

# A copula-based bivariate integer-valued autoregressive process with application

Andrius Buteikis\*, Remigijus Leipus

*Institute of Applied Mathematics, Faculty of Mathematics and Informatics, Vilnius University, Lithuania*

[andrius.buteikis@mif.vu.lt](mailto:andrius.buteikis@mif.vu.lt) (A. Buteikis), [remigijus.leipus@mif.vu.lt](mailto:remigijus.leipus@mif.vu.lt) (R. Leipus)

Received: 21 August 2018, Revised: 12 December 2018, Accepted: 28 January 2019,  
Published online: 12 March 2019

**Abstract** A bivariate integer-valued autoregressive process of order 1 (BINAR(1)) with copula-joint innovations is studied. Different parameter estimation methods are analyzed and compared via Monte Carlo simulations with emphasis on estimation of the copula dependence parameter. An empirical application on defaulted and non-defaulted loan data is carried out using different combinations of copula functions and marginal distribution functions covering the cases where both marginal distributions are from the same family, as well as the case where they are from different distribution families.

**Keywords** Count data, BINAR, Poisson, negative binomial distribution, copula, FGM copula, Frank copula, Clayton copula

**2010 MSC** 60G10, 62M10, 62H12

## 1 Introduction

Different financial institutions that issue loans do this following company-specific (and/or country-defined) rules which act as a safeguard against loans issued to people who are known to be insolvent. However, striving for higher profits might motivate some companies to issue loans to higher risk clients. Usually company's methods for evaluating loan risk are not publicly available. However, one way to evaluate if there aren't too many knowingly very high-risk loans issued, and if insolvent clients

---

\*Corresponding author.

are adequately separated from responsible clients, would be to look at the quantity of defaulted and non-defaulted loans issued each day. The adequacy of company's rules for issuing loans can be analysed by modelling via copulas the dependence between the number of defaulted loans and the number of non-defaulted loans. The advantage of such approach is that copulas allow to model the marginal distributions (possibly from different distribution families) and their dependence structure (which is described via a copula) separately. Because of this feature, copulas were applied to many different fields, including survival analysis, hydrology, insurance risk analysis as well as finance (for examples of copula applications, see [3] or [4]), which also included the analysis of loans and their default rates.

The dependence of the default rate of loans on different credit risk categories was analysed in [5]. To model the dependence, copulas from ten different families were applied and three model selection tests were carried out. Because of the small sample size (24 observations per risk category) most of the copula families were not rejected and a single best copula model was not selected. To analyse whether dependence is affected by time, Fenech et al. [6] estimated the dependence among four different loan default indexes before the global financial crisis and after. They have found that the dependence was different in these periods. Four copula families were used to estimate the dependence between the default index pairs. While these studies were carried out for continuous data, discrete models created with copulas are less investigated: Genest and Nešlehová [8] discussed the differences and challenges of using copulas for discrete data compared to continuous data. Note that the previously mentioned studies assumed that the data does not depend on its own previous values. By using bivariate integer-valued autoregressive models (BINAR) it is possible to account for both the discreteness and autocorrelation of the data. Furthermore, copulas can be used to model the dependence of innovations in the BINAR(1) models: Karlis and Pedeli [10] used the Frank copula and the normal copula to model the dependence of the innovations of the BINAR(1) model.

In this paper we expand on using copulas in BINAR models by analysing additional copula families for the innovations of the BINAR(1) model and analyse different methods for BINAR(1) model parameter estimation. We also present a two-step method for the parameter estimation of the BINAR(1) model, where we estimate the model parameters separately from the dependence parameter of the copula. These estimation methods (including the one used in [10]) are compared via Monte Carlo simulations. Finally, in order to analyse the presence of autocorrelation and copula dependence in loan data, an empirical application is carried out for empirical weekly loan data.

The paper is organized as follows. Section 2 presents the BINAR(1) process and its main properties, Section 3 presents the main properties of copulas as well as some copula functions. Section 4 compares different estimation methods for the BINAR(1) model and the dependence parameter of copulas via Monte Carlo simulations. In Section 5 an empirical application is carried out using different combinations of copula functions and marginal distribution functions. Conclusions are presented in Section 6.

## 2 The bivariate INAR(1) process

The BINAR(1) process was introduced in [18]. In this section we will provide the definition of the BINAR(1) model and will formulate its properties.

**Definition 1.** Let  $\mathbf{R}_t = [R_{1,t}, R_{2,t}]'$ ,  $t \in \mathbb{Z}$ , be a sequence of independent identically distributed (i.i.d.) nonnegative integer-valued bivariate random variables. A bivariate integer-valued autoregressive process of order 1 (BINAR(1)),  $\mathbf{X}_t = [X_{1,t}, X_{2,t}]'$ ,  $t \in \mathbb{Z}$ , is defined as:

$$\mathbf{X}_t = \mathbf{A} \circ \mathbf{X}_{t-1} + \mathbf{R}_t = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \circ \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} R_{1,t} \\ R_{2,t} \end{bmatrix}, \quad t \in \mathbb{Z}, \quad (1)$$

where  $\alpha_j \in [0, 1)$ ,  $j = 1, 2$ , and the symbol ‘ $\circ$ ’ is the thinning operator which also acts as the matrix multiplication. So the  $j$ th ( $j = 1, 2$ ) element is defined as an INAR process of order 1 (INAR(1)):

$$X_{j,t} = \alpha_j \circ X_{j,t-1} + R_{j,t}, \quad t \in \mathbb{Z}, \quad (2)$$

where  $\alpha_j \circ X_{j,t-1} := \sum_{i=1}^{X_{j,t-1}} Y_{j,t,i}$  and  $Y_{j,t,1}, Y_{j,t,2}, \dots$  is a sequence of i.i.d. Bernoulli random variables with  $\mathbb{P}(Y_{j,t,i} = 1) = \alpha_j = 1 - \mathbb{P}(Y_{j,t,i} = 0)$ ,  $\alpha_j \in [0, 1)$ , such that these sequences are mutually independent and independent of the sequence  $\mathbf{R}_t$ ,  $t \in \mathbb{Z}$ . For each  $t$ ,  $\mathbf{R}_t$  is independent of  $\mathbf{X}_s$ ,  $s < t$ .

Properties of the thinning operator are provided in [17] and [19] with proofs for selected few. We present the main properties of the thinning operator which will be used later on in the case of BINAR(1) model. Denote by ‘ $\stackrel{d}{=}$ ’ the equality of distributions.

**Theorem 1** (Thinning operator properties). *Let  $X, X_1, X_2$  be nonnegative integer-valued random variables, such that  $\mathbb{E}Z^2 < \infty$ ,  $Z \in \{X, X_1, X_2\}$ ,  $\alpha, \alpha_1, \alpha_2 \in [0, 1)$  and let ‘ $\circ$ ’ be the thinning operator. Then the following properties hold:*

- (a)  $\alpha_1 \circ (\alpha_2 \circ X) \stackrel{d}{=} (\alpha_1 \alpha_2) \circ X$ ;
- (b)  $\alpha \circ (X_1 + X_2) \stackrel{d}{=} \alpha \circ X_1 + \alpha \circ X_2$ ;
- (c)  $\mathbb{E}(\alpha \circ X) = \alpha \mathbb{E}(X)$ ;
- (d)  $\text{Var}(\alpha \circ X) = \alpha^2 \text{Var}(X) + \alpha(1 - \alpha) \mathbb{E}(X)$ ;
- (e)  $\mathbb{E}((\alpha \circ X_1) X_2) = \alpha \mathbb{E}(X_1 X_2)$ ;
- (f)  $\text{Cov}(\alpha \circ X_1, X_2) = \alpha \text{Cov}(X_1, X_2)$ ;
- (g)  $\mathbb{E}((\alpha_1 \circ X_1)(\alpha_2 \circ X_2)) = \alpha_1 \alpha_2 \mathbb{E}(X_1 X_2)$ .

$X_{j,t}$ , defined in eq. (2), has two random components: the survivors of the elements of the process at time  $t - 1$ , each with the probability of survival  $\alpha_j$ , which are denoted by  $\alpha_j \circ X_{j,t-1}$ , and the elements which enter in the system in the interval  $(t - 1, t]$ , which are called arrival elements and denoted by  $R_{j,t}$ . We can obtain a

moving average representation by substitutions and the properties of the thinning operator as in [1] or [11, p. 180]:

$$X_{j,t} = \alpha_j \circ X_{j,t-1} + R_{j,t} \stackrel{d}{=} \sum_{k=0}^{\infty} \alpha_j^k \circ R_{j,t-k}, \quad j = 1, 2, t \in \mathbb{Z}, \quad (3)$$

where convergence on the right-hand side holds a.s.

Now we present some properties of the BINAR(1) model. They will be used when analysing some of parameter estimation methods. The proofs for these properties can be easily derived and some of them are provided in [17].

