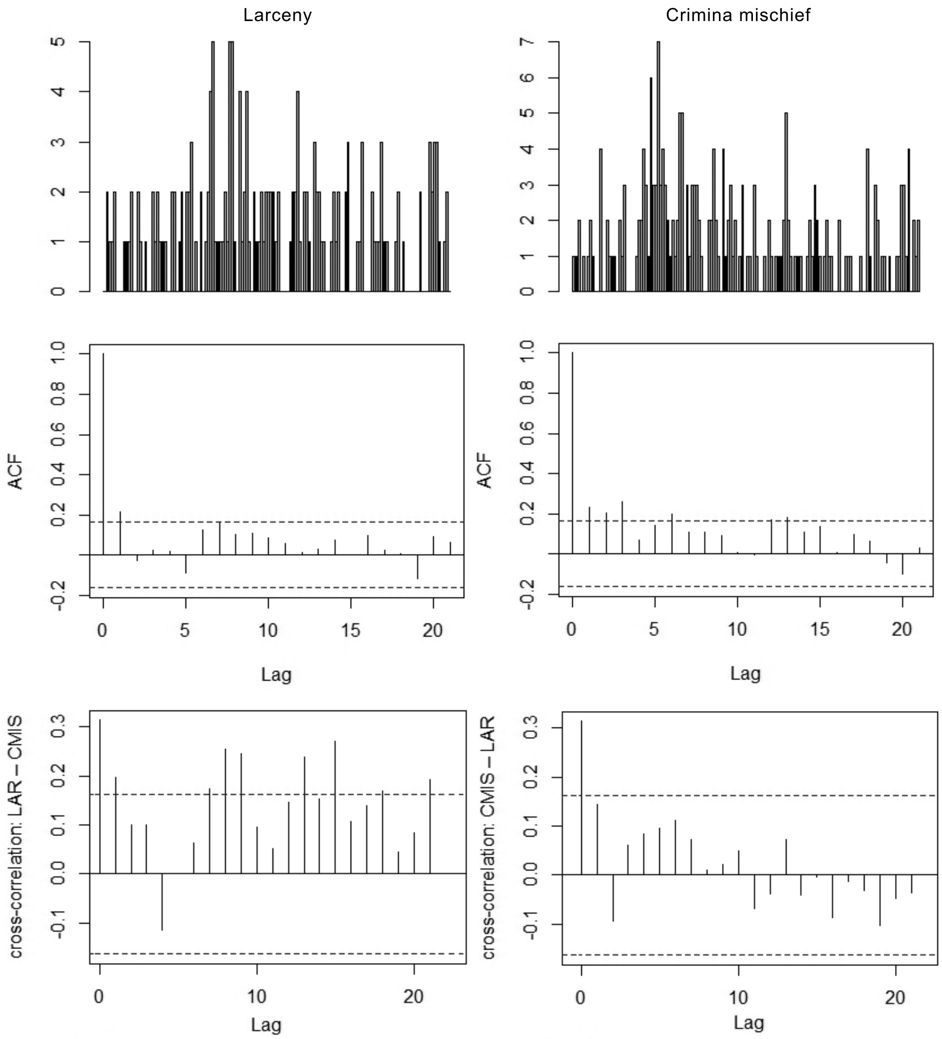

FIG. 5.1: Larceny series and Criminal mischief series

6. Conclusion

In this article we presented a bivariate INAR model with dependent innovation components while the survival component of the model is defined with random coefficients. This model extends the univariate INAR model with a random coefficient to the bivariate case. Also, it introduces the random coefficient concept for bivariate models with dependent innovation processes. The existence and stationarity of the model are proved. We focus on a special case of the model when the innovation processes follow binomial Poisson distribution. Besides the statistical measures of the model, we also discussed conditional maximum likelihood and the method of moments for the parameters estimation. We found out that the first method provides more precise estimates whereas the second one is welcome for big data sets due to the much shorter computation time. In the end, we discussed the application of the model to real data. We gave some comments on the features of the observed time series of counts and suggested for which series the model is appropriate.

Further development of the model might consider an introduction of some other thinning operators and some other bivariate distributions for the innovation processes.

