VILNIAUS UNIVERSITETAS

MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro darbas

Dvimačių sveikareikšmių sezoninių laiko eilučių modeliai: autovarijų skaičiaus atvejis

Modelling Bivariate Integer Valued Time Series with Seasonality: Evidence on Car Accident Data

Aušrinė Rinkevičiūtė

VILNIUS 2020

MATEMATIKOS IR INFORMATIKOS FAKULTETAS

STATISTINĖS ANALIZĖS KATEDRA

Darbo vadovas Prof. Habil. Dr. Remigijus Leipus

Darbo recenzentas Prof. Dr. Vytautas Kazakevičius

Darbas apgintas ______________

Darbas įvertintas ______________

Registravimo Nr. ______________

Contents

Dvimačių sveikareikšmių sezoninių laiko eilučių modeliai: autovarijų skaičiaus atvejis

Santrauka

Nors sveikareikšmių laiko eilučių modeliai yra plačiai nagrinėjami autorių, analizė kaip gali būti modeliuojamos sveikareikšmės sezoninės laiko eilutės yra ribota. Darbas yra skirtas dvimačių sveikareikšmių sezoninių laiko eilučių modelių nagrinėjimui. Darbe yra suformuluojami trys modeliai tokiems duomenims: BINAR$(1)_s$, BINGARCH$_s$ and TV-BINAR$(1)_s$. Aprašyti vertinimo metodai yra patikrinami naudojant simuliuotus duomenis. Empirinėje dalyje modeliai yra pritaikomi autoavarijų skaičiaus Lietuvoje per mėnesį duomenims, kai avarijas sukėlė blaivūs arba neblaivūs vairuotojai.

Raktiniai žodžiai : Sezoniškumas, dvimatis INAR, dvimatis INGARCH, laike kintantis BINAR, dvimatis Puasono skirstinys, avarijų skaičius

Modelling Bivariate Integer Valued Time Series with Seasonality: Evidence on Car Accident Data

Abstract

Although models for time series of counts are currently intensively studied by a number of researchers, analysis on how integer-valued time series with exhibited seasonality can be modelled is still limited. This thesis is focused on formulating models for bivariate integer-valued time series with exhibited seasonality. Hence, three models suitable to such are formulated: BINAR$(1)_s$, BINGARCH$_s$ and TV-BINAR$(1)_s$. The considered estimation methods are tested on the simulated data. All three models are applied on the bivariate car accident data. Time series consist of number of car accidents caused by the alcohol intoxicated drivers and a number of car accidents caused by the sober drivers in Lithuania per month.

Key words : Seasonality, Bivariate INAR, Bivariate INGARCH, Time Variant BINAR, bivariate Poisson distribution, number of car accidents

1 Introduction

Various processes in a number of disciplines are recorded as series of counts. Integer-valued data can be found in areas such as economics, medicine, biology, criminology. For such data the usual models (e.g. ARMA) are not suitable as they usually consider continuous distributions. Models for time series of counts are currently intensively studied by a number of researchers. Most studies consider binomial thinning operator based integer-valued AR model (INAR), see Al-Osh and Alzaid (1987), Jin-Guan and Yuan (1991), Scotto et al. (2015). Although most articles consider Poisson distribution for the marginals, model has also been analysed in terms of negative binomial distribution (see Ristić et al. (2012)) or geometric distribution (see Ristić et al. (2009)). Similarly to INAR, INMA model has been analysed by Al-Osh and Alzaid (1988). Although most authors consider univariate models, based on the fact that many real-life processes has a bivariate structure, Pedeli and Karlis (2011) formulated a bivariate INAR type model. Although bivariate distributions are often used for the bivariate models, copula based models are also considered for the matter (see Buteikis and Leipus (2019).

In addition to thinning operator based INAR type models, observation based integer-valued GARCH models are considered (also called Poisson regression). Bivariate Poisson INGARCH(p, q) model was constructed by Liu (2012) with proofs of the stationarity and ergodicity under certain conditions.

Having the above, analysis on how integer-valued time series with exhibited seasonality can be modelled is still limited. This thesis is focused on the question how bivariate integer-valued time series with expressed seasonality can be modelled. The aim of the thesis is to adapt the models for integer-valued time series to the process with seasonality. To do so we will formulate INAR and INGARCH type models suitable for bivariate seasonal data linked by the bivariate Poisson distribution and define new model mixing properties of the two above mentioned models (TV-BINAR(1)$_s$). To test the efficiency of chosen estimation methods for the models we will use the simulated data. Finally we will estimate models on real-life car accident data.

2 The bivariate seasonal models

In this thesis we will define 3 models that can be used for the bivariate integer-valued time series with seasonality. We will also investigate certain properties of the models, describe parameter estimation methods and test them on simulated data to test the efficiency of the described estimators.

The bivariate distribution for count data considered in the upcoming sections for the estimation of models is a bivariate Poisson distribution with parameters λ_1, λ_2 and ϕ: $BP(\lambda_1, \lambda_2, \phi)$ and probability mass function (pmf):

$$\begin{aligned} f(k,l) = P(Z_1 = k, Z_2 = l) =& e^{-(\lambda_1-\lambda_2-\phi)} \frac{(\lambda_1-\phi)^k}{k!} \frac{(\lambda_2-\phi)^l}{l!} \\ & \times \sum_{m=0}^{\min(k,l)} \binom{k}{m} \binom{l}{m} m! (\frac{\phi}{(\lambda_1-\phi)(\lambda_2-\phi)})^m. \end{aligned} \tag{2.1}$$

Bivariate Poisson distribution can also be understood as a combination of three independent random sequences $X_1 \sim P(\lambda_1)$, $X_2 \sim P(\lambda_2)$ and $X_3 \sim P(\phi)$, where $Z_1 = X_1 + X_3$ and $Z_2 = X_2 + X_3$. Consequently:

$$\mu(Z_i) = \mathbb{V}\text{ar}(Z_i) = \mathbb{E}(Z_i) = \lambda_i + \phi, \quad i = 1, 2; \tag{2.2}$$

$$\mathbb{C}\text{ov}(Z_1, Z_2) = \phi. \tag{2.3}$$

Furthermore, such construction leads to a logical conclusion that $\phi \in [0, \min\{\lambda_1, \lambda_2\}]$

2.1 The bivariate seasonal INAR model: BINAR$_s$

2.1.1 Model

As first suggested by Al-Osh and Alzaid (1987) the AR process of stationary non-negative integer-valued random vectors X_t and $R_t, t \in \mathbb{Z}$ can be defined as (the INAR(1) model):

$$X_t = \alpha \circ X_{t-1} + R_t, \quad t \in \mathbb{Z}.$$

The given model suggests that realization of X at the time t consists of two components: the survival elements from time $t-1$ and an innovation R_t - i.i.d. non-negative integer-valued random variables with certain mean μ and finite variance σ^2 that arrive during the time $[t-1; t)$ and does not depend on past values of X_t. The survival component is sum of Y_i, $i \in \mathbb{Z}$ - a sequence of i.i.d. Bernoulli random variables where $\mathbb{P}(Y_i = 1) = \alpha = 1 - \mathbb{P}(Y_i = 0), \alpha \in [0, 1)$ (i.e. $Y_i = 1$ with probability α and $Y_i = 0$ with probability $1 - \alpha$) and can be denoted as:

$$\alpha \circ X_{t-1} := \sum_{i=1}^{X_{t-1}} Y_i.$$

Here '$\circ$' is a binomial thinning operator introduced by Steutel and Van Harn (1979).

The INAR process in real life example can be imagined as a queue of people, where each person queuing at a time $t - 1$ will remain in the queue at a time t with the probability α (survival component) and R_t number of new people will join the queue during the interval $[t - 1; t)$ (innovation component).

A further modification of the model to account for seasonality of time series as investigated by Bourguignon et al. (2016) and Buteikis and Leipus (2020) can be defined as (the INAR$(1)_s$ model):

$$X_t = \alpha \circ X_{t-s} + R_t, \quad t \in \mathbb{Z}.$$

The model simply assumes survival dependence with the time $t - s$ rather than time $t - 1$, where s is the seasonal period.

Similarly as presented above, we will introduce a model for bivariate integer-valued time series with seasonality.

Definition 2.1. *Let $\mathbf{X}_t = [X_{1,t}, X_{2,t}]', t \in \mathbb{Z}$ be stationary non-negative integer-valued bivariate time series and $\mathbf{R}_t = [R_{1,t}, R_{2,t}]', t \in \mathbb{Z}$ be a non-negative integer-valued bivariate random sequence independent from $\mathbf{X}_t$. Then the process $\mathbf{X}_t$ is a seasonal bivariate INAR process with seasonal period s (BINAR(1)$_s$), if it satisfies the equation:*

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

where $\alpha_i \in [0, 1),\ i = 1, 2$.

Worth noting that in the particular case, the parameter matrix $\mathbf{A}$ is diagonal for the sake of simplicity. Nevertheless, other forms could also be considered. It is meaningful to analyse a bivariate model if either coefficient matrix $\mathbf{A}$ is non-diagonal or innovations $\mathbf{R}_t$ are from the bivariate distribution (or joint in other way). The model is similar to a univariate case in a way it also consist of both survival and innovation components. What is more, the thinning operator '$\circ$' serves as a matrix multiplicator as well. Given the diagonality of coefficient matrix $\mathbf{A}$ in the particular case, time series of the BINAR$(1)_s$ can be expressed separately:

$$X_{t,j} = \alpha_j \circ X_{t-s,j} + R_{t,j}, \quad t \in \mathbb{Z},$$

where $j = 1, 2$.

Even though the latest equation is the same as in the univariate case of seasonal INAR model, the two models have different properties as to account for the fact the time series are of a bivariate process. The distributional properties of the model are revealed via distribution of the innovations $\mathbf{R}_t$.

It is interesting to note that BINAR$(1)_s$ model can be understood as a set of S independent BINAR(1) processes that share same coefficient matrix $\mathbf{A}$ and same innovations $\mathbf{R}_t$ distribution parameters. Moreover, the realization of $\mathbf{X}_t$ can also be expressed as a sum of innovations:

$$\begin{aligned} X_{t,j} &= \alpha_j \circ X_{t-s,j} + R_{t,j} = \alpha_j \circ (\alpha_j \circ X_{t-2s,j} + R_{t-s,j}) + R_{t,j} \\ &\stackrel{d}{=} \alpha_j^2 \circ X_{t-2s,j} + \alpha_j \circ R_{t-s,j} + R_{t,j} \stackrel{d}{=} \ldots \stackrel{d}{=} \sum_{k=0}^{\infty} \alpha_j^k \circ R_{t-ks,j}. \end{aligned} \tag{2.4}$$

2.1.2 Properties

In this section we will investigate some properties of the BINAR$(1)_s$ model. Properties have been derived by using similar methods as for BINAR(1) model by Pedeli (2011). Here we will assume the innovations $\mathbf{R}_t = [R_{1,t}, R_{2,t}]', t \in \mathbb{Z}$ to have a bivariate Poisson distribution.

Conditional mean:

$$\mathbb{E}(X_{i,t}|X_{i,t-s}) = \alpha_i X_{i,t-s} + \lambda_i + \phi, \quad i = 1, 2.$$

Proof:

$$\begin{aligned} \mathbb{E}(X_{i,t}|X_{i,t-s}) &= \mathbb{E}(\alpha_i \circ X_{i,t-s} + R_{i,t}|X_{i,t-s}) = \mathbb{E}(\alpha_i \circ X_{i,t-s}|X_{i,t-s}) + \mathbb{E}(R_{i,t}|X_{i,t-s}) \\ &= \alpha_i \mathbb{E}(X_{i,t-s}|X_{i,t-s}) + \mathbb{E}(R_{i,t}) = \alpha_i X_{i,t-s} + \lambda_i + \phi. \quad \square \end{aligned}$$

Unconditional mean:

$$\mathbb{E}(X_{t,i}) = \frac{\lambda_i + \phi}{1 - \alpha_i}, \quad i = 1, 2. \tag{2.5}$$

Proof:

$$\mathbb{E}(X_{t,i}) = \mathbb{E}(\sum_{k=0}^{\infty} \alpha_i^k \circ R_{t-ks,i}) = \sum_{k=0}^{\infty} \mathbb{E}(\alpha_i^k \circ R_{t-ks,i}) = \sum_{k=0}^{\infty} \alpha_i^k \mathbb{E} R_{t-ks,i} = \sum_{k=0}^{\infty} \alpha_i^k (\lambda_i + \phi) = \frac{\lambda_i + \phi}{1 - \alpha_i}. \quad \square$$

Unconditional variance:

$$\mathbb{V}\text{ar}(X_{t,i}) = \frac{\lambda_i + \phi}{1 - \alpha_i}, \quad i = 1, 2. \tag{2.6}$$

Proof:

$$\begin{aligned}
\mathbb{V}\text{ar}(X_{t,i}) &= \mathbb{V}\text{ar}(\sum_{k=0}^{\infty} \alpha_i^k \circ R_{t-ks,i}) = \sum_{k=0}^{\infty} \mathbb{V}\text{ar}(\alpha_i^k \circ R_{t-ks,i}) \\
&= \sum_{k=0}^{\infty} (\alpha_i^{2k} \mathbb{V}\text{ar}(R_{t-ks,i}) + \alpha_i^k (1 - \alpha_i^k) \mathbb{E}(R_{t-ks,i}) \\
&= \sum_{k=0}^{\infty} (\alpha_i^{2k} (\lambda_i + \phi) + \alpha_i^k (1 - \alpha_i^k)(\lambda_i + \phi)) \\
&= \frac{(\lambda_i + \phi)}{1 - \alpha_i^2} + \frac{(\lambda_i + \phi)}{1 - \alpha_i} - \frac{(\lambda_i + \phi)}{1 - \alpha_i^2} = \frac{\lambda_i + \phi}{1 - \alpha_i}. \quad \square
\end{aligned}$$

Covariance:

$$\mathbb{C}\text{ov}(X_{t,i}, X_{t,j}) = \frac{\phi}{1 - \alpha_i \alpha_j}, \quad i \neq j.$$

Proof:

$$\begin{aligned}
\mathbb{C}\text{ov}(X_{t,i}, X_{t,j}) &= \mathbb{C}\text{ov}(\sum_{k=0}^{\infty} \alpha_i^k \circ R_{t-ks,i}, \sum_{l=0}^{\infty} \alpha_i^l \circ R_{t-ls,j}) \\
&= \sum_{k,l=0}^{\infty} \alpha_i^k \alpha_j^l \mathbb{C}\text{ov}(R_{t-ks,i}, R_{t-ks,j}) = \sum_{k=0}^{\infty} \alpha_i^k \alpha_j^k \mathbb{C}\text{ov}(R_{t-ks,i}, R_{t-ks,j}) \\
&= \frac{\mathbb{C}\text{ov}(R_{t-ks,i}, R_{t-ks,j})}{1 - \alpha_i \alpha_j} = \frac{\phi}{1 - \alpha_i \alpha_j}. \quad \square
\end{aligned}$$

As it can be noted from the equations (2.5) and (2.6), unconditional variance and mean of the model (assuming bivariate Poisson distribution of the innovations) are equal. Thus model does not exhibit overdispersion.