**Theorem 2** (Properties of the BINAR(1) process). *Let  $\mathbf{X}_t = (X_{1,t}, X_{2,t})'$  be a non-negative integer-valued time series given in Def. 1 and  $\alpha_j \in [0, 1)$ ,  $j = 1, 2$ . Let  $\mathbf{R}_t = (R_{1,t}, R_{2,t})'$ ,  $t \in \mathbb{Z}$ , be nonnegative integer-valued random variables with  $\mathbb{E}(R_{j,t}) = \lambda_j$  and  $\text{Var}(R_{j,t}) = \sigma_j^2 < \infty$ ,  $j = 1, 2$ . Then the following properties hold:*

- (a)  $\mathbb{E}X_{j,t} = \mu_{X_j} = \frac{\lambda_j}{1-\alpha_j}$ ;
- (b)  $\mathbb{E}(X_{j,t}|X_{j,t-1}) = \alpha_j X_{j,t-1} + \lambda_j$ ;
- (c)  $\text{Var}(X_{j,t}) = \sigma_{X_j}^2 = \frac{\sigma_j^2 + \alpha_j \lambda_j}{1 - \alpha_j^2}$ ;
- (d)  $\text{Cov}(X_{i,t}, R_{j,t}) = \text{Cov}(R_{i,t}, R_{j,t})$ ,  $i \neq j$ ;
- (e)  $\text{Cov}(X_{j,t}, X_{j,t+h}) = \alpha_j^h \sigma_{X_j}^2$ ,  $h \geq 0$ ;
- (f)  $\text{Corr}(X_{j,t}, X_{j,t+h}) = \alpha_j^h$ ,  $h \geq 0$ ;
- (g)  $\text{Cov}(X_{i,t}, X_{j,t+h}) = \frac{\alpha_j^h}{1 - \alpha_i \alpha_j} \text{Cov}(R_{i,t}, R_{j,t})$ ,  $i \neq j$ ,  $h \geq 0$ ;
- (h)  $\text{Corr}(X_{i,t+h}, X_{j,t}) = \frac{\alpha_i^h \sqrt{(1 - \alpha_i^2)(1 - \alpha_j^2)}}{(1 - \alpha_i \alpha_j) \sqrt{(\sigma_i^2 + \alpha_i \lambda_i)(\sigma_j^2 + \alpha_j \lambda_j)}} \text{Cov}(R_{i,t}, R_{j,t})$ ,  
 $i \neq j$ ,  $h \geq 0$ ;

Similarly to (3), we have that

$$\mathbf{X}_t \stackrel{d}{=} \sum_{k=0}^{\infty} \mathbf{A}^k \circ \mathbf{R}_{t-k},$$

where convergence on the right-hand side holds a.s.

Hence, the distributional properties of the BINAR(1) process can be studied in terms of  $\mathbf{R}_t$  values. Note also, that according to [12], if  $\alpha_j \in [0, 1)$ ,  $j = 1, 2$ , then there exists a unique stationary nonnegative integer-valued sequence  $\mathbf{X}_t$ ,  $t \in \mathbb{Z}$ , satisfying (1).

From the covariance and correlation (see (g) and (h) in Theorem 2) of the BINAR(1) process we see that the dependence between  $X_{1,t}$  and  $X_{2,t}$  depends on the joint distribution of the innovations  $R_{1,t}, R_{2,t}$ . Pedeli and Karlis [18] analysed BINAR(1) models when the innovations were linked by either a bivariate Poisson or a bivariate negative binomial distribution, where the covariance of the innovations can be easily expressed in terms of their joint distribution parameters. Karlis and Pedeli [10] analysed two cases when the distributions of innovations of a BINAR(1) model are linked by either the Frank copula or a normal copula with either Poisson or negative binomial marginal distributions. We will expand their work by analysing additional copulas for the BINAR(1) model innovation distribution as well as estimation methods for the distribution parameters.

### 3 Copulas

In this section we recall the definition and main properties of bivariate copulas, mainly following [8, 15] and [21] for the continuous and discrete settings.

#### 3.1 Copula definition and properties

Copulas are used for modelling the dependence between several random variables. The main advantage of using copulas is that they allow to model the marginal distributions separately from their joint distribution. In this paper we are using two-dimensional copulas which are defined as follows:

**Definition 2.** A 2-dimensional copula  $C : [0, 1]^2 \rightarrow [0, 1]$  is a function with the following properties:

(i) for every  $u, v \in [0, 1]$ :

$$C(u, 0) = C(0, v) = 0; \tag{4}$$

(ii) for every  $u, v \in [0, 1]$ :

$$C(u, 1) = u, \quad C(1, v) = v; \tag{5}$$

(iii) for any  $u_1, u_2, v_1, v_2 \in [0, 1]$  such that  $u_1 \leq u_2$  and  $v_1 \leq v_2$ :

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \tag{6}$$

(this is also called *the rectangle inequality*).

The theoretical foundation of copulas is given by Sklar’s theorem:

**Theorem 3 ([20]).** Let  $H$  be a joint cumulative distribution function (cdf) with marginal distributions  $F_1, F_2$ . Then there exists a copula  $C$  such that for all  $(x_1, x_2) \in [-\infty, \infty]^2$ :

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{7}$$

If  $F_i$  is continuous for  $i = 1, 2$  then  $C$  is unique; otherwise  $C$  is uniquely determined only on  $\text{Ran}(F_1) \times \text{Ran}(F_2)$ , where  $\text{Ran}(F)$  denotes the range of the cdf  $F$ . Conversely, if  $C$  is a copula and  $F_1, F_2$  are distribution functions, then the function  $H$ , defined by equation (7) is a joint cdf with marginal distributions  $F_1, F_2$ .

If a pair of random variables  $(X_1, X_2)$  has continuous marginal cdfs  $F_i(x)$ ,  $i = 1, 2$ , then by applying the probability integral transformation one can transform them into random variables  $(U_1, U_2) = (F_1(X_1), F_2(X_2))$  with uniformly distributed marginals which can then be used when modelling their dependence via a copula. More about Copula theory, properties and applications can be found in [15] and [9].

### 3.2 Copulas with discrete marginal distributions

Since innovations of a BINAR(1) model are nonnegative integer-valued random variables, one needs to consider copulas linking discrete distributions. In this section we will mention some of the key differences when copula marginals are discrete rather than continuous.

Firstly, as mentioned in Theorem 3, if  $F_1$  and  $F_2$  are discrete marginals then a unique copula representation exists only for values in the range of  $\text{Ran}(F_1) \times \text{Ran}(F_2)$ . However, the lack of uniqueness does not pose a problem in empirical applications because it implies that there may exist more than one copula which describes the distribution of the empirical data. Secondly, regarding concordance and discordance, the discrete case has to allow for ties (i.e. when two variables have the same value), so the concordance measures (Spearman's rho and Kendal's tau) are margin-dependent, see [21]. There are several modifications proposed for Spearman's rho, however, none of them are margin-free. Furthermore, Genest and Nešlehová [8] state that estimators of the dependence parameter  $\theta$  based on Kendall's tau or its modified versions are biased, and estimation techniques based on maximum likelihood are recommended. As such, we will not examine estimation methods based on concordance measures. Another difference from the continuous case is the use of the probability mass function (pmf) instead of the probability density function when estimating the model parameters which will be seen in Section 4.

### 3.3 Some concrete copulas

In this section we will present several bivariate copulas, which will be used later when constructing and evaluating the BINAR(1) model. For all the copulas discussed, the following notation is used:  $u_1 := F_1(x_1)$ ,  $u_2 := F_2(x_2)$ , where  $F_1, F_2$  are marginal cumulative distribution functions (cdfs) of discrete random variables, and  $\theta$  is the dependence parameter.

#### Farlie–Gumbel–Morgenstern copula

The Farlie–Gumbel–Morgenstern (FGM) copula has the following form:

$$C(u_1, u_2; \theta) = u_1 u_2 (1 + \theta(1 - u_1)(1 - u_2)). \quad (8)$$

The dependence parameter  $\theta$  can take values from the interval  $[-1, 1]$ . If  $\theta = 0$ , then the FGM copula collapses to independence. Note that the FGM copula can only model weak dependence between two marginals (see [15]). The copula when  $\theta = 0$  is called a product (or independence) copula:

$$C(u_1, u_2) = u_1 u_2. \quad (9)$$

Since the product copula corresponds to independence, it is important as a benchmark.

*Frank copula*

The Frank copula has the following form:

$$C(u_1, u_2; \theta) = -\frac{1}{\theta} \log \left( 1 + \frac{(\exp(-\theta u_1) - 1)(\exp(-\theta u_2) - 1)}{\exp(-\theta) - 1} \right).$$

The dependence parameter  $\theta$  can take values from  $(-\infty, \infty) \setminus \{0\}$ . The Frank copula allows for both positive and negative dependence between the marginals.

*Clayton copula*

The Clayton copula has the following form:

$$C(u_1, u_2; \theta) = \max\{u_1^{-\theta} + u_2^{-\theta} - 1, 0\}^{-\frac{1}{\theta}}, \tag{10}$$

with the dependence parameter  $\theta \in [-1, \infty) \setminus \{0\}$ . The marginals become independent when  $\theta \rightarrow 0$ . It can be used when the correlation between two random variables exhibits a strong left tail dependence – if smaller values are strongly correlated and high values are less correlated. The Clayton copula can also account for negative dependence when  $\theta \in [-1, 0)$ . For more properties of this copula, see the recent paper by Manstavičius and Leipus [14].

**4 Parameter estimation of the copula-based BINAR(1) model**

In this section we examine different BINAR(1) model parameter estimation methods and provide a two-step method for separate estimation of the copula dependence parameter. Estimation methods are compared via Monte Carlo simulations. Let  $\mathbf{X}_t = (X_{1,t}, X_{2,t})'$  be a non-negative integer-valued time series given in Def. 1, where the joint distribution of  $(R_{1,t}, R_{2,t})'$ , with marginals  $F_1, F_2$ , is linked by a copula  $C(\cdot, \cdot)$ :

$$\mathbb{P}(R_{1,t} \leq x_1, R_{2,t} \leq x_2) = C(F_1(x_1), F_2(x_2))$$

and let  $C(u_1, u_2) = C(u_1, u_2; \theta)$ , where  $\theta$  is a dependence parameter.