7. Appendix

The elements of matrix $\boldsymbol{G}$ are

$$
\begin{aligned}
G_{11} &= \frac{1}{1-\alpha_1 p_1}\\
G_{13} &= \frac{p_1\lambda_1}{(1-\alpha_1 p_1)^2} G_{15} = \frac{\alpha_1\lambda_1}{(1-\alpha_1 p_1)^2}\\
G_{12} &= G_{14} = G_{16} = G_{17} = 0\\
G_{21} &= \frac{1}{1-\alpha_1 p_1} + \frac{2\alpha_1^2 p_1(1-p_1)\lambda_1}{(1-\alpha_1^2 p_1)(1-\alpha_1 p_1)^2}\\
G_{23} &= \frac{p_1\lambda_1}{(1-\alpha_1 p_1)^2} + \frac{\alpha_1 p_1(1-p_1)\lambda_1^2(2-\alpha_1 p_1(1+\alpha_1^2 p_1))}{(1-\alpha_1^2 p_1)^2(1-\alpha_1 p_1)^2}\\
G_{25} &= \frac{\alpha_1\lambda_1}{(1-\alpha_1 p_1)^2} + \frac{\alpha_1^2\lambda_1^2(1-2p_1+\alpha_1 p_1^2(1+\alpha_1-\alpha_1^2))}{(1-\alpha_1^2 p_1)^2(1-\alpha_1 p_1)^2}\\
G_{22} &= G_{24} = G_{26} = G_{27} = 0\\
G_{31} &= \frac{\alpha_1 p_1}{1-\alpha_1 p_1} + \frac{2\alpha_1^3 p_1^2(1-p_1)\lambda_1}{(1-\alpha_1^2 p_1)(1-\alpha_1 p_1)}\\
G_{33} &= \frac{p_1\lambda_1}{(1-\alpha_1 p_1)^2} + \frac{\alpha_1^2 p_1^2(1-p_1)(3-\alpha_1 p_1(2+\alpha_1))\lambda_1^2}{(1-\alpha_1^2 p_1)^2(1-\alpha_1 p_1)^2}\\
G_{35} &= \frac{\alpha_1\lambda_1}{(1-\alpha_1 p_1)^2} + \frac{\alpha_1^3 p_1\lambda_1^2(2-(3+\alpha_1+\alpha_1^2)p_1+2\alpha_1(1+\alpha_1)p_1^2-\alpha_1^3 p_1^3)}{(1-\alpha_1^2 p_1)^2(1-\alpha_1 p_1)^2}\\
G_{32} &= G_{34} = G_{36} = G_{37} = 0\\
G_{42} &= \frac{1}{1-\alpha_2 p_2}\\
G_{44} &= \frac{p_2\lambda_2}{(1-\alpha_2 p_2)^2}
\end{aligned}
$$

$$G_{46} = \frac{\alpha_2\lambda_2}{(1-\alpha_2 p_2)^2}$$
$$G_{41} = G_{43} = G_{45} = G_{47} = 0$$
$$G_{52} = \frac{1}{1-\alpha_2 p_2} + \frac{2\alpha_2^2 p_2(1-p_2)\lambda_2}{(1-\alpha_2^2 p_2)(1-\alpha_2 p_2)^2}$$
$$G_{54} = \frac{p_2\lambda_2}{(1-\alpha_2 p_2)^2} + \frac{\alpha_2 p_2(1-p_2)\lambda_2^2(2-\alpha_2 p_2(1+\alpha_2^2 p_2))}{(1-\alpha_2^2 p_2)^2(1-\alpha_2 p_2)^2}$$
$$G_{56} = \frac{\alpha_2\lambda_2}{(1-\alpha_2 p_2)^2} + \frac{\alpha_2^2\lambda_2^2(1-2p_2+\alpha_2 p_2^2(1+\alpha_2-\alpha_2^2))}{(1-\alpha_2^2 p_2)^2(1-\alpha_2 p_2)^2}$$
$$G_{51} = G_{53} = G_{55} = G_{57} = 0$$
$$G_{62} = \frac{\alpha_2 p_2}{1-\alpha_2 p_2} + \frac{2\alpha_2^3 p_2^2(1-p_2)\lambda_2}{(1-\alpha_2^2 p_2)(1-\alpha_2 p_2)}$$
$$G_{64} = \frac{p_2\lambda_2}{(1-\alpha_2 p_2)^2} + \frac{\alpha_2^2 p_2^2(1-p_2)(3-\alpha_2 p_2(2+\alpha_2))\lambda_2^2}{(1-\alpha_2^2 p_2)^2(1-\alpha_2 p_2)^2}$$
$$G_{66} = \frac{\alpha_2\lambda_2}{(1-\alpha_2 p_2)^2} + \frac{\alpha_2^3 p_2\lambda_2^2(2-(3+\alpha_2+\alpha_2^2)p_2+2\alpha_2(1+\alpha_2)p_2^2-\alpha_2^3 p_2^3)}{(1-\alpha_2^2 p_2)^2(1-\alpha_2 p_2)^2}$$
$$G_{61} = G_{63} = G_{65} = G_{67} = 0$$
$$G_{71} = G_{72} = 0$$
$$G_{73} = \frac{p_1\phi}{1-\alpha_1\alpha_2 p_1 p_2} + \frac{\alpha_1\alpha_2 p_1^2 p_2\phi}{(1-\alpha_1\alpha_2 p_1 p_2)^2}$$
$$G_{74} = \frac{\alpha_1^2 p_1^2 p_2\phi}{(1-\alpha_1\alpha_2 p_1 p_2)^2}$$
$$G_{75} = \frac{\alpha_1\phi}{1-\alpha_1\alpha_2 p_1 p_2} + \frac{\alpha_1^2\alpha_2 p_1 p_2\phi}{(1-\alpha_1\alpha_2 p_1 p_2)^2}$$
$$G_{76} = \frac{\alpha_1^2\alpha_2 p_1^2\phi}{(1-\alpha_1\alpha_2 p_1 p_2)^2}$$
$$G_{77} = \frac{\alpha_1 p_1}{1-\alpha_1\alpha_2 p_1 p_2}$$