2.1.3 Conditional Least Squares estimation

Parameters of BINAR$(1)_s$ model can be estimated using Conditional Least Squares (CLS). For the particular estimation, first we will construct the conditional expectations where the expectations depend on the information available at the time $t-s$:

$$\mu_{t|t-s} = \begin{bmatrix} \mathbb{E}(X_{1,t}|X_{1,t-s}, X_{2,t-s}) \\ \mathbb{E}(X_{2,t}|X_{1,t-s}, X_{2,t-s}) \end{bmatrix} = \begin{bmatrix} \alpha_1 X_{1,t-s} + \lambda_1 \\ \alpha_2 X_{2,t-s} + \lambda_2 \end{bmatrix},$$

here $\lambda_i = \mathbb{E}(R_{i,t})$ $i = 1, 2$.

Then we will calculate the difference between the actual realization at the time t and the conditional expectation based on information available at the time $t-s$:

$$X_t - \mu_{t|t-s} = \begin{bmatrix} X_{1,t} - \alpha_1 X_{1,t-s} - \lambda_1 \\ X_{2,t} - \alpha_2 X_{2,t-s} - \lambda_2 \end{bmatrix}.$$

To derive CLS estimates we will minimize the squared differences:

$$Q(\alpha_j, \lambda_j) = \min_{\alpha_j, \lambda_j} \sum_{n=s+1}^{N} (X_{j,t} - \alpha_j X_{j,t-s} - \lambda_j)^2, \quad j = 1, 2.$$

Taking derivatives of $Q(\alpha_j, \lambda_j)$ with respect to α_j and λ_j and equating them to 0 leads to the following system of equations:

$$\begin{cases} \dfrac{\partial Q}{\partial \alpha_j} = \displaystyle\sum_{n=s+1}^{N} -2X_{j,t-s}(X_{j,t} - \alpha_j X_{j,t-s} - \lambda_j) = 0 \\ \dfrac{\partial Q}{\partial \lambda_j} = \displaystyle\sum_{n=s+1}^{N} -2(X_{j,t} - \alpha_j X_{j,t-s} - \lambda_j) = 0 \end{cases}, \quad j = 1, 2. \tag{2.7}$$

From (2.7) we have the following equations:

$$\sum_{n=s+1}^{N} X_{j,t} X_{j,t-s} - \alpha_j \sum_{n=s+1}^{N} X_{j,t-s}^2 - \lambda_j \sum_{n=s+1}^{N} X_{j,t-s} = 0 \tag{2.8}$$

and

$$\lambda_j = \frac{1}{N-s} \sum_{n=s+1}^{N} (X_{j,t} - \alpha_j X_{j,t-s}) = \frac{1}{N-s} \sum_{n=s+1}^{N} X_{j,t} - \frac{\alpha_j}{N-s} \sum_{n=s+1}^{N} X_{j,t-s}. \tag{2.9}$$

Substituting (2.9) in (2.8):

$$\begin{aligned} &\sum_{n=s+1}^{N} X_{j,t} X_{j,t-s} - \alpha_j \sum_{n=s+1}^{N} X_{j,t-s}^2 - \frac{1}{N-s} \sum_{n=s+1}^{N} X_{j,t} \sum_{n=s+1}^{N} X_{j,t-s} \\ &+ \frac{\alpha_j}{N-s} \sum_{n=s+1}^{N} X_{j,t-s} \sum_{n=s+1}^{N} X_{j,t-s} = \sum_{n=s+1}^{N} X_{j,t} X_{j,t-s} - \frac{1}{N-s} \sum_{n=s+1}^{N} X_{j,t} \sum_{n=s+1}^{N} X_{j,t-s} \\ &- \alpha_j (\sum_{n=s+1}^{N} X_{j,t-s}^2 - \frac{1}{N-s} \sum_{n=s+1}^{N} X_{j,t-s} \sum_{n=s+1}^{N} X_{j,t-s}) = 0. \end{aligned} \tag{2.10}$$

From (2.10) we have:

$$\begin{aligned}&\alpha_j(\sum_{n=s+1}^{N} X_{j,t-s}^2 - \frac{1}{N-s}\sum_{n=s+1}^{N} X_{j,t-s}\sum_{n=s+1}^{N} X_{j,t-s})\\&= \sum_{n=s+1}^{N} X_{j,t}X_{j,t-s} - \frac{1}{N-s}\sum_{n=s+1}^{N} X_{j,t}\sum_{n=s+1}^{N} X_{j,t-s}.\end{aligned} \tag{2.11}$$

Denoting mean as follows:

$$\bar{X}_j = \frac{1}{N-s}\sum_{n=s+1}^{N} X_{j,t-s},$$

(2.11) becomes

$$\alpha_j = \frac{\sum_{n=s+1}^{N} X_{j,t}X_{j,t-s} - \sum_{n=s+1}^{N} X_{j,t}\bar{X}_j}{\sum_{n=s+1}^{N} X_{j,t-s}^2 - (N-s)\bar{X}_j^2} = \frac{\sum_{n=s+1}^{N}(X_{j,t}(X_{j,t-s} - \bar{X}_j))}{\sum_{n=s+1}^{N}(X_{j,t-s}^2 - \bar{X}_j^2)}.$$

Hence,

$$\begin{aligned}\hat{\alpha}_j^{CLS} &= \frac{\sum_{n=s+1}^{N}(X_{j,t}(X_{j,t-s} - \bar{X}_j))}{\sum_{n=s+1}^{N}(X_{j,t-s}^2 - \bar{X}_j^2)}, \quad j = 1, 2,\\ \hat{\lambda}_j^{CLS} &= \frac{1}{N-s}\sum_{n=s+1}^{N}(X_{j,t} - \hat{\alpha}_j^{CLS}X_{j,t-s}), \quad j = 1, 2.\end{aligned} \tag{2.12}$$

Assuming the innovations to have bivariate distribution, we can see that CLS estimates do not reveal the dependence parameter. As suggested by Pedeli (2011), it can be shown that the residuals of the model are equal to the covariance of the innovation terms. Hence, the dependence parameter ϕ can be estimated by minimizing the squared differences of the model residuals and the covariance between the innovations:

$$Q(\mathbb{C}\text{ov}(R_1, R_2)) = \min_{\mathbb{C}\text{ov}(R_1,R_2)} \sum_{t=s+1}^{N}\{(X_{1,t} - \alpha_1 X_{1,t-s} - \lambda_1)(X_{2,t} - \alpha_2 X_{2,t-s} - \lambda_2) - \mathbb{C}\text{ov}(R_1, R_2)\}^2.$$

If we assume bivariate Poisson distribution between the innovations, from (2.3) $\mathbb{C}\text{ov}(R_1, R_2) = \phi$, and (2.1.3) becomes:

$$Q(\phi) = \min_{\phi} \sum_{t=s+1}^{N}\{(X_{1,t} - \alpha_1 X_{1,t-s} - \lambda_1)(X_{2,t} - \alpha_2 X_{2,t-s} - \lambda_2) - \phi\}^2. \tag{2.13}$$

Taking derivative of (2.13) with respect to ϕ leads to:

$$\frac{\partial Q}{\partial \phi} = 2\sum_{t=s+1}^{N}\{(X_{1,t} - \alpha_1 X_{1,t-s} - \lambda_1)(X_{2,t} - \alpha_2 X_{2,t-s} - \lambda_2) - \phi\} = 0$$

$$
\begin{aligned}
\phi =& \frac{1}{N-s}\{\sum_{t=s+1}^{N}(X_{1,t}-\alpha_1 X_{1,t-s})(X_{2,t}-\alpha_2 X_{2,t-s})- \\
&-\lambda_1\sum_{t=s+1}^{N}(X_{2,t}-\alpha_2 X_{2,t-s})-\lambda_2\sum_{t=s+1}^{N}(X_{1,t}-\alpha_1 X_{1,t-s})\}+\lambda_1\lambda_2.
\end{aligned} \tag{2.14}
$$

Since from (2.12)

$$
\hat{\lambda}_j^{CLS}=\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{j,t}-\hat{\alpha}_j^{CLS}X_{j,t-s})
$$

Subsituting $\hat{\lambda}_j^{CLS}$ in (2.14) we have:

$$
\begin{aligned}
\hat{\phi}^{CLS} =& \frac{1}{N-s}\sum_{t=s+1}^{N}(X_{1,t}-\hat{\alpha}_1^{CLS}X_{1,t-s})(X_{2,t}-\hat{\alpha}_2^{CLS}X_{2,t-s}) \\
&-\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{1,t}-\hat{\alpha}_1^{CLS}X_{1,t-s})\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{2,t}-\hat{\alpha}_2^{CLS}X_{2,t-s}) \\
&-\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{2,t}-\hat{\alpha}_2^{CLS}X_{2,t-s})\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{1,t}-\hat{\alpha}_1^{CLS}X_{1,t-s}) \\
&+\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{1,t}-\hat{\alpha}_1^{CLS}X_{1,t-s})\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{2,t}-\hat{\alpha}_2^{CLS}X_{2,t-s}) \\
=& \frac{1}{N-s}\{\sum_{t=s+1}^{N}(X_{1,t}-\hat{\alpha}_1^{CLS}X_{1,t-s})(X_{2,t}-\hat{\alpha}_2^{CLS}X_{2,t-s}) \\
&-\frac{1}{N-s}\sum_{t=s+1}^{N}(X_{1,t}-\hat{\alpha}_1^{CLS}X_{1,t-s})\sum_{t=s+1}^{N}(X_{2,t}-\hat{\alpha}_2^{CLS}X_{2,t-s})\}.
\end{aligned}
$$

Worth noting that $\hat{\lambda}_j^{CLS}$ here represents $\mathbb{E}(R_{j,t})$. Assuming the innovations term in the BINAR(1)$_s$ model are distributed under bivariate Poisson distribution $(R_1, R_2) \sim BPo(\lambda_1, \lambda_2, \phi)$, the expectations are $\mathbb{E}(R_{j,t}) = \lambda_j + \phi$. Thus $\hat{\lambda}_j^{*CLS}$, representing the actual estimate of λ_j from the bivariate Poisson distribution is:

$$
\hat{\lambda}_j^{*CLS}=\hat{\lambda}_j^{CLS}-\hat{\phi}^{CLS}.
$$

2.1.4 Conditional Maximum Likelihood estimation

The following estimation of BINAR(1)$_s$ via Conditional Maximum Likelihood (CML) is an approach used in Pedeli and Karlis (2011) and Pedeli and Karlis (2013) for BINAR(1) and adapted to for the seasonal model. The estimator of density can be written as a convolution of two binomials (survival elements) and a bivariate distribution (innovations). As mentioned, for this instance we will assume the Bivariate Poisson distribution for the innovations. The pmf of binomials $\mathbf{X}_t = [X_{1,t}, X_{2,t}]'$ can be written as follows:

$$f_1(x_1) = \mathbb{P}(X_{1,t} = x_{1,t} | X_{1,t-s} = x_{1,t-s}) = \binom{x_{1,t-s}}{x_{1,t}} \alpha_1^{x_{1,t}} (1-\alpha_1)^{x_{1,t-s}-x_{1,t}}, \tag{2.15}$$

$$f_2(x_2) = \mathbb{P}(X_{2,t} = x_{2,t} | X_{2,t-s} = x_{2,t-s}) = \binom{x_{2,t-s}}{x_{2,t}} \alpha_2^{x_{2,t}} (1-\alpha_2)^{x_{2,t-s}-x_{2,t}}, \tag{2.16}$$

where $\binom{a}{b}$ is a binomial coefficient and is calculated as follows:

$$\binom{a}{b} = \frac{a!}{b!\,(a-b)!}.$$

Furthermore, considering distributional choice, pmf of innovations $\mathbf{R}_t = [R_{1,t}, R_{2,t}]'$ can be written (as indicated in the (2.1)):

$$\begin{aligned} f_3(k,l) = P(R_{1,t} = k, R_{2,t} = l) = & e^{-(\lambda_1-\lambda_2-\phi)} \frac{(\lambda_1-\phi)^k}{k!} \frac{(\lambda_2-\phi)^l}{l!} \\ & \times \sum_{m=0}^{\min(k,l)} \binom{k}{m} \binom{l}{m} m! (\frac{\phi}{(\lambda_1-\phi)(\lambda_2-\phi)})^m. \end{aligned}$$

Given the above, the joint pmf (jpmf) for the BINAR$(1)_s$ process becomes:

$$\begin{aligned} & f(x_{1,t}, x_{2,t} | x_{1,t-s}, x_{2,t-s}, \alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi) = \sum_{k=0}^{g_1} \sum_{l=0}^{g_2} f_1(x_{1,t}-k) f_2(x_{2,t}-l) f_3(k,l) \\ = & \sum_{k=0}^{g_1} \sum_{l=0}^{g_2} e^{-(\lambda_1-\lambda_2-\phi)} \frac{(\lambda_1-\phi)^k}{k!} \frac{(\lambda_2-\phi)^l}{l!} \sum_{m=0}^{\min(k,l)} \binom{k}{m} \binom{l}{m} m! (\frac{\phi}{(\lambda_1-\phi)(\lambda_2\phi)})^m \\ & \times \binom{x_{1,t-s}}{x_{1,t}-k} \alpha_1^{x_{1,t}-k} (1-\alpha_1)^{x_{1,t-s}-x_{1,t}+k} \binom{x_{2,t-s}}{x_{2,t}-l} \alpha_2^{x_{2,t}-l} (1-\alpha_2)^{x_{2,t-s}-x_{2,t}+l}, \end{aligned}$$

where $g_j = \min(x_{j,t}, x_{j,t-s}), \quad j = 1, 2$. From the jpmf we can now define the conditional log-likelihood function:

$$\ell(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi) = \sum_{t=s+1}^{N} \log f(x_{1,t}, x_{2,t} | x_{1,t-s}, x_{2,t-s}, \alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi).$$

To obtain CLM estimates of the unknown parameters, we maximize the conditional log-likelihood:

$$\hat{\theta}^{CLM} = \max_{\alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi} \ell(\alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi).$$

Here θ indicates vector of the unknown parameters. Having formulated the maximization problem, statistical software can be used to obtain the actual estimates of the parameters. In the particular instance, function *optim* in R was employed for the calculations. CLS estimates were used as the initial values for the CML estimations. Although CML method is used to find global extreme, unreasonably selected initial values of the parameters may lead to finding local rather than global extremes and, hence, biased results.