*4.1 Conditional least squares estimation*

The Conditional least squares (CLS) estimator minimizes the squared distance between  $\mathbf{X}_t$  and its conditional expectation. Similarly to the method in [19] for the INAR(1) model, we construct the CLS estimator in the case of the BINAR(1) model.

Using Theorem 1 we can write the vector of conditional means as

$$\boldsymbol{\mu}_{t|t-1} := \begin{bmatrix} \mathbb{E}(X_{1,t}|X_{1,t-1}) \\ \mathbb{E}(X_{2,t}|X_{2,t-1}) \end{bmatrix} = \begin{bmatrix} \alpha_1 X_{1,t-1} + \lambda_1 \\ \alpha_2 X_{2,t-1} + \lambda_2 \end{bmatrix}, \tag{11}$$

where  $\lambda_j := \mathbb{E}R_{j,t}, j = 1, 2$ . In order to calculate the CLS estimators of  $(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$  we define the vector of residuals as the difference between the observations and their conditional expectation:

$$\mathbf{X}_t - \boldsymbol{\mu}_{t|t-1} = \begin{bmatrix} X_{1,t} - \alpha_1 X_{1,t-1} - \lambda_1 \\ X_{2,t} - \alpha_2 X_{2,t-1} - \lambda_2 \end{bmatrix}.$$

Then, given a sample of  $N$  observations,  $\mathbf{X}_1, \dots, \mathbf{X}_N$ , the CLS estimators of  $\alpha_j, \lambda_j, j = 1, 2$ , are found by minimizing the sum

$$Q_j(\alpha_j, \lambda_j) := \sum_{t=2}^N (X_{j,t} - \alpha_j X_{j,t-1} - \lambda_j)^2 \longrightarrow \min_{\alpha_j, \lambda_j}, \quad j = 1, 2.$$

By taking the derivatives with respect to  $\alpha_j$  and  $\lambda_j, j = 1, 2$ , and equating them to zero we get:

$$\hat{\alpha}_j^{\text{CLS}} = \frac{\sum_{t=2}^N (X_{j,t} - \bar{X}_j)(X_{j,t-1} - \bar{X}_j)}{\sum_{t=2}^N (X_{j,t-1} - \bar{X}_j)^2} \tag{12}$$

and

$$\hat{\lambda}_j^{\text{CLS}} = \frac{1}{N-1} \left( \sum_{t=2}^N X_{j,t} - \hat{\alpha}_j^{\text{CLS}} \sum_{t=2}^N X_{j,t-1} \right). \tag{13}$$

The asymptotic properties of the CLS estimators for the INAR(1) model case are provided in [13, 19, 2] and can be applied to the BINAR(1) parameter estimates, specified via equations (12) and (13). By the fact that the  $j$ -th component of the BINAR(1) process is an INAR(1) itself, we can formulate the following theorem for the marginal parameter vector distributions (see [2]):

**Theorem 4.** Let  $\mathbf{X}_t = (X_{1,t}, X_{2,t})'$  be defined in Def. 1 and let the parameter vector of (2) be  $(\alpha_j, \lambda_j)'$ . Assume that  $\hat{\alpha}_j^{\text{CLS}}$  and  $\hat{\lambda}_j^{\text{CLS}}$  are the CLS estimators of  $\alpha_j$  and  $\lambda_j, j = 1, 2$ . Then:

$$\sqrt{N} \begin{pmatrix} \hat{\alpha}_j^{\text{CLS}} - \alpha_j \\ \hat{\lambda}_j^{\text{CLS}} - \lambda_j \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0}_2, \mathbf{B}_j),$$

where

$$\mathbf{B}_j = \begin{bmatrix} \mathbb{E}X_{j,t}^2 & \mathbb{E}X_{j,t} \\ \mathbb{E}X_{j,t} & 1 \end{bmatrix}^{-1} \mathbf{A}_j \begin{bmatrix} \mathbb{E}X_{j,t}^2 & \mathbb{E}X_{j,t} \\ \mathbb{E}X_{j,t} & 1 \end{bmatrix}^{-1},$$

$$\mathbf{A}_j = \alpha_j(1 - \alpha_j) \begin{bmatrix} \mathbb{E}X_{j,t}^3 & \mathbb{E}X_{j,t}^2 \\ \mathbb{E}X_{j,t}^2 & \mathbb{E}X_{j,t} \end{bmatrix} + \sigma_j^2 \begin{bmatrix} \mathbb{E}X_{j,t}^2 & \mathbb{E}X_{j,t} \\ \mathbb{E}X_{j,t} & 1 \end{bmatrix}, \quad j = 1, 2.$$

Here, according to BINAR(1) properties in Theorem 2,

$$\mathbb{E}X_{j,t} = \frac{\lambda_j}{1 - \alpha_j}, \quad \mathbb{E}X_{j,t}^2 = \frac{\sigma_j^2 + \alpha_j \lambda_j}{1 - \alpha_j^2} + \frac{\lambda_j^2}{(1 - \alpha_j)^2},$$

$$\mathbb{E}X_{j,t}^3 = \frac{\mathbb{E}R_{j,t}^3 - 3\sigma_j^2(1 + \lambda_j) - \lambda_j^3 + 2\lambda_j}{1 - \alpha_j^3} + 3 \frac{\sigma_j^2 + \alpha_j \lambda_j}{1 - \alpha_j^2} - 2 \frac{\lambda_j}{1 - \alpha_j}$$

$$+ 3 \frac{\lambda_j(\sigma_j^2 + \alpha_j \lambda_j)}{(1 - \alpha_j)(1 - \alpha_j^2)} + \frac{\lambda_j^3}{(1 - \alpha_j)^3}.$$



For the Poisson marginal distribution case the asymptotic variance matrix can be expressed as (see [7])

$$B_j = \begin{bmatrix} \frac{\alpha_j(1 - \alpha_j)^2}{\lambda_j} + 1 - \alpha_j^2 & -(1 + \alpha_j)\lambda_j \\ -(1 + \alpha_j)\lambda_j & \lambda_j + \frac{1 + \alpha_j}{1 - \alpha_j}\lambda_j^2 \end{bmatrix}, \quad j = 1, 2.$$

Furthermore, for a more general case, [12] proved that the CLS estimators of a multivariate generalized integer-valued autoregressive process (GINAR) are asymptotically normally distributed.

Note that

$$\mathbb{E}(X_{1,t} - \alpha_1 X_{1,t-1} - \lambda_1)(X_{2,t} - \alpha_2 X_{2,t-1} - \lambda_2) = \text{Cov}(R_{1,t}, R_{2,t}), \quad (14)$$

which follows from

$$\begin{aligned} &\mathbb{E}(X_{1,t} - \alpha_1 X_{1,t-1} - \lambda_1)(X_{2,t} - \alpha_2 X_{2,t-1} - \lambda_2) \\ &= \mathbb{E}(\alpha_1 \circ X_{1,t-1} - \alpha_1 X_{1,t-1})(\alpha_2 \circ X_{2,t-1} - \alpha_2 X_{2,t-1}) \\ &\quad + \mathbb{E}(\alpha_1 \circ X_{1,t-1} - \alpha_1 X_{1,t-1})(R_{2,t} - \lambda_2) \\ &\quad + \mathbb{E}(\alpha_2 \circ X_{2,t-1} - \alpha_2 X_{2,t-1})(R_{1,t} - \lambda_1) \\ &\quad + \mathbb{E}(R_{1,t} - \lambda_1)(R_{2,t} - \lambda_2) \end{aligned}$$

since the first three summands are zeros.

**Example 4.1.** Assume that the joint pmf of  $(R_{1,t}, R_{2,t})$  is given by bivariate Poisson distribution:

$$\mathbb{P}(R_{1,t} = k, R_{2,t} = l) = \sum_{i=0}^{\min\{k,l\}} \frac{(\lambda_1 - \lambda)^{k-i} (\lambda_2 - \lambda)^{l-i} \lambda^i}{(k-i)!(l-i)!i!} e^{-(\lambda_1 + \lambda_2 - \lambda)},$$

where  $k, l = 0, 1, \dots, \lambda_j > 0, j = 1, 2, 0 \leq \lambda < \min\{\lambda_1, \lambda_2\}$ . Then, for each  $j = 1, 2$ , the marginal distribution of  $R_{j,t}$  is Poisson with parameter  $\lambda_j$  and  $\text{Cov}(R_{1,t}, R_{2,t}) = \lambda$ . If  $\lambda = 0$  then the two variables are independent.

**Example 4.2.** Assume that the joint pmf of  $(R_{1,t}, R_{2,t})$  is bivariate negative binomial distribution given by

$$\begin{aligned} \mathbb{P}(R_{1,t} = k, R_{2,t} = l) &= \frac{\Gamma(\beta + k + l)}{\Gamma(\beta)k!l!} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2 + \beta}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2 + \beta}\right)^l \\ &\quad \times \left(\frac{\beta}{\lambda_1 + \lambda_2 + \beta}\right)^\beta, \end{aligned}$$

where  $k, l = 0, 1, \dots, \lambda_j > 0, j = 1, 2, \beta > 0$ . Then, for each  $j = 1, 2$ , the marginal distribution of  $R_{j,t}$  is negative binomial with parameters  $\beta$  and  $p_j = \beta/(\lambda_j + \beta)$  and  $\mathbb{E}R_{j,t} = \lambda_j, \text{Var}(R_{j,t}) = \lambda_j(1 + \beta^{-1}\lambda_j), \text{Cov}(R_{1,t}, R_{2,t}) = \beta^{-1}\lambda_1\lambda_2$ . Thus, bivariate negative binomial distribution is more flexible than bivariate Poisson due to overdispersion parameter  $\beta$ .