REFERENCES

1. M.A. Al-Osh, A.A. Alzaid, *First-order Integer-valued Autoregressive (INAR(1))* Process, Journal of Time Series Analysis 8 (1987), 261–275.
2. H.S. Bakouch, M.M. Ristić, *Zero truncated Poisson integer-valued AR(1) model*, Metrika 72 (2010), 265–280.
3. K. Brännäs, J. Nordström, *A Bivariate Integer Valued Allocation Model for Guest Nights in Hotels and Cottages*, in: UmeåEconomic Studies (2000), 547.
4. J. Franke, S.T. Rao, *Multivariate First-Order Integer-Valued Autoregressions*, Technical Report, Mathematical Department, UMIST.
5. D. Karlis, X. Pedeli, *Flexible Bivariate INAR(1) Processes Using Copulas*, Communications in Statistics - Theory and Methods 42 (2013), 723–740.
6. A. Latour, *The Multivatiate GINAR(p) Process*, Advances in Applied Probability 29 (1997), 228–248.
7. E. McKenzie, *Discrete variate time series*, in: Rao, C.R. Shanbhag, D.N. ed., Handbook of Statistics, Amsterdam: Elsevier (2003), 573–606.
8. A.S. Nastić, *On suitability of negative binomial marginals and geometric counting sequence in some applications of Combined INAR(p) model*, Facta Universitatis 29 (2014), 25–42.
9. A.S. Nastić, M.M. Ristić, P.M. Popović, *Estimation in a Bivariate Integer-Valued Autoregressive Process*, Communication in Statistics - Theory and Methods (2015), to appear.
10. X. Pedeli, D. Karlis, *A Bivariate INAR(1) Process with Application*, Statistical Modelling 11 (2011), 325–349.

11. X. Pedeli, D. Karlis, *On Estimation of the Bivariate Poisson INAR Process*, Communications in Statistics - Simulation and Computation 42 (2013), 514–533.

12. X. Pedeli, D. Karlis, *Some properties of multivariate INAR(1) processes*, Computational Statistics and Data Analysis 67 (2013), 213–225.

13. X. Pedeli, D. Karlis, *On composite likelihood estimation of a multivariate INAR(1) model*, Journal of Time Series Analysis 34 (2013), 206–220.

14. A.M.M.S. Quoreshi, *Bivariate time series modelling of financial count data*, Communication in Statistics - Theory and Methods 35 (2006), 1343–1358.

15. M.M. Ristić, A.S. Nastić, K. Jayakumar, H.S. Bakouch, *A bivariate INAR(1) time series model with geometric marginals*, Applied Mathematics Letters 25 (2012), 481–485.

16. M.G. Scotto, C.H. Weiss, M.E. Silva, I. Pereira, *Bivariate binomial autoregressive models*, Journal of Multivariate Analysis 125 (2014), 233–251.

17. C.H. Weiss, *Thinning operations for modeling time series of counts - a survey*, Advances in Statistical Analysis 92 (2008), 319–341.

18. H. Zheng, I.V. Basawa, S. Datta, *First-order random coefficient integer-valued autoregressive processes*, Journal of Statistical Planning and Inference 137 (2007), 212–229.

Predrag M. Popović
University of Niš
Department of Mathematics, Informatics and Physics
Faculty of Civil Engineering and Architecture
Niš, Serbia
popovicpredrag@yahoo.com