2.1.5 Simulation

Simulations were performed to test if the constructed estimates are plausible. Both CLS and CML methods were tested on the same sets of simulated data. Two different lengths of the time series have been chosen to see how methods perform on small and larger samples. For the first set of simulations, the length of the time series was $N = 100$, for the second - $N = 500$. For each set a total of 200 independent experiments were performed. Actual parameters chosen were the same for both sets of simulations. Parameters α_1 and α_2 indicating marginal dependence of the series $\mathbf{X}_t$ respectively were chosen to be 0.7 and 0.5. Worth noting that data was simulated to have a seasonal dependence with $s = 12$ to resemble seasonal dependence of the data used in the empirical part. Parameters λ_1, λ_2 and ϕ of the Bivariate Poisson distribution accordingly are 2, 3 and 0.5.

To evaluate and compare parameter estimates certain measures of goodness need to be introduced. For the numerical evaluation we will calculate Bias (2.17) and Root Mean Square Error (RMSE) (2.18). For the visual comparison we will employ box plots.

$$Bias = \frac{1}{T}\sum_{t=1}^{T}(\hat{\theta}_t - \theta_t), \tag{2.17}$$

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{\theta}_t - \theta_t)^2}. \tag{2.18}$$

Table 1 represents results of both simulations. One can notice that within the set of shorter simulations ($N = 100$) CML estimate does not significantly over-perform CLS estimate. For both estimators BIAS and RMSE are relatively similar, and for the estimate of λ_2 one can notice that CLS estimator was even better. Nevertheless, in the set of simulations where $N = 500$ results are slightly different. Even though BIAS is relatively similar for both estimators, RMSE reveals better performance CML estimator.

Table 1: Simulation results of BINAR(1)$_{s=12}$

		CLS estimation		CML estimation	
Size	Real parameter	BIAS	RMSE	BIAS	RMSE
$N = 100$	α_1=0.7	0.0233	0.1303	-0.0056	0.1578
	α_2=0.5	-0.0264	0.0923	-0.0376	0.1206
	λ_1=2	-0.1636	1.0204	-0.0280	0.8797
	λ_2=3	0.1833	0.8423	0.2976	0.9505
	$\phi = 0.5$	-0.0027	0.4728	0.0206	0.4223
$N = 500$	α_1=0.7	-0.0028	0.0337	0.0020	0.0151
	α_2=0.5	-0.0010	0.0385	0.0002	0.0149
	λ_1=2	0.0074	0.3526	0.0064	0.2071
	λ_2=3	-0.0143	0.3347	-0.0045	0.2107
	$\phi = 0.5$	0.0161	0.2166	-0.0129	0.1700

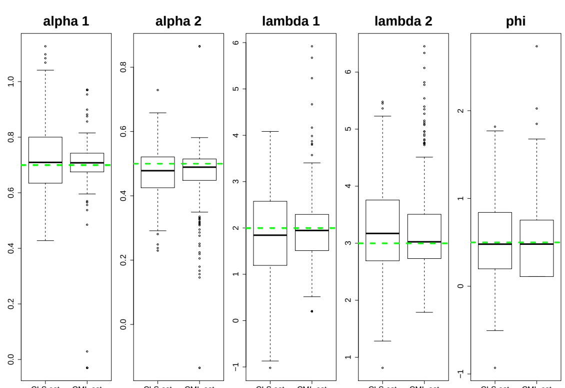

Figure 1: Simulation results of BINAR(1)$_{s=12}$; $N = 100$

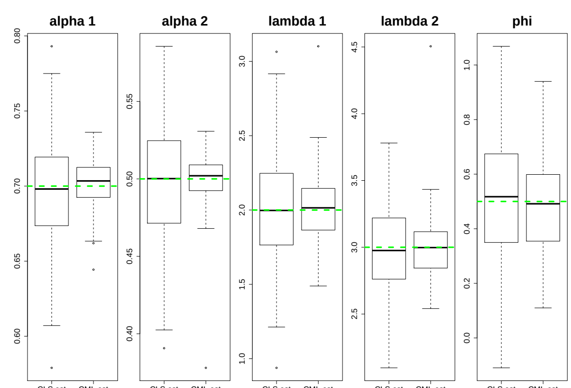

Figure 2: Simulation results of BINAR(1)$_{s=12}$; $N = 500$

Similar ideas are supported by the box plots. From Figure 1 we see that median values of both estimators are close to the value of the real parameter used for the simulations (marked in green dashed line). Additionally, even if more estimations are closer to the median value, CML produced quite a number of outliers. Also, some cases had α parameters greater than 1. Having a bigger sample solves for the problem of the outliers in the CML estimators as well as unreasonably large

α (see Figure 2). In the set of simulations where $N = 500$, CLM estimates were closer to the median value as the box plots are significantly shorter than those of CLS estimates.

2.2 The bivariate seasonal INGARCH model: BINGARCH$_\mathrm{S}$

2.2.1 Model

Another model for integer-valued time series is a GARCH type model first suggested by Ferland et al. (2006). The main idea behind the model is to propose a distribution suitable for integer data and assume time dependence of the parameters. Namely, univariate INGARCH(p, q) model, assuming Poisson distribution, can be written as follows:

$$\begin{cases} X_t|\mathcal{F}_{t-1} \sim P(\lambda_t), \quad t \in \mathbb{Z}, \\ \lambda_t = \delta + \sum_{i=1}^{q} \alpha_i \lambda_{t-i} + \sum_{j=1}^{p} \beta_j X_{t-j}. \end{cases}$$

Model assumes that conditional mean at the time t depends on the past values of the conditional mean and the previous observations of the variable itself. The model is known to be easily estimated via CML. Moreover, positive value of λ_t can be ensured by keeping all parameters α_i, $i = 1, ..., q$ and β_j, $j = 1, ..., p$ non-negative and parameter δ positive. The aforementioned model is also called a Poisson autoregression (e.g. Fokianos et al. (2009)). Such models are not limited to Poisson distribution. Zhu (2011) formulated INGARCH(p, q) process for negative binomial distribution, Zhu (2012) elaborated on zero-inflated Poisson distribution.

Based on the fact that a number of count data processes are of a multivariate nature, Liu (2012) proposed a bivariate observation-driven count data model. Bivariate Poisson autoregression or a BINGARCH(p, q) is defined as:

$$\begin{cases} \mathbf{X}_t|\mathcal{F}_{t-1} \sim BP(\lambda_{1,t}, \lambda_{2,t}, \phi), \quad t \in \mathbb{Z}, \\ \boldsymbol{\lambda}_t = \boldsymbol{\delta} + \sum_{i=1}^{q} \mathbf{A}_i \boldsymbol{\lambda}_{t-i} + \sum_{j=1}^{p} \mathbf{B}_j \mathbf{X}_{t-j}. \end{cases}$$

Here $\boldsymbol{\lambda}_t = [\lambda_{1,t}, \lambda_{2,t}]'$, $t \in \mathbb{Z}$, and coefficient matrices are as follows:

$$\boldsymbol{\delta} = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}; \quad \mathbf{A}_i = \begin{bmatrix} \alpha_{i1} & \alpha_{i2} \\ \alpha_{i3} & \alpha_{i4} \end{bmatrix}; \quad \mathbf{B}_j = \begin{bmatrix} \beta_{j1} & \beta_{j2} \\ \beta_{j3} & \beta_{j4} \end{bmatrix}.$$

Worth noting that structure of the coefficient matrices can be chosen for each case individually. However, it is reasonable to consider a bivariate model if at least one of the coefficient matrices is not diagonal or ϕ parameter is not equal to 0. Particular model establishes time dependent λ parameters. In addition, time variant ϕ could also be analysed.

BINGARCH model for data with seasonal structure can be formulated as an instance of the BINGARCH(p,q) model including first and the seasonal lags of the conditional mean and the variables. Such form is suggested accounting for the fact that most data that exhibits seasonality also depends from the first lag and, hence, such choice should help to explain the process better.

Definition 2.2. *Let $\mathbf{X}_t = [X_{1,t}, X_{2,t}]', t \in \mathbb{Z}$ be stationary non-negative integer-valued bivariate time series. Then the process $\mathbf{X}_t$ is a seasonal bivariate INGARCH process with seasonal period s (BINGARCH$_s$), if it satisfies:*

$$\begin{cases} \mathbf{X}_t|\mathcal{F}_{t-1} \sim BP(\lambda_{1,t}, \lambda_{2,t}, \phi), \quad t \in \mathbb{Z} \\ \boldsymbol{\lambda}_t = [\lambda_{1,t}, \lambda_{2,t}]' = \boldsymbol{\delta} + \mathbf{A}_1\boldsymbol{\lambda}_{t-1} + \mathbf{A}_2\boldsymbol{\lambda}_{t-s} + \mathbf{B}_1\mathbf{X}_{t-1} + \mathbf{B}_2\mathbf{X}_{t-s} \\ \quad = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{12} \end{bmatrix} \begin{bmatrix} \lambda_{1,t-1} \\ \lambda_{2,t-1} \end{bmatrix} + \begin{bmatrix} \alpha_{21} & 0 \\ 0 & \alpha_{22} \end{bmatrix} \begin{bmatrix} \lambda_{1,t-s} \\ \lambda_{2,t-s} \end{bmatrix} \\ \quad + \begin{bmatrix} \beta_{11} & 0 \\ 0 & \beta_{12} \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} \beta_{21} & 0 \\ 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} X_{1,t-s} \\ X_{2,t-s} \end{bmatrix} \end{cases}$$

where $\delta_i > 0,\ i = 1, 2,\ \alpha_{i,j} \geq 0,\ i = 1, 2,\ j = 1, 2, \beta_{i,j} \geq 0,\ i = 1, 2,\ j = 1, 2.$

As evident from (2.2), in the particular case the coefficient matrices were chosen to be diagonal leaving the dependence between time series to be revealed via parameter ϕ. Such structure has been chosen to simplify the model and reduce the number of parameters to be estimated. It also suggests that realisation of data at the time t depends from certain factors that apply to both variables of the data and affects them in the same way. Moreover, such structural form allows to express each variable similarly as in the univariate case and examine properties accordingly:

$$\begin{cases} X_{i,t}|\mathcal{F}_{t-1} \sim P(\lambda_{i,t} + \phi), \quad t \in \mathbb{Z}, \\ \lambda_{i,t} = \delta_i + \alpha_{1i}\lambda_{i,t-1} + \alpha_{2i}\lambda_{i,t-s} + \beta_{1i}X_{i,t-1} + \beta_{2i}X_{i,t-s} \end{cases} \qquad i = 1, 2.$$

Stability theory for BINGARCH(p,q) model and, hence, applicable to BINGARCH$_s$ as well (being an instance of the former model) has been provided by Liu (2012).

2.2.2 Properties

In this section we will investigate some properties of the BINGARCH$_s$ model.