Assume now that the Poisson innovations  $R_{1,t}$  and  $R_{2,t}$  with parameters  $\lambda_1$  and  $\lambda_2$ , respectively, are linked by a copula with the dependence parameter  $\theta$ . Taking into account equality (14), we can estimate  $\theta$  by minimizing the sum of squared differences

$$S = \sum_{t=2}^N (R_{1,t}^{\text{CLS}} R_{2,t}^{\text{CLS}} - \gamma(\hat{\lambda}_1^{\text{CLS}}, \hat{\lambda}_2^{\text{CLS}}; \theta))^2, \tag{15}$$

where

$$R_{j,t}^{\text{CLS}} := X_{j,t} - \hat{\alpha}_j^{\text{CLS}} X_{j,t-1} - \hat{\lambda}_j^{\text{CLS}}, \quad j = 1, 2,$$

$$\gamma(\lambda_1, \lambda_2; \theta) := \text{Cov}(R_{1,t}, R_{2,t}) = \sum_{k,l=1}^{\infty} kl c(F_1(k; \lambda_1), F_2(l; \lambda_2); \theta) - \lambda_1 \lambda_2.$$

Here,  $c(F_1(k; \lambda_1), F_2(s; \lambda_2); \theta)$  is the joint pmf:

$$\begin{aligned} c(F_1(k; \lambda_1), F_2(l; \lambda_2); \theta) &= \mathbb{P}(R_{1,t} = k, R_{2,t} = l) \\ &= C(F_1(k; \lambda_1), F_2(s; \lambda_2); \theta) \\ &\quad - C(F_1(k-1; \lambda_1), F_2(l; \lambda_2); \theta) \\ &\quad - C(F_1(k; \lambda_1), F_2(l-1; \lambda_2); \theta) \\ &\quad + C(F_1(k-1; \lambda_1), F_2(l-1; \lambda_2); \theta), \quad k \geq 1, l \geq 1. \end{aligned} \tag{16}$$

Our estimation method is based on the approximation of covariance  $\gamma(\hat{\lambda}_1^{\text{CLS}}, \hat{\lambda}_2^{\text{CLS}}; \theta)$  by

$$\gamma^{(M_1, M_2)}(\hat{\lambda}_1^{\text{CLS}}, \hat{\lambda}_2^{\text{CLS}}; \theta) = \sum_{k=1}^{M_1} \sum_{l=1}^{M_2} kl c(F_1(k; \hat{\lambda}_1^{\text{CLS}}), F_2(l; \hat{\lambda}_2^{\text{CLS}}); \theta) - \hat{\lambda}_1^{\text{CLS}} \hat{\lambda}_2^{\text{CLS}}. \tag{17}$$

For example, if the marginals are Poisson with parameters  $\lambda_1 = \lambda_2 = 1$  and their joint distribution is given by the FGM copula in (8), then the covariance  $\gamma^{(M_1, M_2)}(1, 1; \theta)$  stops changing significantly after setting  $M_1 = M_2 = M = 8$ , regardless of the selected dependence parameter  $\theta$ . We used this approximation methodology when carrying out a Monte Carlo simulation in Section 4.4.

For the FGM copula, if we take the derivative of the sum

$$S^{(M_1, M_2)} = \sum_{t=2}^N (R_{1,t}^{\text{CLS}} R_{2,t}^{\text{CLS}} - \gamma^{(M_1, M_2)}(\hat{\lambda}_1^{\text{CLS}}, \hat{\lambda}_2^{\text{CLS}}; \theta))^2, \tag{18}$$

equate it to zero and use equation (17), we get

$$\hat{\theta}^{\text{FGM}} = \frac{\sum_{t=2}^N (X_{1,t} - \hat{\alpha}_1^{\text{CLS}} X_{1,t-1} - \hat{\lambda}_1^{\text{CLS}})(X_{2,t} - \hat{\alpha}_2^{\text{CLS}} X_{2,t-1} - \hat{\lambda}_2^{\text{CLS}})}{(N-1) \sum_{k=1}^{M_1} k(F_{1,k} \bar{F}_{1,k} - F_{1,k-1} \bar{F}_{1,k-1}) \sum_{l=1}^{M_2} l(F_{2,l} \bar{F}_{2,l} - F_{2,l-1} \bar{F}_{2,l-1})}, \tag{19}$$

where  $F_{j,k} := F_j(k; \hat{\lambda}_j^{\text{CLS}})$ ,  $\bar{F}_{j,k} := 1 - F_{j,k}$ ,  $j = 1, 2$ . The derivation of equation (19) is straightforward and thus omitted.

Depending on the selected copula family, calculation of (16) to get the analytical expression of the estimator  $\hat{\theta}$  may be difficult. However, we can use the function `optim` in the R statistical software to minimize (15). For other cases, where the marginal distribution has parameters other than expected value  $\lambda_j$ , equation (15) would need to be minimized by those additional parameters. For example, in the case of negative binomial marginals with corresponding mean  $\lambda_j$  and variance  $\sigma_j^2$ , i.e. when

$$\mathbb{P}(R_{j,t} = k) = \frac{\Gamma(k + \frac{\lambda_j^2}{\sigma_j^2 - \lambda_j})}{\Gamma(\frac{\lambda_j^2}{\sigma_j^2 - \lambda_j})k!} \left(\frac{\lambda_j}{\sigma_j^2}\right)^{\frac{\lambda_j^2}{\sigma_j^2 - \lambda_j}} \left(\frac{\sigma_j^2 - \lambda_j}{\sigma_j^2}\right)^k, \quad k = 0, 1, \dots, j = 1, 2,$$

the additional parameters are  $\sigma_1^2, \sigma_2^2$ , and the minimization problem becomes

$$S^{(M_1, M_2)} \longrightarrow \min_{\sigma_1^2, \sigma_2^2, \theta} .$$

#### 4.2 Conditional maximum likelihood estimation

BINAR(1) models can be estimated via conditional maximum likelihood (CML) (see [18] and [10]). The conditional distribution of the BINAR(1) process is:

$$\begin{aligned} \mathbb{P}(X_{1,t} = x_{1,t}, X_{2,t} = x_{2,t} | X_{1,t-1} = x_{1,t-1}, X_{2,t-1} = x_{2,t-1}) \\ = \mathbb{P}(\alpha_1 \circ x_{1,t-1} + R_{1,t} = x_{1,t}, \alpha_2 \circ x_{2,t-1} + R_{2,t} = x_{2,t}) \\ = \sum_{k=0}^{x_{1,t}} \sum_{l=0}^{x_{2,t}} \mathbb{P}(\alpha_1 \circ x_{1,t-1} = k) \mathbb{P}(\alpha_2 \circ x_{2,t-1} = l) \mathbb{P}(R_{1,t} = x_{1,t} - k, R_{2,t} = x_{2,t} - l). \end{aligned}$$

Here,  $\alpha_j \circ x$  is the sum of  $x$  independent Bernoulli trials. Hence,

$$\mathbb{P}(\alpha_j \circ x_{j,t-1} = k) = \binom{x_{j,t-1}}{k} \alpha_j^k (1 - \alpha_j)^{x_{j,t-1} - k}, \quad k = 0, \dots, x_{j,t-1}, j = 1, 2.$$

In the case of copula-based BINAR(1) model with Poisson marginals,

$$\mathbb{P}(R_{1,t} = x_{1,t} - k, R_{2,t} = x_{2,t} - l) = c(F_1(x_{1,t} - k, \lambda_1), F_2(x_{2,t} - l, \lambda_2); \theta).$$

Thus, we obtain

$$\begin{aligned} \mathbb{P}(X_{1,t} = x_{1,t}, X_{2,t} = x_{2,t} | X_{1,t-1} = x_{1,t-1}, X_{2,t-1} = x_{2,t-1}) \\ = \sum_{k=0}^{x_{1,t}} \sum_{l=0}^{x_{2,t}} \binom{x_{1,t-1}}{k} \alpha_1^k (1 - \alpha_1)^{x_{1,t-1} - k} \binom{x_{2,t-1}}{l} \alpha_2^l (1 - \alpha_2)^{x_{2,t-1} - l} \\ \times c(F_1(x_{1,t} - k, \lambda_1), F_2(x_{2,t} - l, \lambda_2); \theta) \end{aligned}$$

and the log conditional likelihood function, for estimating the marginal distribution parameters  $\lambda_1, \lambda_2$ , the probabilities of the Bernoulli trial successes  $\alpha_1, \alpha_2$  and the dependence parameter  $\theta$ , is

$$\ell(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta) = \sum_{t=2}^N \log \mathbb{P}(X_{1,t} = x_{1,t}, X_{2,t} = x_{2,t} | X_{1,t-1} = x_{1,t-1}, X_{2,t-1} = x_{2,t-1})$$

for some initial values  $x_{1,1}$  and  $x_{2,1}$ . In order to estimate the unknown parameters we maximize the log conditional likelihood:

$$\ell(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta) \longrightarrow \max_{\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta}. \quad (20)$$

Numerical maximization is straightforward with the `optim` function in the R statistical software.

As for the CLS estimator, in other cases, where the marginal distribution has parameters other than  $\lambda_j$ , equation (20) would need to be maximized by those additional parameters. The CML estimator is asymptotically normally distributed under standard regularity conditions and its variance matrix is the inverse of the Fisher information matrix [18].