Conditional mean and variance:

$$\mu_{X_{i,t}|\mathcal{F}_{t-1}} = \mathbb{V}\text{ar}[X_{i,t}|\mathcal{F}_{t-1}] = \mathbb{E}(X_{i,t}|\mathcal{F}_{t-1}) = \lambda_{i,t} + \phi$$
$$= \delta_i + \alpha_{1,i}\lambda_{i,t-1} + \alpha_{2,i}\lambda_{i,t-s} + \beta_{1,i}X_{i,t-1} + \beta_{2,i}X_{i,t-s} + \phi, \quad i = 1, 2.$$

Covariance (by the property of bivariate Poisson distribution from (2.3)):

$$\mathbb{C}\text{ov}(X_1, X_2) = \phi.$$

Unconditional mean:

$$\mathbb{E}(X_{i,t}) = \mu_i = \mathbb{E}(X_{i,t}) = \frac{\delta_i + \phi(1 - \alpha_1 - \alpha_2)}{1 - (\alpha_{1,i} + \alpha_{2,i} + \beta_{1,i} + \beta_{2,i})}, \quad i = 1, 2. \tag{2.19}$$

Proof:

$$\mathbb{E}(X_{i,t}) = \delta_i + \phi + \alpha_{1,i}\mathbb{E}(\lambda_{t-1}) + \alpha_{2,i}\mathbb{E}(\lambda_{t-s}) + \beta_{1,i}\mathbb{E}(X_{i,t-1}) + \beta_{2,i}\mathbb{E}(X_{i,t-s}). \tag{2.20}$$

Assuming the process is stationary, we can write:

$$\mathbb{E}(X_{i,t}) = \mathbb{E}(X_{i,t-1}) = \mathbb{E}(X_{i,t-s})$$

and

$$\mathbb{E}(\lambda_{t-1}) = \mathbb{E}(\lambda_{t-s}).$$

What is more, from the properties of the bivariate Poisson distribution we can note that:

$$\mathbb{E}(\lambda_{t-1}) = \mathbb{E}(X_{i,t} - \phi).$$

Hence, (2.20) can be rewritten:

$$\mathbb{E}(X_{i,t}) = \delta_i + \phi + \alpha_{1,i}\mathbb{E}(X_{i,t} - \phi) + \alpha_{2,i}\mathbb{E}(X_{i,t} - \phi) + \beta_{1,i}\mathbb{E}(X_{i,t}) + \beta_{2,i}\mathbb{E}(X_{i,t}). \tag{2.21}$$

Collecting the terms in (2.21) we obtain:

$$\mathbb{E}(X_{i,t}) = \frac{\delta_i + \phi(1 - \alpha_1 - \alpha_2)}{1 - (\alpha_{1,i} + \alpha_{2,i} + \beta_{1,i} + \beta_{2,i})}. \quad \square$$

Unconditional variance:

$$\mathbb{V}\text{ar}[X_{i,t}] = \frac{\mu_i(1 - (\alpha_{1,i} + \beta_{1,i})^2 - (\alpha_{2,i} + \beta_{2,i})^2 + (\alpha_{1,i}^2 + \alpha_{2,i}^2))}{1 - ((\alpha_{1,i} + \beta_{1,i})^2 + (\alpha_{2,i} + \beta_{2,i})^2)}. \tag{2.22}$$

Proof:

Equation can be proven in a similar manner as for APC(1,1) model (see Heinen (2003)) by applying the following property on conditional variance:

$$\mathbb{V}\text{ar}[y] = \mathbb{E}_x[\mathbb{V}\text{ar}_{y|x}(y|x)] + \mathbb{V}\text{ar}_x[\mathbb{E}_{y|x}(y|x)].$$

Here

$$\mathbb{E}_x[\mathbb{V}\text{ar}_{y|x}(y|x)] = \mathbb{E}[(X_{i,t} - \mu_{X_{i,t}})^2] = \mu_i, \tag{2.23}$$

as defined in (2.19). And

$$\mathbb{V}\text{ar}_x[\mathbb{E}_{y|x}(y|x)] = \mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2]. \tag{2.24}$$

To calculate the later expression (2.24), following steps can be taken:

$$\mu_{X_{i,t}} - \mu_i = \alpha_{1,i}(X_{i,t-1} - \mu_i) + \beta_{1,i}(\mu_{X_{i,t-1}} - \mu_i) + \alpha_{2,i}(X_{i,t-s} - \mu_i) + \beta_{2,i}(\mu_{X_{i,t-s}} - \mu_i$$

$$\begin{aligned}\mu_{X_{i,t}} - \mu_i = {} & \alpha_{1,i}(X_{i,t-1} - \mu_{X_{i,t}}) + (\alpha_{1,i} + \beta_{1,i})(\mu_{X_{i,t-1}} - \mu_i) \\ & + \alpha_{2,i}(X_{i,t-s} - \mu_{X_{i,t}}) + (\alpha_{2,i} + \beta_{2,i})(\mu_{X_{i,t-s}} - \mu_i).\end{aligned} \tag{2.25}$$

Squaring and taking the expectation of (2.25):

$$\begin{aligned}\mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2] = {} & \alpha_{1,i}^2\mathbb{E}[(X_{i,t-1} - \mu_{X_{i,t-1}})^2] + (\alpha_{1,i} + \beta_{1,i})^2\mathbb{E}[(\mu_{X_{i,t-1}} - \mu_i)^2] \\ & + \alpha_{2,i}^2\mathbb{E}[(X_{i,t-s} - \mu_{X_{i,t}})^2] + (\alpha_{2,i} + \beta_{2,i})^2\mathbb{E}[(\mu_{X_{i,t-s}} - \mu_i)^2].\end{aligned} \tag{2.26}$$

Replacing conditional variance in (2.26) by its expression:

$$\mathbb{E}[(X_{i,t-1} - \mu_{X_{i,t-1}})^2] = \mathbb{E}[(X_{i,t-s} - \mu_{X_{i,t-s}})^2] = \mu_i,$$

we obtain:

$$\begin{aligned}\mathbb{V}\text{ar}[\mu_{X_{i,t}}] = {} & \mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2] = (\alpha_{1,i}^2 + \alpha_{2,i}^2)\mu_i \\ & + (\alpha_{1,i} + \beta_{1,i})^2\mathbb{E}[(\mu_{X_{i,t-1}} - \mu_i)^2] + (\alpha_{2,i} + \beta_{2,i})^2\mathbb{E}[(\mu_{X_{i,t-s}} - \mu_i)^2].\end{aligned} \tag{2.27}$$

Here, assuming the process is stationary, we can write:

$$\mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2] = \mathbb{E}[(\mu_{X_{i,t-1}} - \mu_i)^2] = \mathbb{E}[(\mu_{X_{i,t-s}} - \mu_i)^2].$$

Accordingly, we can rewrite (2.27):

$$\begin{aligned}\mathbb{V}\text{ar}[\mu_{X_{i,t}}] =&\mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2] = (\alpha_{1,i}^2 + \alpha_{2,i}^2)\mu_i \\ &+ (\alpha_{1,i} + \beta_{1,i})^2\mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2] + (\alpha_{2,i} + \beta_{2,i})^2\mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2].\end{aligned} \tag{2.28}$$

Collecting the terms in (2.28) we obtain:

$$\mathbb{V}\text{ar}[\mu_{X_{i,t}}] = \mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2] = \frac{\mu_i(\alpha_{1,i}^2 + \alpha_{2,i}^2)}{1 - (\alpha_{1,i} + \beta_{1,i})^2 - (\alpha_{2,i} + \beta_{2,i})^2}. \tag{2.29}$$

Finally, from (2.23) and (2.29):

$$\begin{aligned}\mathbb{V}\text{ar}[X_{i,t}] = \mathbb{E}[(X_{i,t} - \mu_i)^2] &= \mathbb{E}[(X_{i,t} - \mu_{X_{i,t}})^2] + \mathbb{E}[(\mu_{X_{i,t}} - \mu_i)^2] \\ &= \mu_i + \frac{\mu_i(\alpha_{1,i}^2 + \alpha_{2,i}^2)}{1 - (\alpha_{1,i} + \beta_{1,i})^2 - (\alpha_{2,i} + \beta_{2,i})^2} \\ &= \frac{\mu_i(1 - (\alpha_{1,i} + \beta_{1,i})^2 - (\alpha_{2,i} + \beta_{2,i})^2 + (\alpha_{1,i}^2 + \alpha_{2,i}^2))}{1 - (\alpha_{1,i} + \beta_{1,i})^2 - (\alpha_{2,i} + \beta_{2,i})^2}. \quad \square\end{aligned}$$

From expression of unconditional variance in equation (2.22), it is not difficult to see that $\mathbb{V}\text{ar}[\mu_{X_{i,t}}] > \mu_i$ if $(\alpha_{1,i}^2 + \alpha_{2,i}^2) \neq 0$. Hence BINGARCH$_s$ model exhibits overdispersion, even if equidispersed marginal distribution (i.e. bivariate Poisson) is considered. Moreover, from (2.22) and (2.19) one can notice that model has finite unconditional mean and variance if $(\alpha_{1,i} + \alpha_{2,i} + \beta_{1,i} + \beta_{2,i}) < 1$.

2.2.3 Conditional Maximum Likelihood estimation

Approach how to estimate BINGARCH(p, q) model via CML has been described by Liu (2012). Method is rather straight forward having the pmf of the data distribution can be used for estimation of BINGARCH$_s$ as well:

$$\begin{aligned}f(k, l) = P(X_{1,t} = k, X_{2,t} = l) =& e^{-(\lambda_{1,t} - \lambda_{2,t} - \phi)} \frac{(\lambda_{1,t} - \phi)^k}{k!} \frac{(\lambda_{2,t} - \phi)^l}{l!} \\ &\times \sum_{m=0}^{\min(k,l)} \binom{k}{m}\binom{l}{m} m! (\frac{\phi}{(\lambda_{1,t} - \phi)(\lambda_{2,t} - \phi)})^m,\end{aligned} \tag{2.30}$$

where

$$\boldsymbol{\lambda}_t = \boldsymbol{\delta} + \mathbf{A}_1\boldsymbol{\lambda}_{t-1} + \mathbf{A}_2\boldsymbol{\lambda}_{t-s} + \mathbf{B}_1\mathbf{X}_{t-1} + \mathbf{B}_2\mathbf{X}_{t-s},$$

as described in (2.2).

From (2.30) we can define the conditional log-likelihood function:

$$\ell(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \phi)$$
$$= \sum_{t=s+1}^{N} \log f(x_{1,t}, x_{2,t} | x_{1,t-1}, x_{2,t-1}, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \phi).$$

To obtain the CML estimates we maximize:

$$\hat{\theta}^{CLM} = \max_{\alpha_{11},\alpha_{12},\alpha_{21},\alpha_{22},\beta_{11},\beta_{12},\beta_{21},\beta_{22},\phi} \ell(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \phi).$$

Although the maximisation problem seems to be difficult considering number of parameters to be optimized, actual maximisation process is less intense compared to that of the BINAR$(1)_s$. Function *nlminb* in R was used for the estimation of BINGARCH$_s$. The initial parameter values for maximisation process were selected randomly (in such a way that log-likelihood can be estimated at the given values).

2.2.4 Simulation

Similarly as in Section 2.1.5 to sets of 200 independent simulations were performed. Length of series for the first set was $N = 100$, for the second - $N = 500$. Seasonal lag has been chosen $s = 12$, once again, to account for the structure of data later used in the empirical part. Bias and RMSE (as introduced in the equations (2.17) and (2.18)) of the simulations are provided in the Table 2. Box plots of the CML estimates are also available in the Figures 3 and 4.

As it can be noted from the Bias and RMSE of the estimated parameters, CML is rather accurate, as the estimates are close to the real parameter values (Bias and RMSE are small). Although in this instance initial parameter values were selected randomly (since CLS estimates were not available), accuracy did not deteriorate. Having larger sample ($N = 500$) has even improved the accuracy of the estimates. Similarly as in the instance of the BINAR$(1)_s$, having smaller sample size led to having more outliers of the parameter estimations as can be seen from Figure 3. Larger sample size reduced this problem. However, from Figure 4 it can be noted that a number of outliers were estimated in terms of intercept (δ_1 and δ_2). Nevertheless, in the set of simulations with the larger sample size, more parameters were estimated closer to the median and their real values as can be seen from the shorter box plots.

Table 2: Simulation results of $\mathrm{BINGARCH}_{s=12}$

Size	$N = 100$		$N = 500$	
Real parameter	BIAS	RMSE	BIAS	RMSE
δ_1=2	0.1160	1.1055	0.1366	0.6808
δ_1=1	-0.1031	0.5152	-0.0318	0.3329
α_{11}=0.15	0.0492	0.2069	0.0033	0.1025
α_{12}=0.2	0.0960	0.2220	-0.0945	0.1297
α_{21}=0.1	0.0535	0.2062	0.0767	0.1305
α_{22}=0.2	0.0202	0.2037	0.0032	0.1211
β_{11}=0.3	-0.0152	0.1496	-0.0019	0.0443
β_{12}=0.3	-0.0317	0.1321	-0.0094	0.0464
β_{21}=0.1	0.0096	0.1221	-0.0001	0.0411
β_{22}=0.05	0.0411	0.0798	0.0007	0.0384
ϕ=0.1	0.2593	0.4394	0.0507	0.1756

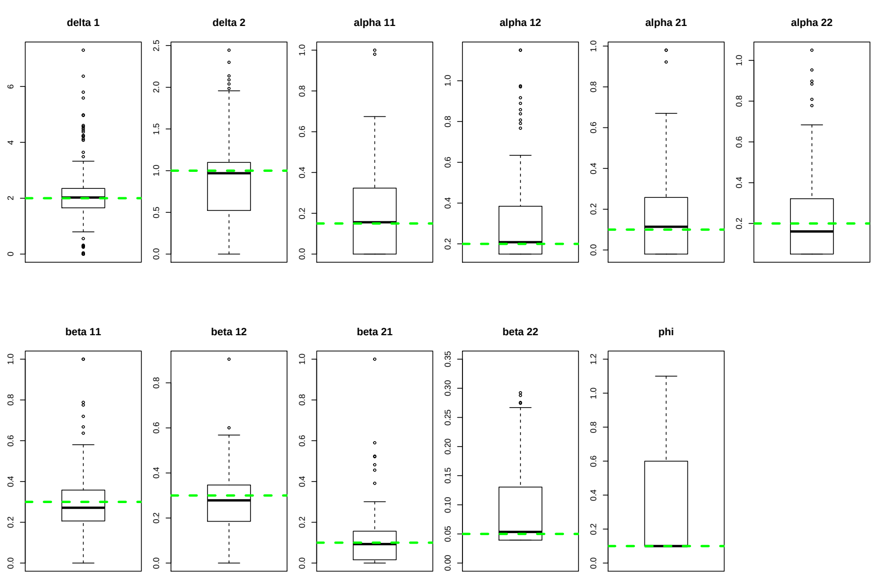

Figure 3: Simulation results of $\mathrm{BINGARCH}_{s=12}$; $N = 100$

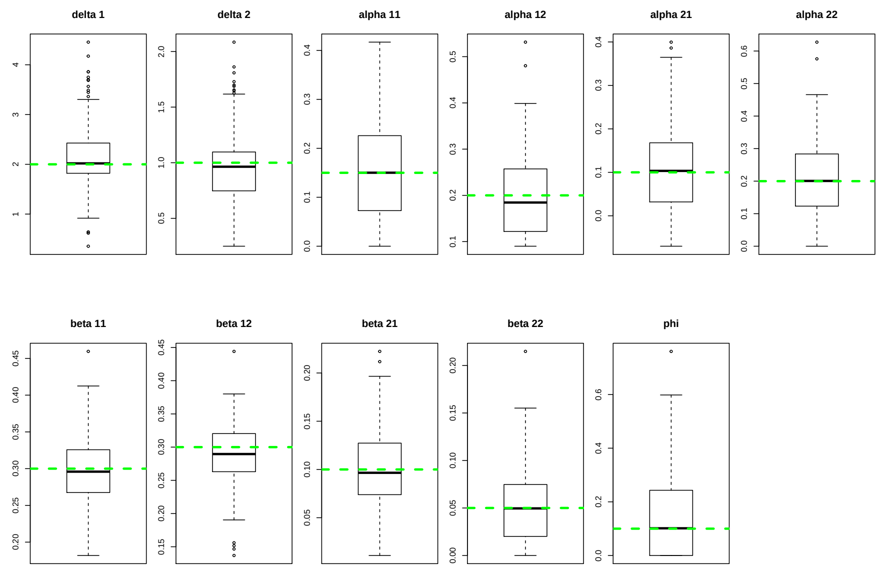

Figure 4: Simulation results of BINGARCH$_{s=12}$; $N = 500$

2.3 The bivariate time variant seasonal INAR model: TV-BINAR$_\mathrm{s}$

2.3.1 Model

In this section we will formulate new model for the bivariate count data with seasonality. As noted Section 2.2.1, most data with expressed seasonality also exhibits significant dependence on the first lag. BINGARCH model is capable of including additional lags without complicating estimation of the model significantly. However, this model is limited to analysing the survival information of the data. BINAR model on the other hand has a capability to model process in terms of both survival and arrival elements. Although it technically is possible to extend this model to BINAR(p) and introduce more lags (and e.g. construct a model in a similar manner that the BINGARCH$_s$ was formulated for seasonal data), estimation of BINAR type models via CML even when having only one lag included is computationally intense since. Considering the above, we will formulate the model by mixing properties of both BINAR and BINGARCH models that would be able to model both first and seasonal lag dependencies without intensifying the estimation too much.

Definition 2.3. *Let* $\mathbf{X}_t = [X_{1,t}, X_{2,t}]', t \in \mathbb{Z}$ *be stationary non-negative integer-valued bivariate time series and* $\mathbf{R}_t = [R_{1,t}, R_{2,t}]', t \in \mathbb{Z}$ *be a non-negative integer-valued bivariate sequence.*

Then the process $\mathbf{X}_t$ is a bivariate time variant seasonal INAR process with seasonal period s (TV-BINAR(1)$_s$), if it satisfies the equation:

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

and $\mathbf{R}_t$ are innovations generated by a Poisson BINGARCH$(1,1)$ process

$$\begin{cases} \mathbf{R}_t|\mathcal{F}_{t-1} \sim BP(\lambda_{1,t}, \lambda_{2,t}, \phi), \\ \boldsymbol{\lambda}_t = \boldsymbol{\delta} + \mathbf{G}_1\boldsymbol{\lambda}_{t-1} + \mathbf{B}_1\mathbf{X}_{t-1} = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix} \begin{bmatrix} \lambda_{1,t-1} \\ \lambda_{2,t-1} \end{bmatrix} + \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix}, \end{cases} \tag{2.31}$$

where $\alpha_i \in [0,1),\ i = 1,2$; $\delta_i > 0,\ i = 1,2$; $\gamma_i \geq 0,\ i = 1,2$; $\beta_i \geq 0,\ i = 1,2$.

It can be noted that this model varies in time in terms of the bivariate Poisson parameters λ_1 and λ_2 of the innovations. In this case the variability is expressed via realization of $\mathbf{X}_{t-1}$. Although we do not set the direct dependency, we assume that innovations carry information about the past. In terms of the example given in Section 2.1.1 that BINAR process can be imagined as a queue of people, in this case we suggest that number of people joining the queue during the time interval $[t-s;t]$ is influenced by how many people were in the queue at the time $t-1$.

Once again, we will assume diagonality of all coefficient matrices and dependence of the time series via parameter ϕ that is time invariant (i.e. remains constant over the time). Worth noting that equation (2.4) holds for TV-BINAR(1)$_s$ as well.

2.3.2 Properties

In this section we will investigate some properties of the TV-BINAR(1)$_s$ model.

Conditional mean:

$$\mathbb{E}(X_{i,t}|X_{i,t-1}) = \alpha_i X_{i,t-s} + \delta_i + \gamma_i\lambda_{i,t-1} + \beta_i X_{i,t-1} + \phi, \quad i = 1,2.$$

Proof:

$$\begin{aligned} \mathbb{E}(X_{i,t}|X_{i,t-1}) &= \mathbb{E}(\alpha_i \circ X_{i,t-s} + R_{i,t}|X_{i,t-1}) = \mathbb{E}(\alpha_i \circ X_{i,t-s}|X_{i,t-1}) + \mathbb{E}(R_{i,t}|X_{i,t-1}) \\ &= \alpha_i X_{i,t-s} + \delta_i + \gamma_i\lambda_{i,t-1} + \beta_i X_{i,t-1} + \phi. \quad \square \end{aligned}$$

Unonditional mean:

$$\mathbb{E}(X_i) = \frac{\delta_i + \phi(1-\gamma_i)}{(1-\alpha_i)(1-\gamma_i)-\beta_i}, \quad i = 1, 2.$$

Proof:

Similarly as for BINAR(1)$_s$ model, we can express an observation as a sum of innovations

$$\mathbb{E}(X_{t,i}) = \mathbb{E}(\sum_{k=0}^{\infty} \alpha_i^k \circ R_{t-ks,i}) = \sum_{k=0}^{\infty} \mathbb{E}(\alpha_i^k \circ R_{t-ks,i}) = \sum_{k=0}^{\infty} \alpha_i^k \mathbb{E} R_{t-ks,i} = \frac{\mathbb{E}(R_{t-ks,i})}{1-\alpha_i}. \quad (2.32)$$

We can also note that:

$$\mathbb{E}(R_{t,i}) = \phi + \delta_i + \gamma_i \mathbb{E}(\lambda_{i,t-1}) + \beta_i \mathbb{E}(X_{i,t-1}). \quad (2.33)$$

Assuming the process is stationary, we can disregard the time indexes in the expectations and substitute (2.33) in (2.32):

$$\mathbb{E}(X_i) = \frac{\phi + \delta_i + \gamma_i \mathbb{E}(\lambda_i) + \beta_i \mathbb{E}(X_i)}{1-\alpha_i}. \quad (2.34)$$

We can also note that:

$$\mathbb{E}(X_i) = \alpha_i \mathbb{E}(X_i) + \mathbb{E}(\lambda_i) + \phi.$$

Hence,

$$\mathbb{E}(\lambda_i) = (1-\alpha_i)\mathbb{E}(X_i) - \phi.$$

Therefore, we can rewrite (2.34) as follows:

$$\mathbb{E}(X_i) = \frac{\phi + \delta_i + \gamma_i((1-\alpha_i)\mathbb{E}(X_i) - \phi) + \beta_i \mathbb{E}(X_i)}{1-\alpha_i}. \quad (2.35)$$

Here, after collecting the terms in (2.35) we finally obtain the result:

$$\mathbb{E}(X_i) = \frac{\delta_i + \phi(1-\gamma_i)}{(1-\alpha_i)(1-\gamma_i)-\beta_i}. \quad \square$$

It can be noted that X_i has a finite mean if $\alpha_i < 1, \ i = 1, 2; \ \gamma_i < 1, \ i = 1, 2, \ \beta_i \neq (1-\alpha_i)(1-\gamma_i), \ i = 1, 2$ and has a positive mean if $\beta_i < (1-\alpha_i)(1-\gamma_i), \ i = 1, 2.$

Covariance:

$$\mathbb{C}\text{ov}(X_{t,i}, X_{t,j}) = \frac{\phi}{1-\alpha_i \alpha_j}, \quad i \neq j.$$

Proof:

Same as in Section 2.1.2

2.3.3 Conditional Maximum Likelihood estimation

TV-BINAR$(1)_s$ model can be estimated similarly as a BINAR model via CML. Here we will also formulate density for each time period t as a convolution of two binomials and a bivariate Poisson distribution. The pmfs of binomials $\mathbf{X}_t$ will be considered as shown in the equations (2.15) and (2.16). The pmf of innovations will be used similarly as shown in the equation (2.1):

$$\begin{aligned} f(k,l) = P(R_{1,t}=k, R_{2,t}=l) = & e^{-(\lambda_{1,t}-\lambda_{2,t}-\phi)} \frac{(\lambda_{1,t}-\phi)^k}{k!} \frac{(\lambda_{2,t}-\phi)^l}{l!} \\ & \times \sum_{m=0}^{\min(k,l)} \binom{k}{m}\binom{l}{m} m! (\frac{\phi}{(\lambda_{1,t}-\phi)(\lambda_{2,t}-\phi)})^m, \end{aligned} \tag{2.36}$$

where $\lambda_{1,t}$ and $\lambda_{2,t}$ from the equation (2.31) are:

$$\boldsymbol{\lambda}_t = \boldsymbol{\delta} + \mathbf{G}_1 \boldsymbol{\lambda}_{t-1} + \mathbf{B}_1 \mathbf{X}_{t-1}. \tag{2.37}$$

Having the above, we can now define a jpmf for TV-BINAR$(1)_s$:

$$\begin{aligned} f(x_{1,t}, x_{2,t} | x_{1,t-s}, x_{2,t-s}, \alpha_1, \alpha_2, \lambda_1, \lambda_2, \phi) = \sum_{k=0}^{g_1}\sum_{l=0}^{g_2} f_1(x_{1,t}-k) f_2(x_{2,t}-l) f_3(k,l) = \\ = \sum_{k=0}^{g_1}\sum_{l=0}^{g_2} e^{-(\lambda_{1,t}-\lambda_{2,t}-\phi)} \frac{(\lambda_{1,t}-\phi)^k}{k!} \frac{(\lambda_{2,t}-\phi)^l}{l!} \sum_{m=0}^{\min(k,l)} \binom{k}{m}\binom{l}{m} m! (\frac{\phi}{(\lambda_{1,t}-\phi)(\lambda_{2,t}-\phi)})^m \\ \times \binom{x_{1,t-s}}{x_{1,t}-k} \alpha_1^{x_{1,t}-k} (1-\alpha_1)^{x_{1,t-s}-x_{1,t}+k} \binom{x_{2,t-s}}{x_{2,t}-l} \alpha_2^{x_{2,t}-l} (1-\alpha_2)^{x_{2,t-s}-x_{2,t}+l}, \end{aligned}$$

where $g_j = \min(x_{j,t}, x_{j,t-s}),\ j = 1,2$ and $\boldsymbol{\lambda}_t$ as defined in the equation (2.37). The conditional log-likelihood then is:

$$\ell(\alpha_1, \alpha_2, \delta_1, \delta_2, \gamma_1, \gamma_2, \beta_1, \beta_2, \phi) = \sum_{t=s+1}^{N} \log f(x_{1,t}, x_{2,t} | x_{1,t-1}, x_{2,t-1}, \alpha_1, \alpha_2, \delta_1, \delta_2, \gamma_1, \gamma_2, \beta_1, \beta_2, \phi).$$

To obtain the CML estimates we maximize:

$$\hat{\theta}^{CLM} = \max_{\alpha_1,\alpha_2,\delta_1,\delta_2,\gamma_1,\gamma_2,\beta_1,\beta_2,\phi} \ell(\alpha_1, \alpha_2, \delta_1, \delta_2, \gamma_1, \gamma_2, \beta_1, \beta_2, \phi).$$

Function *optim* in R has been used for the maximisation problem. Similarly as for estimation of BINGARCH, since CLS estimates were not available, the initial parameter values for maximisation process were selected randomly in a such way that log-likelihood can be estimated at the given values.

2.3.4 Simulation

Similar simulation set up as for the previous models has been used to test how suggested parameter estimation method works. Two sets of 200 independent simulations were created and estimated with two different sample sizes: N $= 100$ and N $= 500$. Seasonal lag has been chosen s $= 12$. Results of Bias and RMSE are presented in Table 3, box plots are available in Figures 5 and 6.

Table 3: Simulation results of TV-BINAR(1)$_{s=12}$

Size	$N = 100$		$N = 500$	
Real parameter	BIAS	RMSE	BIAS	RMSE
α_1=0.4	0.1974	0.2001	0.0032	0.0135
α_2=0.5	0.1326	0.1360	0.0039	0.0134
δ_1=2	-0.7720	0.9320	-0.1638	0.4172
δ_2=3	-0.9394	1.1538	0.0918	0.4860
γ_1=0.2	-0.0847	0.1824	-0.0018	0.1513
γ_2=0.1	0.0180	0.1458	-0.0048	0.1087
β_1=0.15	-0.0540	0.0796	-0.0178	0.0357
β_2=0.2	-0.1039	0.1198	0.0090	0.0319
ϕ=0.1	0.1247	0.3222	0.1257	0.1881

Simulation results of TV-BINAR(1)$_s = 12$ reveal that this model is not estimated as accurately as the previous two. From the box plots in Tables 5 and 6 we see the median values of the estimates visibly differ from the real parameter values for most of the parameters in both sets of simulations. On the other hand we see that having a larger sample visibly reduced Bias and RMSE (e.g. for α_1 Bias dropped from 0.1974 where $N = 100$ to 0.0032 where $N = 500$, RMSE respectively reduced from 0.2001 to 0.0135). Worse accuracy could be explained by the previously mentioned idea that result of the maximisation problem heavily relies on the initial values chosen as local extremes may be found instead of global. Such problem did not occur for the BINGARCH model as the estimation is computationally less intense. For the particular model, estimation is more complex and not having initial CLS values influences the results.