### 4.3 Two-step estimation based on CLS and CML

Depending on the range of attainable values of the parameters and the sample size, CML maximization might take some time to compute. On the other hand, since CLS estimators of  $\alpha_j$  and  $\lambda_j$  are easily derived (compared to the CLS estimator of  $\theta$ , which depends on the copula pmf form and needs to be numerically maximized), we can substitute the parameters of the marginal distributions in eq. (20) with CLS estimates from equations (12) and (13). Then we will only need to maximize  $\ell$  with respect to a single dependence parameter  $\theta$  for the Poisson marginal distribution case.

Summarizing, the two-step approach to estimating unknown parameters is to find

$$(\hat{\alpha}_j^{\text{CLS}}, \hat{\lambda}_j^{\text{CLS}}) = \arg \min Q_j(\alpha_j, \lambda_j), \quad j = 1, 2,$$

and to take these values as given in the second step:

$$\hat{\theta}^{\text{CML}} = \arg \max \ell(\hat{\alpha}_1^{\text{CLS}}, \hat{\alpha}_2^{\text{CLS}}, \hat{\lambda}_1^{\text{CLS}}, \hat{\lambda}_2^{\text{CLS}}, \theta).$$

For other cases of marginal distribution, any additional parameters, other than  $\alpha_j$  and  $\lambda_j$  would be estimated in the second step.

### 4.4 Comparison of estimation methods via Monte Carlo simulation

We carried out a Monte Carlo simulation 1000 times to test the estimation methods with sample size 50 and 500. The generated model was a BINAR(1) with innovations joined by either the FGM, Frank or Clayton copula with Poisson marginal distributions, as well as with marginal distributions from different families: one is a

**Table 1.** Monte Carlo simulation results for a BINAR(1) model with Poisson innovations linked by the FGM, Frank or Clayton copula

Copula	Sample size	Parameter	True value	CLS		CML		Two-Step	
				MSE	Bias	MSE	Bias	MSE	Bias
FGM	N = 50	$\alpha_1$	0.6	0.01874	-0.05823	0.00887	-0.01789	-	-
		$\alpha_2$	0.4	0.02033	-0.05223	0.01639	-0.02751	-	-
		$\lambda_1$	1	0.12983	0.13325	0.06514	0.03366	-	-
		$\lambda_2$	2	0.25625	0.16029	0.19939	0.07597	-	-
		$\theta$	-0.5	<b>0.29789</b>	0.12568	0.33840	0.07568	0.3311	0.0876
	N = 500	$\alpha_1$	0.6	0.00147	-0.00432	0.00073	-0.00122	-	-
		$\alpha_2$	0.4	0.00184	-0.00505	0.00129	-0.00157	-	-
		$\lambda_1$	1	0.01012	0.00968	0.00556	0.00215	-	-
		$\lambda_2$	2	0.02413	0.01843	0.01763	0.00678	-	-
		$\theta$	-0.5	0.04679	0.00668	0.04271	-0.00700	<b>0.04265</b>	-0.00443
Frank	N = 50	$\alpha_1$	0.6	0.02023	-0.06039	0.00950	-0.01965	-	-
		$\alpha_2$	0.4	0.02005	-0.05251	0.01630	-0.02858	-	-
		$\lambda_1$	1	0.13562	0.13536	0.06740	0.03625	-	-
		$\lambda_2$	2	0.25687	0.16392	0.19975	0.08291	-	-
		$\theta$	-1	<b>1.83454</b>	0.12394	2.05786	0.00860	1.97515	0.04216
	N = 500	$\alpha_1$	0.6	0.00153	-0.00595	0.00075	-0.00249	-	-
		$\alpha_2$	0.4	0.00181	-0.00582	0.00129	-0.00132	-	-
		$\lambda_1$	1	0.01033	0.01269	0.00550	0.00421	-	-
		$\lambda_2$	2	0.02442	0.02129	0.01785	0.00629	-	-
		$\theta$	-1	0.22084	0.01746	0.20138	-0.01779	<b>0.20070</b>	-0.01342
Clayton	N = 50	$\alpha_1$	0.6	0.01826	-0.05489	0.00799	-0.013295	-	-
		$\alpha_2$	0.4	0.01976	-0.05057	0.01585	-0.02427	-	-
		$\lambda_1$	1	0.12679	0.12104	0.06080	0.01743	-	-
		$\lambda_2$	2	0.25725	0.15704	0.19934	0.06499	-	-
		$\theta$	1	0.71845	0.02621	0.72581	0.22628	<b>0.62372</b>	0.13283
	N = 500	$\alpha_1$	0.6	0.00146	-0.00518	0.00070	0.00016	-	-
		$\alpha_2$	0.4	0.00189	-0.00350	0.00120	-0.00049	-	-
		$\lambda_1$	1	0.00973	0.01137	0.00513	-0.00150	-	-
		$\lambda_2$	2	0.02447	0.01113	0.01707	0.00065	-	-
		$\theta$	1	0.11578	0.03556	0.05864	0.04250	<b>0.03199</b>	-0.01342

Poisson distribution and the other is a negative binomial one. Note that for the two-step method only the estimates of  $\theta$  and  $\sigma_2^2$  are included because estimated values of  $\alpha_1^{CLS}, \alpha_2^{CLS}, \lambda_1^{CLS}, \lambda_2^{CLS}$  are used in order to estimate the remaining parameters via CML.

The results for the Poisson marginal distribution case are provided in Table 1. The results for the case when one innovation follows a Poisson distribution and the other follows a negative binomial one are provided in Table 2. The lowest MSE values of  $\hat{\theta}$  are highlighted in bold. It is worth noting that CML estimation via numerical maximization depends heavily on the initial parameter values. If the initial values are selected too low or too high from the actual value, then the global maximum may not be found. In order to overcome this, we have selected the starting values equal to the CLS parameter estimates.

As can be seen in Table 1, the estimated values of  $\alpha_j$  and  $\lambda_j, j = 1, 2$ , have a smaller bias and MSE when parameters are estimated via CML. On the other hand, estimation of  $\theta$  via CLS exhibits a smaller MSE in the Frank copula case for smaller samples. For larger samples, the estimates of  $\theta$  via the Two-step estimation method

**Table 2.** Monte Carlo simulation results for a BINAR(1) model with one innovation following a Poisson distribution and the other – a negative binomial one, where both innovations are linked by the FGM, Frank or Clayton copula

Copula	Sample size	Parameter	True value	CLS		CML		Two-Step	
				MSE	Bias	MSE	Bias	MSE	Bias
FGM	N = 50	$\alpha_1$	0.6	0.01895	-0.05858	0.00845	-0.01513	-	-
		$\alpha_2$	0.4	0.01936	-0.04902	0.00767	-0.01953	-	-
		$\lambda_1$	1	0.12940	0.12812	0.05424	0.01879	-	-
		$\lambda_2$	2	0.39724	0.15151	0.24138	0.04833	-	-
		$\theta$	-0.5	0.31467	0.14070	<b>0.29415</b>	0.06674	0.29949	0.09693
		$\sigma_2^2$	9	27.87327	1.15731	15.12863	-0.14888	21.68229	0.72326
		$\sigma_2^2$	9	27.87327	1.15731	15.12863	-0.14888	21.68229	0.72326
	N = 500	$\alpha_1$	0.6	0.00156	-0.00695	0.00076	-0.00153	-	-
		$\alpha_2$	0.4	0.00194	-0.00373	0.00053	0.00016	-	-
		$\lambda_1$	1	0.01041	0.01201	0.00543	0.00290	-	-
		$\lambda_2$	2	0.03882	0.01843	0.02362	-0.00057	-	-
		$\theta$	-0.5	0.06670	-0.02014	<b>0.04298</b>	-0.00268	0.04313	0.00562
		$\sigma_2^2$	9	6.24237	-1.99232	1.81265	0.00611	1.85222	-0.03506
		$\sigma_2^2$	9	6.24237	-1.99232	1.81265	0.00611	1.85222	-0.03506
Frank	N = 50	$\alpha_1$	0.6	0.02049	-0.06064	0.00912	-0.01594	-	-
		$\alpha_2$	0.4	0.01951	-0.04936	0.00772	-0.02070	-	-
		$\lambda_1$	1	0.13769	0.13467	0.05748	0.02280	-	-
		$\lambda_2$	2	0.40626	0.15408	0.23717	0.05534	-	-
		$\theta$	-1	1.81788	0.12516	1.75638	-0.01239	<b>1.68019</b>	0.06211
		$\sigma_2^2$	9	25.10400	0.49423	14.86812	-0.10034	21.92090	0.74026
		$\sigma_2^2$	9	25.10400	0.49423	14.86812	-0.10034	21.92090	0.74026
	N = 500	$\alpha_1$	0.6	0.00161	-0.00702	0.00075	-0.00239	-	-
		$\alpha_2$	0.4	0.00187	-0.00364	0.00050	-0.00046	-	-
		$\lambda_1$	1	0.01093	0.01652	0.00562	0.00501	-	-
		$\lambda_2$	2	0.03728	0.01217	0.02335	0.00203	-	-
		$\theta$	-1	0.31942	-0.05593	<b>0.18960</b>	-0.01481	0.1902	-0.0079
		$\sigma_2^2$	9	4.82620	-1.75765	1.83082	0.02144	1.85852	-0.02690
		$\sigma_2^2$	9	4.82620	-1.75765	1.83082	0.02144	1.85852	-0.02690
Clayton	N = 50	$\alpha_1$	0.6	0.01987	-0.06159	0.00903	-0.01671	-	-
		$\alpha_2$	0.4	0.01879	-0.04928	0.00632	-0.01644	-	-
		$\lambda_1$	1	0.13479	0.14072	0.06096	0.03052	-	-
		$\lambda_2$	2	0.40675	0.14807	0.23171	0.02871	-	-
		$\theta$	1	0.78497	0.07464	0.67837	0.21235	<b>0.57454</b>	0.10972
		$\sigma_2^2$	9	24.40051	0.17321	15.29879	-0.08379	23.73506	0.73754
		$\sigma_2^2$	9	24.40051	0.17321	15.29879	-0.08379	23.73506	0.73754
	N = 500	$\alpha_1$	0.6	0.00153	-0.00722	0.00075	-0.00197	-	-
		$\alpha_2$	0.4	0.00196	-0.00385	0.00047	-0.00083	-	-
		$\lambda_1$	1	0.01036	0.01745	0.00517	0.00409	-	-
		$\lambda_2$	2	0.03999	0.01227	0.02304	0.00110	-	-
		$\theta$	1	0.09927	0.04408	<b>0.05557</b>	0.03556	0.05559	0.02310
		$\sigma_2^2$	9	2.95995	-0.68733	1.79836	0.01348	1.87740	-0.02407
		$\sigma_2^2$	9	2.95995	-0.68733	1.79836	0.01348	1.87740	-0.02407