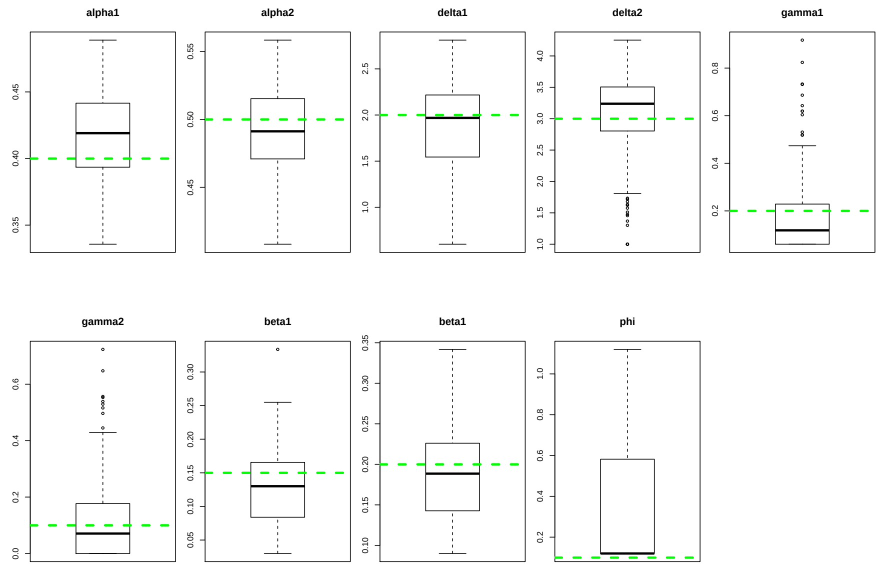

Figure 5: Simulation results of TV-BINAR(1)$_{s=12}$; $N = 100$

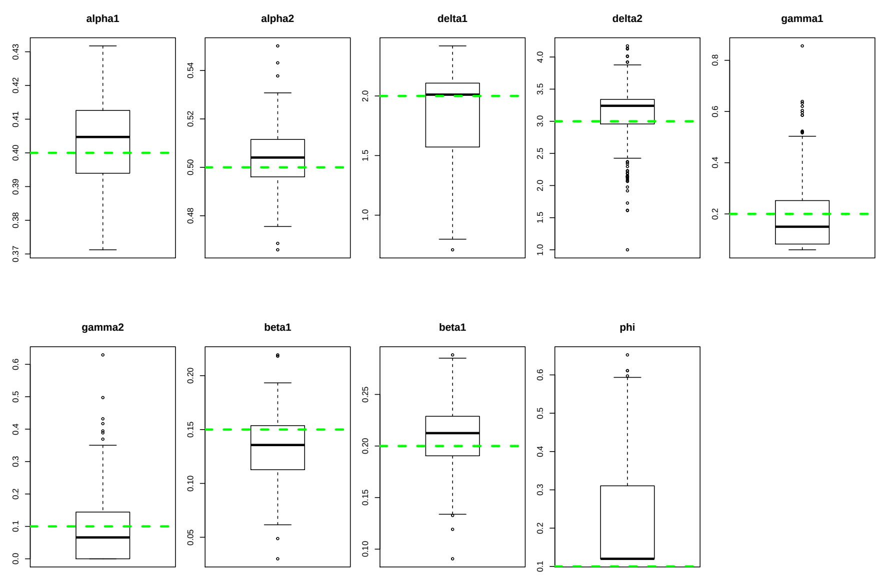

Figure 6: Simulation results of TV-BINAR(1)$_{s=12}$; $N = 500$

3 Evidence on car accident data

3.1 Data

Data used for the empirical part of this thesis is the car accident data from the Official Lithuanian Statistics Portal. The available data has a monthly frequency and ranges from January 2004 to August 2019. Hence, a total of 188 observations are analysed. The data is classified as total number of accidents occurred during a given month and a number of accidents caused by the alcohol intoxicated drivers. For the following application we will consider two time series - a number of accidents caused by the alcohol intoxicated drivers and a number of accidents caused by the sober drivers. The latter time series is derived by deducting a number of accidents caused by the alcohol intoxicated drivers from the total number of accidents. In further analysis, following notation will be used:

- $\mathrm{X}_{1,t}$ - a number of accidents caused by the sober drivers at the time t;
- $\mathrm{X}_{2,t}$ - a number of accidents caused by the alcohol intoxicated drivers at the time t.

Though it is evident that a number of accidents caused by the alcohol intoxicated drivers has an additional factor of causality, both time series also share same contribution factors such as infrastructure, weather conditions etc. Thus, it is reasonable to analyse the two time series as a bivariate process. Visual representation of the data is available in Figure 7. It is interesting to note that the number of accidents during the winter months is the lowest as one could assume vice versa. The lowest number of accidents is usually recorded during February. This is not surprising considering February is the shortest month of the year.

As can be seen from Figure 7, data has a structural break. The change is caused by the new system of documenting the car accidents introduced as of 9 April 2008. Based on the the new regulations, drivers may fill a traffic accident declaration and provide it directly to the insurance company without the involvement of the police, if during the accident:

- no more than two vehicles were involved;
- nobody was injured of perished;

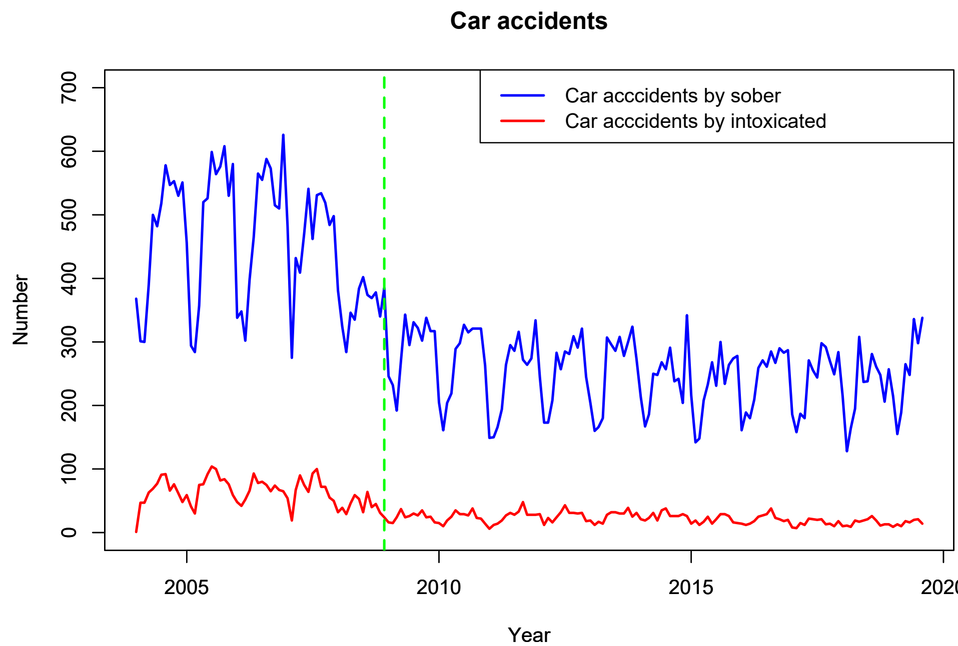

Figure 7: Car incident data: number of car accidents caused by the alcohol intoxicated and sober drivers

- both participants agree on the circumstances;
- both participants were sober.

If such declaration is filled, police does not receive information about the accident. And, hence, such accident is not included in the statistics. Although involvement of police is mandatory were one of the drivers is intoxicated, reduced number of accidents is also visible in the time series of accidents caused by the alcohol intoxicated drivers. Even if such reduction formally should not be supported by the new system, it could be reasoned that the drivers may mutually agree not to report such accidents intentionally or, given low level of intoxication, it may not be visible.

To account for the change in the data, cumulative sum (CUSUM) has been performed on both time series separately test to detect change in mean, as proposed by Horváth et al. (2019). The test is available in R in the package CPAT. Change in mean for the number of accidents caused by sober drivers was detected in the 60^{th} period that of the intoxicated drivers was detected in the 58^{th} period (see Table 4). Considering the results, only observations starting from the 60_{th} period (i.e. December 2008) hve been used for the further analysis leaving the length of data set 129. Dashed green line in Figure 7 represents the point of change.

Table 4: CUSUM test results

CUSUM test			
Variable	Period	Test statistics	p-value
X_1	60	8.3448	$< 2.2e\text{-}16$
X_2	58	8.7932	$< 2.2e\text{-}16$

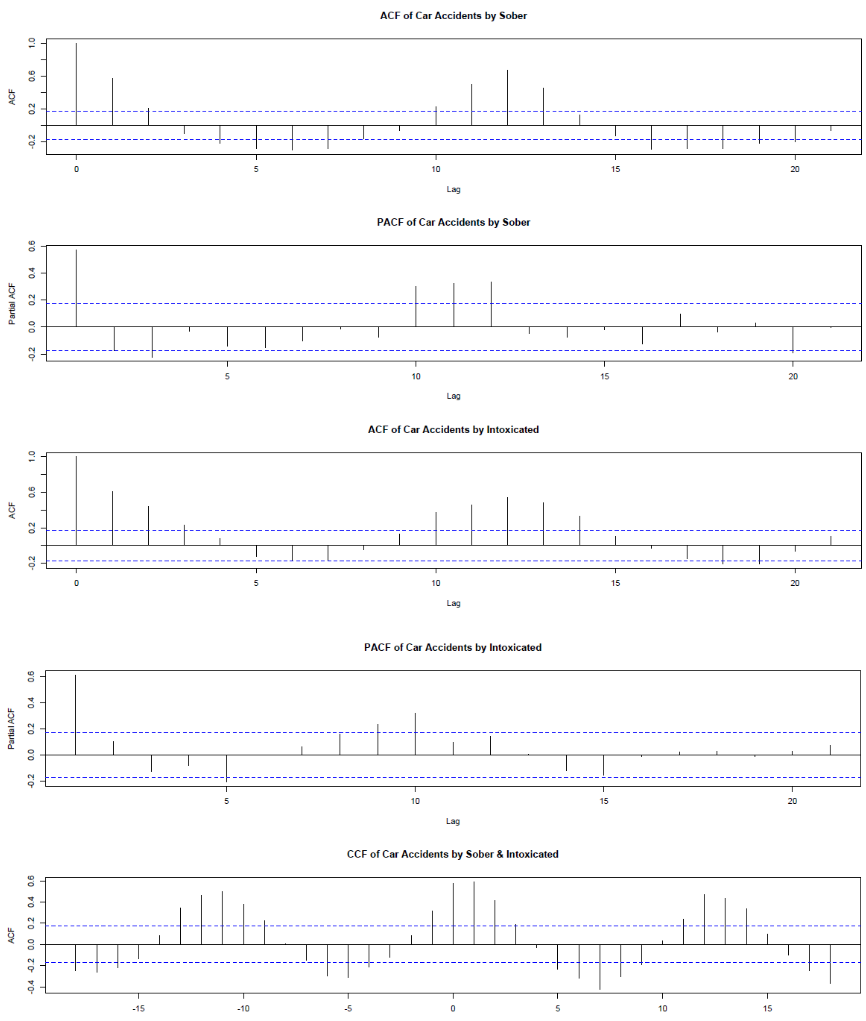

Figure 8: ACF, PACF and CCF of the number of car accidents caused by the alcohol intoxicated and sober drivers

From the autocorrelation and the partial autocorrelation plots of the time series we can note that data exhibits dependence from the 12^{th} lag. Finding is not surprising as data has a monthly frequency and a seasonal dependence visible from Figure 7. Moreover, from the partial autocorrelation plots it is visible that data has a first lag dependency. Cross correlation plot suggests that the data is strongly correlated. Correlation between the two time series is 0.5765. Hence, considering bivariate models is reasonable.

Descriptive statistics of the two time series are provided in Table 5. Mean value of car accidents caused by the sober drivers is 253 per month, variance 3109.1470, values range from 128 to 385. Meanwhile number of car accidents caused by intoxicated ranges from 6 to 48 per month with mean and variance 22 and 69.1165 respectively. It can be noted that both time series exhibits overdispersion with mean values lower than the variance.

Table 5: Descriptive statistics

Variable	Mean	Variance	Min	Max
X_1	253	3109.1470	128	385
X_2	22	69.1165	6	48

ADF and KPSS test were performed to see if processes are stationary (see Table 6). Both tests are available in R. Based on the ADF test results, both time series does not have a unit root (null hypothesis is rejected in both cases having p-value less than 0.01). Result is supported by the KPSS test results - for both time series null hypothesis indicating level stationarity is not rejected having p-value greater than 0.1 (larger p-values are not provided in the output).