are very close to the CML estimates in terms of MSE and bias, and are closer to the true parameter values than the CLS estimates. Furthermore, since in the Two-step estimation numerical maximization is only carried out via a single parameter  $\theta$ , the initial parameter values have less effect on the numerical maximization.

Table 2 demonstrates the estimation results when one innovation has a Poisson distribution and the other has a negative binomial one. With the inclusion of an additional variance parameter, the CLS estimation methods exhibit larger MSE and bias than the CML and Two-step estimation methods, for both the dependence and variance parameter estimates. Furthermore, the MSE of  $\hat{\sigma}_2^2$  is smallest when the CML estimation method is used. On the other hand, both the Two-step and CML estima-

tion methods produce similar estimates of  $\theta$  in terms of MSE, regardless of sample size and copula function.

We can conclude that it is possible to accurately estimate the dependence parameter via CML using the CLS estimates of  $\hat{\alpha}_j$  and  $\hat{\lambda}_j$ . The resulting  $\hat{\theta}$  will be closer to the actual value of  $\theta$  than  $\hat{\theta}^{\text{CLS}}$  and will not differ much from  $\hat{\theta}^{\text{CML}}$ . Additional inference on the bias of the estimates can be found in Appendix A.

## 5 Application to loan default data

In this section we estimate a BINAR(1) model with the joint innovation distribution modelled by a copula cdf for empirical data. The data set consists of loan data which includes loans that have defaulted and loans that were repaid without missing any payments (non-defaulted loans). We will analyse and model the dependence between defaulted and non-defaulted loans as well as the presence of autocorrelation.

### 5.1 Loan default data

The data sample used is from Bondora, the Estonian peer-to-peer lending company. In November of 2014 Bondora introduced a loan rating system which assigns loans to different groups, based on their risk level. There are 8 groups ranging from the lowest risk group, ‘AA’, to the highest risk group, ‘HR’. However, the loan rating system could not be applied to most older loans due to a lack of data needed for Bondora’s rating model. Although Bondora issues loans in 4 different countries: Estonia, Finland, Slovakia and Spain, we will only focus on the loans issued in Spain. Since a new rating model indicates new rules for accepting or rejecting loans, we have selected the data sample from 21 October 2013, because from that date forward all loans had a rating assigned to them, to 1 January 2016. The time series are displayed in Figure 1. We are analysing data consisting of 115 weekly records.

- ‘CompletedLoans’ – the amount of non-defaulted loans issued per week which are repaid and have never defaulted (a loan that is 60 or more days overdue is considered defaulted);
- ‘DefaultedLoans’ – the amount of defaulted loans issued per week.

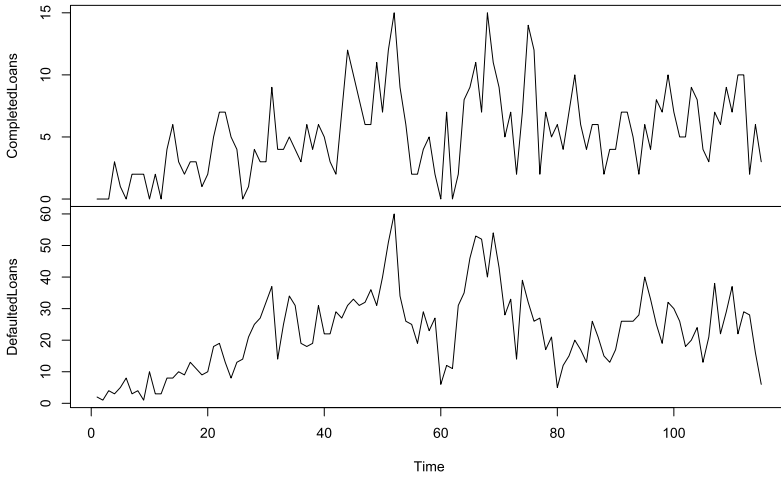
The loan statistics are provided in Table 3:

**Table 3.** Summary statistics of the weekly data of defaulted and non-defaulted loans issued in Spain

	min	max	mean	variance
DefaultedLoans	1.00	60.00	22.60	158.66
CompletedLoans	0.00	15.00	5.30	11.67

The mean, minimum, maximum and variance is higher for defaulted loans than for non-defaulted loans. As can be seen from Figure 2, the numbers of defaulted and non-defaulted loans might be correlated since they both exhibit increase and decrease periods at the same times.

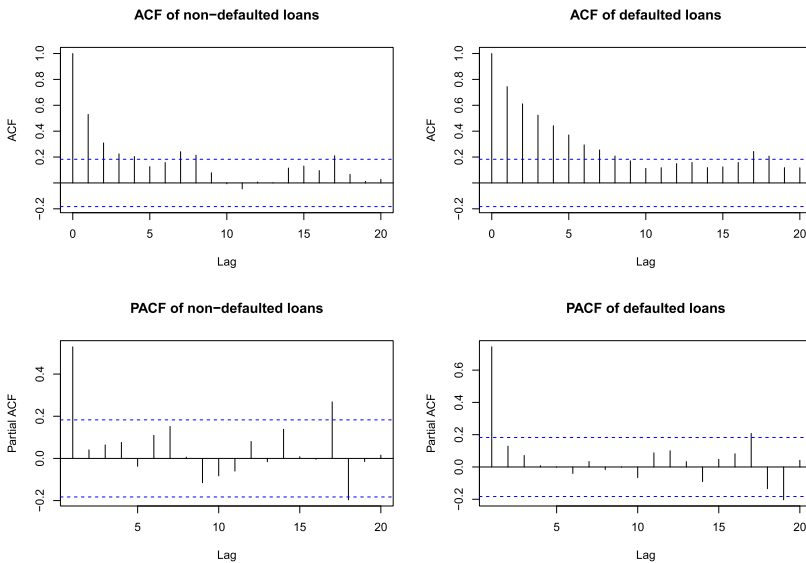
The correlation between the two time series is 0.6684. We also note that the mean and variance are lower in the beginning of the time series. This feature could be



**Fig. 1.** Bondora loan data: non-defaulted and defaulted loans by their issue date

due to various reasons: the effect of the new loan rating system, which was officially implemented in December of 2014, the effect of advertising or the fact that the amount of loans, issued to people living outside of Estonia, increased. The analysis of the significance of these effects is left for future research.

The sample autocorrelation (AC) function and the partial autocorrelation (PAC) function are displayed in Figure 2. We can see that the AC function is decaying over time and the PAC function has a significant first lag which indicates that the non-negative integer-valued time series could be autocorrelated.



**Fig. 2.** AC function and PAC function plots of Bondora loan data



**Table 4.** Parameter estimates for BINAR(1) model via the Two-step estimation method: parameter CLS estimates from the first step with standard errors for the Poisson marginal distribution case in parenthesis

$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
0.53134	0.75581	2.52174	5.58940
(0.08151)	(0.06163)	(0.45012)	(1.41490)

In order to analyse if the amount of defaulted loans depends on the amount of non-defaulted loans on the same week, we will consider a BINAR(1) model with different copulas for the innovations. For the marginal distributions of the innovations we will consider the Poisson distribution as well as the negative binomial one. Our focus is the estimation of the dependence parameter, and we will use the Two-step estimation method, based on the Monte Carlo simulation results presented in Section 4.

### 5.2 Estimated models

We estimated a number of BINAR(1) models with different distributions of innovations which include combinations of:

- different copula functions: FGM, Frank or Clayton;
- different combinations of the Poisson and negative binomial distributions: both marginals are Poisson, both marginals are negative binomial, or a mix of both.

In the first step of the Two-step method, we estimated  $\hat{\alpha}_1$  and  $\hat{\lambda}_1$  for non-defaulted loans, and  $\hat{\alpha}_2$  and  $\hat{\lambda}_2$  for defaulted loans via CLS. The results are provided in Table 4 with standard errors for the Poisson case in parenthesis:

Because the CLS estimation of parameters  $\alpha_j$  and  $\lambda_j$ ,  $j = 1, 2$ , does not depend on the selected copula and the marginal distribution family, these parameters will remain the same for each of the different distribution combinations for innovations. We can see that defaulted loans exhibit a higher degree of autocorrelation than non-defaulted loans do, due to a larger value of  $\hat{\alpha}_2$ . The innovation mean parameter for defaulted loans is also higher, what indicates that random shocks have a larger effect on the number of defaulted loans.