Table 6: Tests of stationarity and unit root

	ADF test		KPSS test	
Variable	Test statistics	p-value	Test statistics	p-value
X_1	-7.0058	0.01	0.3689	0.1
X_2	-6.2023	0.01	0.8889	0.1

3.2 Results

Three models - BINAR$(1)_{s=12}$, BINGARCH$_{s=12}$ and TV-BINAR$(1)_{s=12}$ were fitted on the car accident data. The dynamics of the three models can be written as follows:

1. $\mathbf{X}_t = \mathbf{A} \circ \mathbf{X}_{t-12} + \mathbf{R}_t; \quad \mathbf{R}_t \sim BP(\lambda_1, \lambda_2, \phi)$,

2. $$\begin{cases} \mathbf{X}_t | \mathcal{F}_{t-1} \sim BP(\lambda_{1,t}, \lambda_{2,t}, \phi) \\ \boldsymbol{\lambda}_t = \boldsymbol{\delta} + \mathbf{A}_1 \boldsymbol{\lambda}_{t-1} + \mathbf{A}_2 \boldsymbol{\lambda}_{t-12} + \mathbf{B}_1 \mathbf{X}_{t-1} + \mathbf{B}_2 \mathbf{X}_{t-12} \end{cases},$$

3. $\mathbf{X}_t = \mathbf{A} \circ \mathbf{X}_{t-12} + \mathbf{R}_t$,

   where $\mathbf{R}_t$ is generated by a Poisson BINGARCH$(1,1)$ process
   $$\begin{cases} \mathbf{R}_t | \mathcal{F}_{t-1} \sim BP(\lambda_{1,t}, \lambda_{2,t}, \phi), \\ \boldsymbol{\lambda}_t = \boldsymbol{\delta} + \mathbf{G}_1 \boldsymbol{\lambda}_{t-1} + \mathbf{B}_1 \mathbf{X}_{t-1}. \end{cases}$$

BINAR$(1)_{s=12}$ was estimated using CLS and CML methods, BINGARCH$_{s=12}$ and TV-BINAR$(1)_{s=12}$ models were estimated using CML method. Fitted values of all models are plotted in Figures 9 and 10 for car accidents caused by the sober and intoxicated drivers respectively.

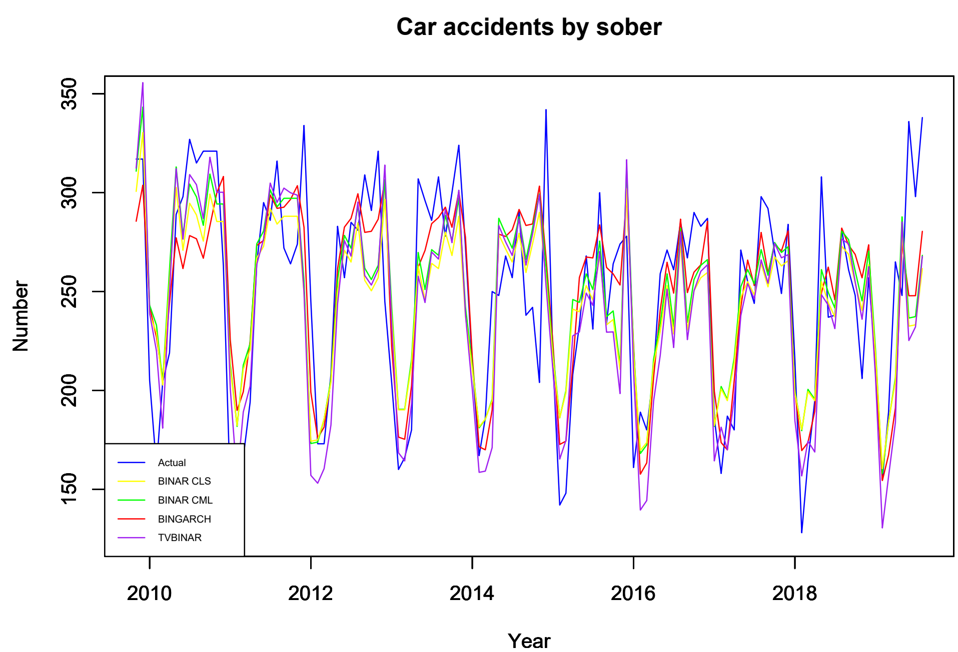

Figure 9: Predictions of the number of car accidents caused by the sober drivers

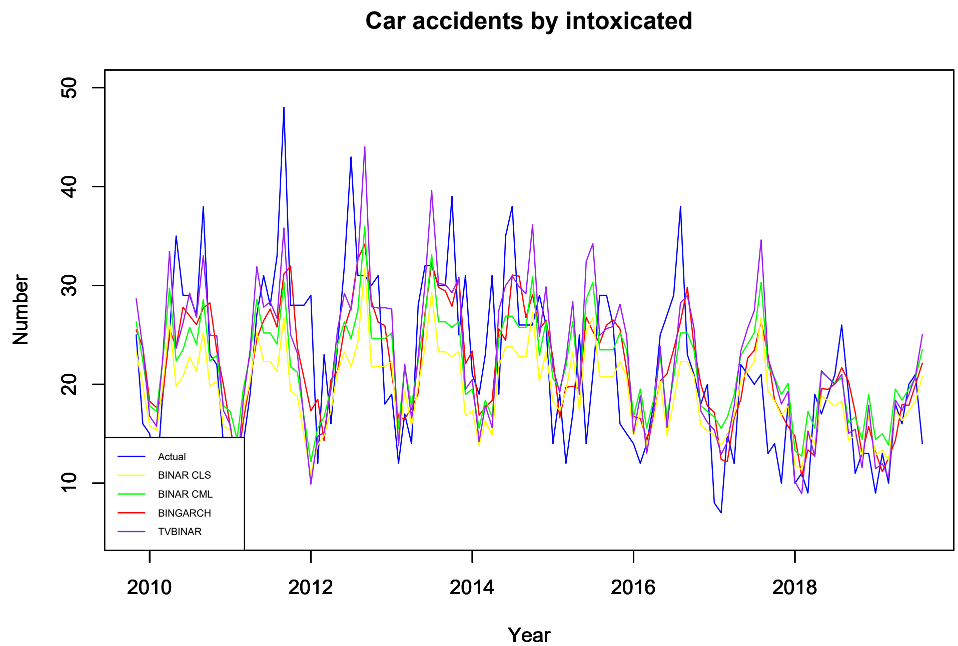

Figure 10: Predictions of the number of car accidents caused by the alcohol intoxicated drivers

From Figures 9 and 10 it is evident that all models managed to capture general movement of data rather well. From visual analysis one can note that TV-BINAR$(1)_{s=12}$ model seems to have the closest fit to the data especially looking in the spikier parts of the data. Though all models seem to have performed rather similarly for the X_1 time series, both BINAR$(1)_{s=12}$ model estimations (CLS and CML) seem to have captured dynamics of the X_2 the least favourably. This is especially noticeable for the CLS predictions that are visibly different from the actual values of the data.

Assumption that TV-BINAR$(1)_{s=12}$ model provided the closest fit to the data is also supported by the AIC score provided in Table 7. TV-BINAR$(1)_{s=12}$ model has the smallest AIC score compared to that of the other two models estimated via CML. This indicates that the model has the smallest out-of-sample prediction error.

Other interesting result is that all estimated models show a relatively similar ϕ value indication that approximately 3-6 car accidents per month are caused by the same factors for both time series (e.g. unusual weather or other environmental conditions). What is more, all three models revealed that car accidents caused by the sober drivers have a higher dependency on its seasonal lag. Hence, has a more expressed survival explanation power. Whereas number of car accidents caused by the

Table 7: Estimation results

	$\text{BINAR(1)}^{\text{CLS}}_{s=12}$	$\text{BINAR(1)}^{\text{CML}}_{s=12}$	$\text{BINGARCH}_{s=12}$	$\text{TV-BINAR(1)}_{s=12}$
α_1	0.7016	0.7211		0.6616
α_2	0.5172	0.5666		0.4972
λ_1	61.8326	60.2403		
λ_2	2.7829	3.3210		
δ_1			0.8389	36.1318
δ_2			0.2154	6.0004
α_{11}			0.1382	
α_{12}			0.0583	
α_{21}			0.3195	
α_{22}			0.1592	
β_{11}			0.0816	0.0675
β_{12}			0.3441	0.0729
β_{21}			0.4429	
β_{22}			0.2971	
γ_1				0.2832
γ_2				0.3158
ϕ	5.8756	5.4394	3.1657	4.1170
Log-likelihood		1802.67	1054.714	414.8945
AIC		3615.34	2131.428	847.789

intoxicated drivers is less predictable and less season dependent; thus, has more random dynamics. Alternatively, number of car accidents caused by the intoxicated drivers showed a higher dependency on period $t-1$ (in $\text{BINGARCH}_{s=12}$ and $\text{TV-BINAR(1)}_{s=12}$ models) indicating a more rigid behaviour. What is more, some interesting insights can be brought from the $\text{TV-BINAR(1)}_{s=12}$ model parameter estimates about the structure of arrival element. From the $\hat{\delta}_1$ and $\hat{\delta}_2$ one can note that respectively 36 and 6 car accidents per month are caused by the sober and intoxicated drivers due to the unusual environmental conditions (constant components of $\lambda_{1,t}$ and $\lambda_{2,t}$ $\forall t$), plus the arrival elements include 28 per cent and 32 per cent of arrival car accidents from the previous month (parameters $\hat{\beta}_{11}$ and $\hat{\beta}_{12}$) and approximately 7 per cent of all car accidents occurred during a previous month (parameters $\hat{\gamma}_1$ and $\hat{\gamma}_2$).