The parameter estimation results from the second-step are provided in Table 5 with standard errors in parenthesis.  $\hat{\sigma}_1^2$  is the innovation variance estimate of non-defaulted loans and  $\hat{\sigma}_2^2$  is the innovation variance estimate of defaulted loans. According to [16], the observed Fisher information is the negative Hessian matrix, evaluated at the maximum likelihood estimator (MLE). The asymptotic standard errors reported in Table 5 are derived under the assumption that  $\alpha_j$  and  $\lambda_j$ ,  $j = 1, 2$ , are known, ignoring that the true values are substituted in the second step with their CLS estimates.

From the results in Table 5 we see that, according to the Akaike information criterion (AIC) and log-likelihood values, in most cases the FGM copula most accurately describes the relationship between the innovations of defaulted and non-defaulted loans, with the Frank copula being very close in terms of the AIC value. The Clayton copula is the least accurate in describing the innovation joint distribution, when compared to the FGM and Frank copula cases, which indicates that defaulted and non-defaulted loans do not exhibit strong left tail dependence.

**Table 5.** Parameter estimates for BINAR(1) model via Two-step estimation method: parameter CML estimates from the second-step for different innovation marginal and joint distribution combinations with standard errors in parenthesis, derived under the assumption that the values  $\hat{\lambda}_j$  and  $\hat{\alpha}_j$ ,  $j = 1, 2$ , from the first step are true

Marginals	Copula	$\hat{\theta}$	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	AIC	Log-likelihood
Both Poisson	FGM	0.89270 (0.18671)	–	–	1763.48096	–880.74048
	Frank	2.38484 (0.53367)	–	–	1760.15692	–879.07846
	Clayton	0.39357 (0.11697)	–	–	1761.12369	–879.56185
Negative binomial and Poisson	FGM	1.00000 (0.22914)	6.46907 (1.01114)	–	1731.57339	–863.78670
	Frank	2.14329 (0.45100)	6.10242 (1.15914)	–	1731.95241	–863.97620
	Clayton	0.34540 (0.12859)	5.73731 (0.52831)	–	1736.47641	–866.23821
Poisson and negative binomial	FGM	1.00000 (0.26357)	–	44.83107 (7.37423)	1498.29563	–747.14782
	Frank	2.01486 (0.61734)	–	44.10555 (7.33169)	1498.81039	–747.40519
	Clayton	0.38310 (0.17376)	–	43.42739 (7.29842)	1503.55388	–749.77694
Both negative binomial	FGM	1.00000 (0.31675)	6.55810 (1.24032)	45.36834 (7.55217)	<b>1466.15418</b>	<b>–730.07709</b>
	Frank	2.21356 (0.68192)	6.58754 (1.26126)	45.42601 (7.57743)	1466.97947	–730.48973
	Clayton	0.55939 (0.24652)	6.64478 (1.25833)	45.78307 (7.66324)	1470.73515	–732.36758

Since the summary statistics of the data sample showed that the variance of the data is larger than the mean, a negative binomial marginal distribution may provide a better fit. Additionally, because copulas can link different marginal distributions, it is interesting to see if copulas with different discrete marginal distributions would also improve the model fit. BINAR(1) models where non-defaulted loan innovations are modelled with negative binomial distributions and defaulted loan innovations are modelled with Poisson marginal distributions, and vice versa, were estimated. In general, changing one of the marginal distributions to a negative binomial provides a better fit in terms of AIC than the Poisson marginal distribution case. However, the smallest AIC value is achieved when both marginal distributions are modelled with negative binomial distributions, linked via the FGM copula. Furthermore, the estimated innovation variance,  $\hat{\sigma}_2^2$ , is much larger for defaulted loans, and this is similar to what we observed from the defaulted loan data summary statistics.

Overall, both Frank and FGM copulas provide similar fit in terms of log-likelihood, regardless of the selected marginal distributions. We note, however, that for some FGM copula cases, the estimated value of parameter  $\theta$  is equal to the maximal attainable value 1. Based on copula descriptions from Section 3, the FGM copula is used to model weak dependence. Given a larger sample size, the Frank copula might be more appropriate because it can capture a stronger dependence than the FGM copula can do. The negative binomial marginal distribution case  $\hat{\theta} \approx 2.21356$  for the Frank cop-

ula indicates that there is a positive dependence between defaulted and non-defaulted loans, just as in the FGM copula case.

## 6 Conclusions

The analysis via Monte Carlo simulations of different estimation methods shows that, although the estimates of BINAR(1) parameters via CML has the smallest MSE and bias, estimates of the dependence parameter has smaller differences of MSE and bias than for other estimation methods, indicating that estimations of the dependence parameter via different methods do not exhibit large differences. While CML estimates exhibit the smallest MSE, their calculation via numerical optimization relies on the selection of the initial parameter values. These values can be selected via CLS estimation.

An empirical application of BINAR models for loan data shows that, regardless of the selected marginal distributions, the FGM copula provides the best model fit in almost all cases. Models with the Frank copula are similar to FGM copula models in terms of AIC values. For some of these cases, the estimated FGM copula dependence parameter value was equal to the maximum that can be attained by an FGM copula. In such cases, a larger sample size could help to determine whether the FGM or Frank copula is more appropriate to model the dependence between amounts of defaulted and non-defaulted loans.

Although selecting marginal distributions from different families (Poisson or negative binomial) provided better models than those with only Poisson marginal distributions, the models with both marginal distributions modelled via negative binomial distributions provide the smallest AIC values which reflects overdispersion in amounts of both defaulted and non-defaulted loans. The FGM copula, which provides the best model fit, models variables which exhibit weak dependence. Furthermore, the estimated copula dependence parameter indicates that the dependence between amounts of defaulted and non-defaulted loans is positive.

Finally, one can apply some other copulas in order to analyse whether the loan data exhibits different forms of dependence from the ones discussed in this paper. Lastly, the approach can be extended by analysing the presence of structural changes within the data, or checking the presence of seasonality as well as extending the BINAR(1) model with copula joined innovations to account for the past values of other time series rather than only itself.

## A Appendix

Let our Monte Carlo simulation data be  $X_{j,1}^{(i)}, \dots, X_{j,N}^{(i)}$  for simulated sample  $i = 1, \dots, M$  and  $j = 1, 2$ . Let  $\eta \in \{\alpha_1, \alpha_2, \lambda_1, \lambda_2, \theta, \sigma_2^2\}$  and let  $\widehat{\eta}^{(i)}$  be either a CLS, CML or Two-step estimate of the true parameter value  $\eta$  for the simulated sample  $i$ .

The mean squared error and the bias are calculated as follows:

$$\text{MSE}(\widehat{\eta}) = \frac{1}{M} \sum_{i=1}^M (\widehat{\eta}^{(i)} - \eta)^2,$$

$$\text{Bias}(\widehat{\eta}) = \frac{1}{M} \sum_{i=1}^M (\widehat{\eta}^{(i)} - \eta).$$

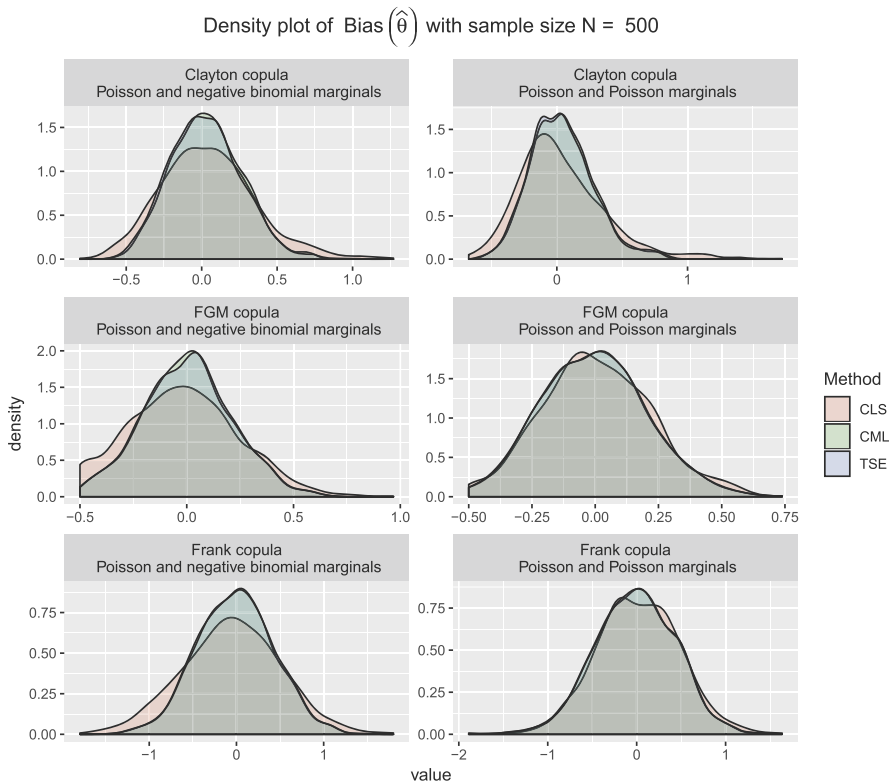
Calculating the per-sample bias for each simulated sample  $i$  would also allow us to calculate the sample variance of biases  $\text{Bias}(\widehat{\eta}^{(i)}) = \widehat{\eta}^{(i)} - \eta, i = 1, \dots, M$ :

$$\widehat{\text{Var}}(\text{Bias}(\widehat{\eta})) = \frac{1}{M - 1} \sum_{i=1}^M [\text{Bias}(\widehat{\eta}^{(i)}) - \text{Bias}(\widehat{\eta})]^2,$$

which we can use to calculate the standard error of the bias.

**Table 6.** Standard errors of the bias of the estimated parameters from the Monte Carlo simulation

Copula	Sample size	Parameter	True value	CLS		CML		Two-Step		
				P-P	P-NB	P-P	P-NB	P-P	P-NB	
FGM	N = 50	$\alpha_1$	0.6	0.12396	0.12465	0.09252	0.09073	–	–	
		$\alpha_2$	0.4	0.13274	0.13029	0.12510	0.08541	–	–	
		$\lambda_1$	1	0.33494	0.33631	0.25311	0.23225	–	–	
		$\lambda_2$	2	0.48040	0.61210	0.44024	0.48916	–	–	
		$\theta$	-0.5	0.53139	0.54330	0.57707	0.53850	0.56899	0.53887	
		$\sigma_2^2$	9	–	5.15368	–	3.88865	–	4.60221	
	N = 500	$\alpha_1$	0.6	0.03813	0.03893	0.02706	0.02745	–	–	
		$\alpha_2$	0.4	0.04258	0.04392	0.03585	0.02306	–	–	
		$\lambda_1$	1	0.10018	0.10076	0.07455	0.07367	–	–	
		$\lambda_2$	2	0.15433	0.19676	0.13266	0.15377	–	–	
		$\theta$	-0.5	0.21631	0.25760	0.20666	0.20741	0.20657	0.20770	
		$\sigma_2^2$	9	–	1.50841	–	1.34701	–	1.36119	
	Frank	N = 50	$\alpha_1$	0.6	0.12882	0.12975	0.09552	0.09420	–	–
			$\alpha_2$	0.4	0.13158	0.13073	0.12448	0.08543	–	–
$\lambda_1$			1	0.34266	0.34594	0.25719	0.23879	–	–	
$\lambda_2$			2	0.47982	0.61879	0.43939	0.48409	–	–	
$\theta$			-1	1.34944	1.34314	1.43522	1.32589	1.40547	1.29538	
$\sigma_2^2$			9	–	4.98845	–	3.85654	–	4.62540	
N = 500		$\alpha_1$	0.6	0.03862	0.03951	0.02734	0.02727	–	–	
		$\alpha_2$	0.4	0.04212	0.04312	0.03591	0.02240	–	–	
		$\lambda_1$	1	0.10091	0.10329	0.07409	0.07481	–	–	
		$\lambda_2$	2	0.15490	0.19278	0.13351	0.15287	–	–	
		$\theta$	-1	0.46985	0.56268	0.44862	0.43540	0.44802	0.43627	
		$\sigma_2^2$	9	–	1.31856	–	1.35359	–	1.36369	
Clayton		N = 50	$\alpha_1$	0.6	0.12352	0.12684	0.08846	0.09360	–	–
			$\alpha_2$	0.4	0.13123	0.12798	0.12361	0.07779	–	–
	$\lambda_1$		1	0.33505	0.33926	0.24609	0.24514	–	–	
	$\lambda_2$		2	0.48252	0.62066	0.44194	0.48075	–	–	
	$\theta$		1	0.84763	0.88328	0.82176	0.79618	0.77890	0.75037	
	$\sigma_2^2$		9	–	4.93912	–	3.91243	–	4.81812	
	N = 500	$\alpha_1$	0.6	0.03782	0.03850	0.02641	0.02742	–	–	
		$\alpha_2$	0.4	0.04337	0.04410	0.03468	0.02176	–	–	
		$\lambda_1$	1	0.09804	0.10033	0.07162	0.07180	–	–	
		$\lambda_2$	2	0.15612	0.19969	0.13071	0.15185	–	–	
		$\theta$	1	0.33857	0.31212	0.23852	0.23316	0.23717	0.23476	
		$\sigma_2^2$	9	–	1.57798	–	1.34163	–	1.37066	



**Fig. 3.** Kernel density estimate for the bias of the dependence parameter estimates in the Monte Carlo simulation

The standard errors of the parameter bias of the Monte Carlo simulation are presented in Table 6. The columns labelled ‘P-P’ indicate the cases where both innovations have Poisson marginal distributions, while columns labelled ‘P-NB’ is for the cases where one innovation component follows the Poisson distribution and the other follows a negative binomial one. The kernel density estimate for the bias of the dependence parameter estimate,  $\hat{\theta}$ , is presented in Figure 3 for the Monte Carlo simulation cases, where the sample size was 500.

The results in Table 6 are in line with the conclusions presented in Section 4.4 – for  $\hat{\alpha}_j$ ,  $\hat{\lambda}_j$ ,  $j = 1, 2$ , and  $\hat{\sigma}_2^2$  the standard error of the bias is smaller for CML than for CLS. On the other hand,  $\hat{\theta}$  has a similar standard error of the bias for CML and Two-step estimation methods. From Figure 3 we see that the CML and Two-step estimates of the dependence parameter  $\theta$  are similar to each other and have a lower standard error of the bias than the CLS estimate.

**Acknowledgement**

The authors would like to thank the anonymous referee for his/her feedback and constructive insights, which helped to improve this paper.

## References

- [1] Al-Osh, M., Alzaid, A.: First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* **8**, 261–275 (1987) [MR0903755](#). <https://doi.org/10.1111/j.1467-9892.1987.tb00438.x>
- [2] Barczy, M., Ispány, M., Pap, G., Scotto, M., Silva, M.E.: Innovational outliers in INAR(1) models. *Commun. Stat., Theory Methods* **39**(18), 3343–3362 (2010) [MR2747588](#). <https://doi.org/10.1080/03610920903259831>
- [3] Brigo, D., Pallavicini, A., Torresetti, R.: *Credit Models and the Crisis: A Journey Into CDOs, Copulas, Correlations and Dynamic Models*. Wiley, United Kingdom (2010)
- [4] Cherubini, U., Mulinacci, S., Gobbi, F., Romagnoli, S.: *Dynamic Copula Methods in Finance*. Wiley, United Kingdom (2011)
- [5] Crook, J., Moreira, F.: Checking for asymmetric default dependence in a credit card portfolio: A copula approach. *J. Empir. Finance* **18**, 728–742 (2011)
- [6] Fenech, J.P., Vosgha, H., Shafik, S.: Loan default correlation using an Archimedean copula approach: A case for recalibration. *Econ. Model.* **47**, 340–354 (2015)
- [7] Freeland, R.K., McCabe, B.: Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Stat. Probab. Lett.* **73**(2), 147–153 (2005) [MR2159250](#). <https://doi.org/10.1016/j.spl.2005.03.006>
- [8] Genest, C., Nešlehová, J.: A primer on copulas for count data. *ASTIN Bull.* **37**(2), 475–515 (2007) [MR2422797](#). <https://doi.org/10.2143/AST.37.2.2024077>
- [9] Joe, H.: *Dependence Modeling with Copulas*. Chapman & Hall/CRC Monographs on Statistics and Applied probability 134 (2015) [MR3328438](#)
- [10] Karlis, D., Pedeli, X.: Flexible bivariate INAR(1) processes using copulas. *Commun. Stat., Theory Methods* **42**, 723–740 (2013) [MR3211946](#). <https://doi.org/10.1080/03610926.2012.754466>
- [11] Kedem, B., Fokianos, K.: *Regression Models for Time Series Analysis*. Wiley-Interscience, New Jersey (2002) [MR1933755](#). <https://doi.org/10.1002/0471266981>
- [12] Latour, A.: The multivariate GINAR(p) process. *Adv. Appl. Probab.* **29**(1), 228–248 (1997) [MR1432938](#). <https://doi.org/10.2307/1427868>
- [13] Latour, A.: Existence and stochastic structure of a non-negative integer-valued autoregressive process. *J. Time Ser. Anal.* **19**(4), 439–455 (1998) [MR1652193](#). <https://doi.org/10.1111/1467-9892.00102>
- [14] Manstavičius, M., Leipus, R.: Bounds for the Clayton copula. *Nonlinear Anal., Model. Control* **22**, 248–260 (2017) [MR3608075](#). <https://doi.org/10.15388/na.2017.2.7>
- [15] Nelsen, R.: *An Introduction to Copulas*, 2nd Edition. Springer (2006) [MR2197664](#). <https://doi.org/10.1007/s11229-005-3715-x>
- [16] Pawitan, Y.: *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press, New York (2001) [MR3668697](#). <https://doi.org/10.1080/00031305.2016.1202140>
- [17] Pedeli, X.: *Modelling multivariate time series for count data*. PhD thesis, Athens University of Economics and Business (2011)
- [18] Pedeli, X., Karlis, D.: A bivariate INAR(1) process with application. *Stat. Model., Int. J.* **11**(4), 325–349 (2011) [MR2906704](#). <https://doi.org/10.1177/1471082X1001100403>
- [19] Silva, I.M.M.: *Contributions to the analysis of discrete-valued time series*. PhD thesis, University of Porto (2005)

- [20] Sklar, M.: Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **8**, 229–231 (1959) [MR0125600](#)
- [21] Trivedi, P.K., Zimmer, D.M.: Copula modelling: An introduction for practitioners. *Found. Trends Econom.* **1**(1), 1–111 (2007)