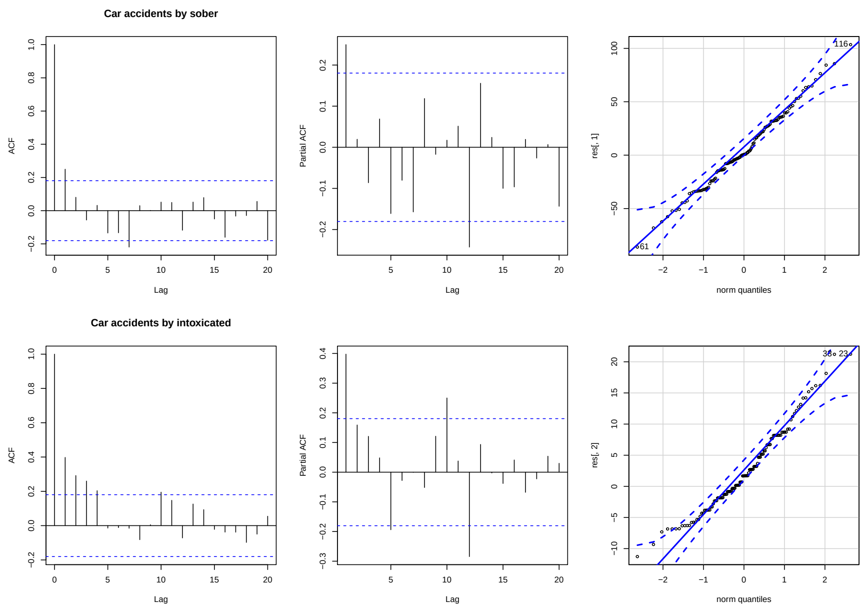

Figure 11: Residual analysis of BINAR(1)$^{\mathrm{CLS}}_{s=12}$ model

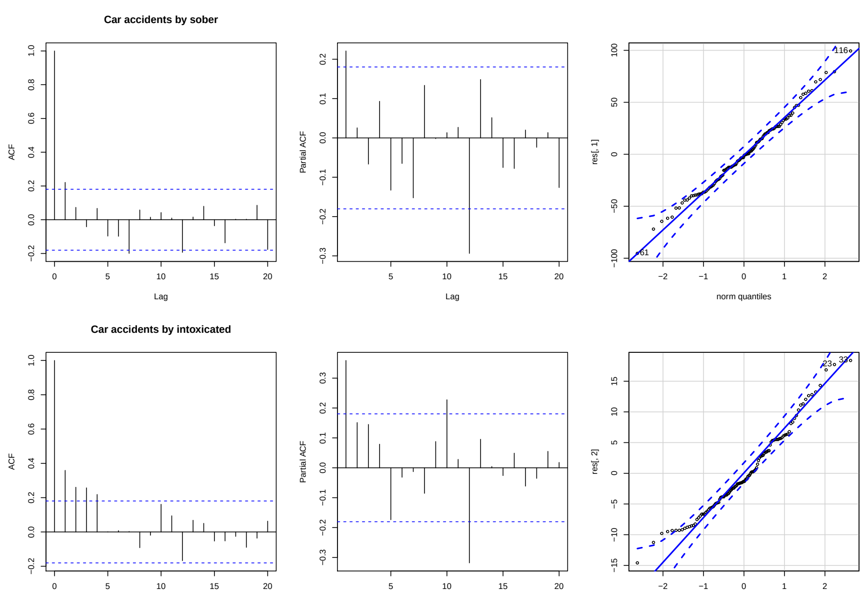

Figure 12: Residual analysis of BINAR(1)$^{\mathrm{CML}}_{s=12}$ model

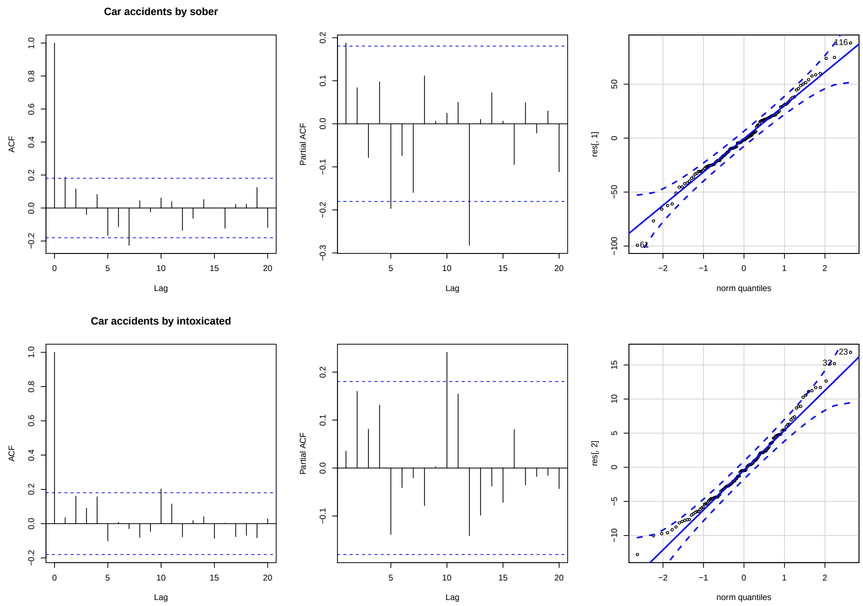

Figure 13: Residual analysis of $\text{BINGARCH}_{s=12}$ model

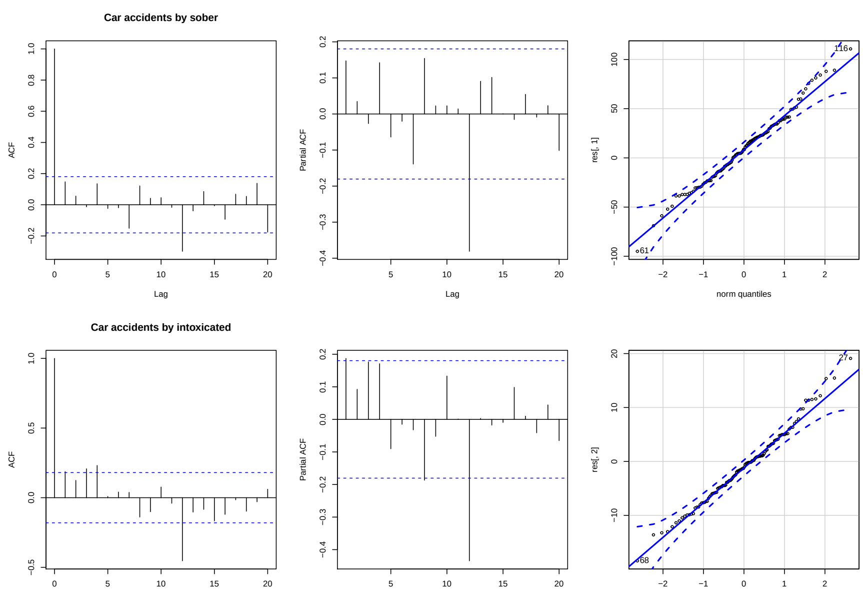

Figure 14: Residual analysis of $\text{TV-BINAR}(1)_{s=12}$ model

From the residual analysis (Figures 11, 12, 13 and 14) it can be noted that none of the three models is capable of describing all dynamics of the data. 12^{th} lags of the residuals are significantly autocorrelated for all estimations. In particular, $\text{BINAR}(1)^{\text{CLS}}_{s=12}$ seems to be the least favourable estimation as the residuals have the most significant autocorrelation lags (see Figure 11). Moreover, they do not seem to fall on a straight line in the Q-Q plot, suggesting that the residuals are not normally distributed and has a number of extreme values. BINAR$(1)^{\text{CML}}_{s=12}$ estimation residuals have less significantly autocorrelated lags (see Figure 12). However, residuals of both time series show a significant 1 lag in the PACF plot. This supports the assumption of limitation of BINAR$(1)_{s=12}$ model in general, that it does not consider significant first order dependency of data, which is the case for the most time series. Looking at the residuals of the BINGARCH$_{s=12}$ estimation (see Figure 13) one can note that model was successful at determining seasonal dependency on the number of car accidents caused by the intoxicated drivers. Finally, from the residuals of the TV-BINAR$(1)_{s=12}$ model, we could support the idea that the model suited data the best (as suggested from the visual analysis and the AIC score) as the residuals have the least number of significant autocorrelation lags (generally, only the 12$^{\text{th}}$ lag is concerning) and the Q-Q plots suggest the most normal distribution of the residuals.

To sum up the above we could suggest that the TV-BINAR$(1)_{s=12}$ model suited data the best. Nevertheless, non of the models captures the actual seasonal dependency properly. This indicates that the dependency differs between the months and 12 different models should rather be estimated (i.e for each month separately). However, this would not be reasonable for the particular dataset as such segregation would leave us with 24 time series, with approximately 11 observations in each. Estimating models on such short time series cannot be accurate and would likely provide false inference.

4 Conclusions and recommendations for future work

In this thesis three models for integer-valued bivariate time series with seasonality were formulated: BINAR$(1)_s$, BINGARCH$_s$ and TV-BINAR$(1)_s$. CML estimation methods were suggested for all three models. In addition, CLS estimation method was formulated for BINAR$(1)_s$ model. Efficiency of the estimators has been tested on simulated data and evaluated using box plots, Bias and RMSE.

Car accident data was used for the application of the models. Bivariate time series consist of number of car accidents caused by the alcohol intoxicated drivers and a number of car accidents caused by the sober drivers in Lithuania per month. Information for period from December 2008 to April 2019 was used for the estimation.

The fitted models revealed that the TV-BINAR$(1)_{s=12}$ provided the best fit for the data having the smallest AIC score. Moreover, residual analysis showed that the model was capable of defining the most information of the data by leaving most of the autocorrelations of the residuals insignificant.

From the parameter estimates it can be concluded that number of car accidents caused by the sober drivers is more predictable than that of the intoxicated drivers as this time series showed stronger seasonal dependency. Moreover, it can be concluded that the data is of a bivariate nature since all three models captured the cross-dependence parameter. In addition, the estimated cross-dependence parameter $\hat{\phi}$ was relatively similar for all models.

Residual analysis has shown that none of the models was able to effectively capture seasonality of the data (12$^{\text{th}}$ autocorrelation lag was significant). This implies that seasonal dependency differs between the months and one model should not be used for the given data. Alternatively, 12 different models should have been estimated. However, considering the length of the analysed dataset, constructing 12 different models would be unreasonable, as it would leave the particular time series very short.

For the future work, different distributional choices could be considered for the models (e.g. negative binomial) to compare the results. Most importantly, it would be necessary to analyse the stability and other properties not considered in the thesis of the TV-BINAR$(1)_s$ model as well as to think of alternative model estimation procedures to improve the accuracy of the estimates.

References

M. A. Al-Osh and A. A. Alzaid. First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, 8(3):261–275, 1987.

M. A. Al-Osh and A. A. Alzaid. Integer-valued moving average (INMA) process. *Statistical Papers*, 29(1):281–300, 1988.

M. Bourguignon, K. L. P. Vasconcellos, V. A. Reisen, and M. Ispány. A Poisson INAR(1) process with a seasonal structure. *Journal of Statistical Computation and Simulation*, 86(2):373–387, 2016.

A. Buteikis and R. Leipus. A copula-based bivariate integer-valued autoregressive process with application. *Modern Stochastics: Theory and Applications*, 6(2):227–249, 2019.

A. Buteikis and R. Leipus. An integer-valued autoregressive process for seasonality. *Journal of Statistical Computation and Simulation*, 90(3):391–411, 2020.

R. Ferland, A. Latour, and D. Oraichi. Integer-valued GARCH process. *Journal of Time Series Analysis*, 27(6):923–942, 2006.

K. Fokianos, A. Rahbek, and D. Tjøstheim. Poisson autoregression. *Journal of the American Statistical Association*, 104(488):1430–1439, 2009.

A. Heinen. Modelling time series count data: an autoregressive conditional Poisson model. *University Library of Munich. MPRA paper, 8113*, 2003.

L. Horváth, C. Miller, and G. Rice. A new class of change point test statistics of Rényi type. *Journal of Business & Economic Statistics*, pages 1–19, 2019.

D. Jin-Guan and L. Yuan. The integer-valued autoregressive (INAR (p)) model. *Journal of Time Series Analysis*, 12(2):129–142, 1991.

H. Liu. *Some models for time series of counts*. PhD thesis, Columbia University, 2012.

X. Pedeli. *Modelling multivariate time series for count data*. PhD thesis, PhD thesis, Athens University of Economis And Business, 2011.

X. Pedeli and D. Karlis. A bivariate INAR(1) process with application. *Statistical Modelling*, 11 (4):325–349, 2011.

X. Pedeli and D. Karlis. On estimation of the bivariate Poisson INAR process. *Communications in Statistics-Simulation and Computation*, 42(3):514–533, 2013.

M. M. Ristić, H. S Bakouch, and A. S. Nastić. A new geometric first-order integer-valued autoregressive (NGINAR (1)) process. *Journal of Statistical Planning and Inference*, 139(7):2218–2226, 2009.

M. M. Ristić, A. S. Nastić, and H. S. Bakouch. Estimation in an integer-valued autoregressive process with negative binomial marginals (NBINAR (1)). *Communications in Statistics-Theory and Methods*, 41(4):606–618, 2012.

M. G. Scotto, C. H. Weiß, and S. Gouveia. Thinning-based models in the analysis of integer-valued time series: A review. *Statistical Modelling*, 15(6):590–618, 2015.

F. W. Steutel and K. Van Harn. Discrete analogues of self-decomposability and stability. *The Annals of Probability*, pages 893–899, 1979.

F. Zhu. A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis*, 32 (1):54–67, 2011.

F. Zhu. Zero-inflated Poisson and negative binomial integer-valued GARCH models. *Journal of Statistical Planning and Inference*, 142(4):826–839, 2012.

A Appendix

R code for calculating BINAR(1)$^{CLS}_{s=12}$ estimates

```
cls_est<-function(y){
rez <- rep(NA,5)
names(rez) <- c("a1","a2","l1","l2","O")
for(i in 1:ncol(y)){
x <- y[,i]
x_a <- rep(NA, (length(x)-12))
for(j in 1:(length(x)-12)){
x_a[j] <- x[j+12]*(x[j]-mean(x))
}
x_b <- rep(NA, (length(x)-12))
for(j in 1:(length(x)-12)){
x_b[j] <- x[j]^2-mean(x)^2
}
a <- sum(x_a)/sum(x_b)
l <- (sum(x[13:length(x)]) - a*sum(x[1:(length(x)-12)]))
    /(length(x)-12)
rez[i] <- a
rez[i+2] <- l
}
O <- (sum((y[13:nrow(y),1]-rez[1]*y[1:(nrow(y)-12),1])
     *(y[13:nrow(y),2]-rez[2]*y[1:(nrow(y)-12),2]))
     -(sum((y[13:nrow(y),1]-rez[1]*y[1:(nrow(y)-12),1]))
     *sum((y[13:nrow(y),2]-rez[2]*y[1:(nrow(y)-12),2])))
     /(length(x)-12))/(length(x)-12)
rez[5] <- O
rez[3] <- rez[3]-rez[5]
rez[4] <- rez[4]-rez[5]
rez
}
```

R codes for calculating log-likelihood.

BINAR(1)$_{s=12}$

```
cml_BP_INAR <- function(vec, y){
a1 <- vec[1]
a2 <- vec[2]
l1 <- vec[3]
l2 <- vec[4]
O <- vec[5]
temp <- rep(0,(nrow(y)-12))
i <- 13
while(i <= nrow(y)){
temp1 <- 0
for(k in 0:min(y[i,1],y[(i-12),1])){
temp2 <- 0
for(s in 0:min(y[i,2],y[(i-12),2])){
temp3 <- dbvpois(k, s, l1, l2, O)
pirmas <- dbinom(x = (y[i,1] - k), size = y[(i-12),1], prob = a1)
antras <- dbinom(x = (y[i,2] - s), size = y[(i-12),2], prob = a2)
temp2 <- temp2 + temp3*pirmas*antras
}
temp1 <- temp1 + temp2
}
temp[i-12] <- ifelse(temp1==0, NA, log(temp1))
i <- i+1
}
temp <- as.vector(na.omit(temp))
return(-sum(temp))
}
```

BINGARCH$_{s=12}$

```
cml_BP_INGARCH <-function(vec, y){
d1 <- vec[1]
d2 <- vec[2]
a11 <- vec[3]
a21 <- vec[4]
a12 <- vec[5]
a22 <- vec[6]
b11 <- vec[7]
b21 <- vec[8]
b12 <- vec[9]
b22 <- vec[10]
O <- vec[11]
temp <- rep(NA,(nrow(y)))
l1 <- rep(0, nrow(y))
l2 <- rep(0, nrow(y))
l1[1:12] <- (d1+O)/(1 - a11 - a12 - b11 - b12)
l2[1:12] <- (d2+O)/(1 - a21 - a22 - b21 - b22)
i <- 13
while (i<= nrow(y)) {
k <- y[(i),1]
s <- y[(i),2]
l1[i] <- (d1 + a11*(l1[i-1]) + a12*(l1[i-12])
         + b11*y[(i-1),1] + b12*y[(i-12),1])
l2[i] <- (d2 + a21*(l2[i-1]) + a22*(l2[i-12])
         + b21*y[(i-1),2] + b22*y[(i-12),2])
temp[i-12] <- log(dbvpois(k,s,l1[i],l2[i],O))
i<-i+1
}
temp <- as.vector(na.omit(temp))
return(-sum(temp))
}
```

TV-BINAR(1)$_{s=12}$

```
cml_BP_TV_INAR <- function(vec, y){
a1 <- vec[1]
a2 <- vec[2]
d1 <- vec[3]
d2 <- vec[4]
g1 <- vec[5]
g2 <- vec[6]
b1 <- vec[7]
b2 <- vec[8]
O <- vec[9]
l1 <- d1/(1 - g1 - b1)
l2 <- d2/(1 - g2 - b2)
temp <- rep(0,nrow(y))
i <- 13
while(i <= nrow(y)){
temp1 <- 0
for(k in 0:min(y[i,1],y[(i-12),1])){
temp2 <- 0
for(s in 0:min(y[i,2],y[(i-12),2])){
temp3 <- dbvpois(k, s, l1, l2, O)
pirmas <- dbinom(x = (y[i,1] - k), size = y[(i-12),1], prob = a1)
antras <- dbinom(x = (y[i,2] - s), size = y[(i-12),2], prob = a2)
temp2 <- temp2 + temp3*pirmas*antras
}
temp1 <- temp1 + temp2
}
temp[i-12] <- ifelse(temp1==0, NA, log(temp1))
l1 <- d1 + g1*l1 + b1*y[i,1]
l2 <- d2 + g2*l2 + b2*y[i,2]
i <- i+1
}
```

```
temp <- as.vector(na.omit(temp))
return(-sum(temp))
}
```